Washington University School of Medicine
Digital Commons@Becker

2020-Current year OA Pubs | Open Access Publications

5-1-2023

Belimumab versus anifrolumab in adults with systemic lupus erythematosus: An indirect comparison of clinical response at 52 weeks

Binod Neupane
Evidera

Pragya Shukla
Evidera

Mahmoud Slim
University of Granada

Amber Martin
Evidera

Michelle Petri
Johns Hopkins University School of Medicine

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/oa_4

 Part of the Medicine and Health Sciences Commons
Please let us know how this document benefits you.

Recommended Citation

Neupane, Binod; Shukla, Pragya; Slim, Mahmoud; Martin, Amber; Petri, Michelle; Bertsias, George K.; Kim, Alfred H J; Fanouriakis, Antonis; Levy, Roger A.; Chauhan, Deven; and Ballew, Nick, "Belimumab versus anifrolumab in adults with systemic lupus erythematosus: An indirect comparison of clinical response at 52 weeks." Lupus Science & Medicine. 10, 1. e000907 (2023).
https://digitalcommons.wustl.edu/oa_4/1825

This Open Access Publication is brought to you for free and open access by the Open Access Publications at Digital Commons@Becker. It has been accepted for inclusion in 2020-Current year OA Pubs by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.

Authors

Binod Neupane, Pragya Shukla, Mahmoud Slim, Amber Martin, Michelle Petri, George K. Bertsias, Alfred H J Kim, Antonis Fanouriakis, Roger A. Levy, Deven Chauhan, and Nick Ballew

This open access publication is available at Digital Commons@Becker: https://digitalcommons.wustl.edu/oa_4/1825

LUPUS SCIENCE & MEDICINE™

Belimumab versus anifrolumab in adults with systemic lupus erythematosus: an indirect comparison of clinical response at 52 weeks

Binod Neupane,[1] Pragya Shukla,[1] Mahmoud Slim,[1,2] Amber Martin,[3] Michelle Petri ,[4] George K Bertsias ,[5] Alfred H J Kim,[6] Antonis Fanouriakis ,[7] Roger A Levy ,[8] Deven Chauhan,[9] Nick Ballew [10]

To cite: Neupane B, Shukla P, Slim M, *et al*. Belimumab versus anifrolumab in adults with systemic lupus erythematosus: an indirect comparison of clinical response at 52 weeks. *Lupus Science & Medicine* 2023;**10**:e000907. doi:10.1136/lupus-2023-000907

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/lupus-2023-000907).

Received 27 January 2023
Accepted 6 April 2023

Check for updates

© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Dr Nick Ballew; nick.g.ballew@gsk.com

ABSTRACT

Objective To generate comparative efficacy evidence of belimumab versus anifrolumab in SLE that can inform treatment practices.

Methods The SLE Responder Index (SRI)-4 response at 52 weeks of belimumab versus anifrolumab was evaluated with an indirect treatment comparison. The evidence base consisted of randomised trials that were compiled through a systemic literature review. A feasibility assessment was performed to comprehensively compare the eligible trials and to determine the most appropriate indirect treatment comparison analysis method. A multilevel network meta-regression (ML-NMR) was implemented that adjusted for differences across trials in four baseline characteristics: SLE Disease Activity Index-2K, anti–double-stranded DNA antibody positive, low complement (C)3 and low C4. Additional analyses were conducted to explore if the results were robust to different sets of baseline characteristics included for adjustment, alternative adjustment methods and changes to the trials included in the evidence base.

Results The ML-NMR included eight trials: five belimumab trials (BLISS-52, BLISS-76, NEA, BLISS-SC, EMBRACE) and three anifrolumab trials (MUSE, TULIP-1, TULIP-2). Belimumab and anifrolumab were comparable in terms of SRI-4 response (OR (95% credible interval), 1.04 (0.74–1.45)), with the direction of the point estimate slightly favouring belimumab. Belimumab had a 0.58 probability of being the more effective treatment. The results were highly consistent across all analysis scenarios.

Conclusions Our results suggest that the SRI-4 response of belimumab and anifrolumab are similar at 52 weeks in the general SLE population, but the level of uncertainty around the point estimate means we cannot rule out the possibility of a clinically meaningful benefit for either treatment. It remains to be seen if specific groups of patients could derive a greater benefit from anifrolumab or from belimumab, and there is certainly an unmet need to identify robust predictors towards more personalised selection of available biological agents in SLE.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Belimumab and anifrolumab are both approved treatments for SLE; the efficacy of each treatment has been demonstrated versus placebo in clinical trials.

⇒ The only results to date on the efficacy of belimumab versus anifrolumab are from a single study that indirectly compared the two treatments, and the study had several limitations.

WHAT THIS STUDY ADDS

⇒ The clinical response of belimumab and anifrolumab at week 52 is generally comparable, and belimumab has a 0.58 probability of being the more effective treatment.

⇒ Our results clearly demonstrate that, despite a recent publication to the contrary, there is no evidence to indicate that patients with SLE would benefit from a change in treatment from belimumab to anifrolumab or vice versa.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our results are a valuable reminder for future research that when population-adjusted indirect comparisons are conducted, the patient-level data informing the population adjustment must be large enough and broad enough that the population-adjusted treatment effects can be accurately estimated.

INTRODUCTION

SLE is characterised by chronic inflammation leading to significant morbidity and mortality.[1] Treatment of SLE aims to minimise disease activity, decrease the incidence of disease flares and prevent organ damage.[2] Conventional treatment options include antimalarials, glucocorticoids and immunosuppressive agents.[2 3] Although such treatments can be initially successful, patients often require

Lupus Sci Med: first published as 10.1136/lupus-2023-000907 on 5 May 2023. Downloaded from http://lupus.bmj.com/ on June 6, 2023 at Washington University School of Medicine Library &. Protected by copyright.

adjunctive therapies or a switch to different immunosuppressives, including biologic drugs.[2]

Belimumab, a human immunoglobulin G1λ (IgG1) monoclonal antibody, inhibits the biologic activity of B-lymphocyte stimulating protein.[4] Belimumab was first approved in 2011 by the US Food and Drug Administration for patients with active, autoantibody positive SLE receiving standard therapy (ST), and is now approved for the treatment of patients ≥5 years of age with SLE in >75 countries.[4 5] Patients treated with belimumab plus ST have consistently demonstrated a reduction in disease activity, glucocorticoid use and frequency of flares versus placebo plus ST in randomised controlled trials (RCTs).[6–9]

Anifrolumab, a fully human IgG1K monoclonal antibody that binds to type I interferon (IFN) receptor subunit 1 and inhibits signalling by all type I IFNs, was approved in the USA in 2021 for the treatment of patients with moderate-to-severe SLE receiving ST.[10 11] Anifrolumab received approval based on evidence across three RCTs, two of which (MUSE[12] and TULIP-2[11]) showed favourable results versus placebo plus ST, while the primary efficacy endpoint of SLE Responder Index-4 (SRI-4) was not met in the TULIP-1 trial.[13]

In the absence of a head-to-head RCT comparing belimumab and anifrolumab, an indirect treatment comparison (ITC) that incorporates results across the available RCTs can generate robust comparative evidence to inform treatment practices. An ITC across RCTs can produce valid evidence when there are no differences across trials in effect modifiers (EMs), or when differences in EMs are appropriately accounted for.[14–17] EMs are characteristics that alter the relative effect of a treatment, so that it is more or less effective than an alternative treatment, depending on the level of the EM (further information on ITCs and EMs provided in online supplemental appendix 1). An ITC that adjusts for differences across trials in EMs is referred to as a population-adjusted indirect comparison (PAIC). See online supplemental appendix 2 for more details on PAICs. One PAIC comparing belimumab and anifrolumab has been published;[18] however, the study did not meet the fundamental requirements for a robust population-adjusted analysis.[19] Several studies[20–22] have demonstrated that PAIC methods can perform poorly and yield inaccurate estimates under scenarios similar to that of the Bruce *et al* study.[18] See Ballew *et al* (and the Discussion section) for further details on the limitations of the Bruce *et al* study.[18 19]

The primary objective of our study was to generate evidence on the comparative efficacy of the approved doses of belimumab versus anifrolumab at 52 weeks. A secondary objective was to examine the validity of the findings reported in the Bruce *et al* study.[18]

METHODS

Compiling and assessing the evidence base

A systematic literature review (SLR) (adhering to the Cochrane Collaboration and Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines) was conducted to identify all trials reporting relevant outcomes at 52 weeks for adults (≥18 years) with SLE receiving belimumab or anifrolumab plus ST, published as of 12 April 2022.[23 24] A standardised data extraction template was used to capture all relevant information from the included trials (detailed information on the SLR can be found in online supplemental appendix 3).

We identified EMs that would need to be balanced to conduct an unbiased ITC based on clinical knowledge, published literature, an evaluation of reported subgroup results within individual trials, exploratory analyses of individual patient data (IPD) in belimumab trials that GSK had on file and several rounds of discussion with lupus experts. When multiple options were available for how to adjust (eg, adjust for proportion with any glucocorticoid use or proportion with a glucocorticoid dosage threshold), we relied on the exploratory regression analyses to inform our decisions (online supplemental appendix 4).

The evidence base compiled from the SLR was compared in terms of study design, trial circumstances, patient population, treatment implementation and outcome definitions (see online supplemental appendix 5 for details on each trial). Special focus was paid to the comparison of baseline characteristics across trials that were identified as potential EMs.[25] Where differences across trials were identified, the expected direction and magnitude of the potential bias was noted, as well as when data limitations precluded a thorough comparison or appropriate adjustment.

Outcomes for analysis

The primary efficacy outcome was proportion of patients achieving SRI-4 response at week 52. In the earlier belimumab trials (BLISS-52, BLISS-76, BLISS-SC and the North East Asia study (NEA)),[6–9] SRI-4 incorporated Safety of Estrogens in Lupus National Assessment–SLE Disease Activity Index (SELENA-SLEDAI) in the original definition. SRI-4 has since been re-analysed in these trials, incorporating modified scoring for proteinuria adapted from SLEDAI-2K. In EMBRACE, SRI-4 was reported both ways.[26] In the anifrolumab trials, SRI-4 incorporated SLEDAI-2K.[11–13] Thus, in an effort to make as close of a like-for-like comparison as possible, the SRI-4 results that incorporated a modified version of SLEDAI-2K were used from the belimumab trials. However, there were unresolved differences between the measures regarding joint scoring that could not be addressed, as SELENA-SLEDAI requires three joints, but SLEDAI-2K just two joints. Further, there were also differences in SRI-4 in terms of the British Isles Lupus Assessment Group (BILAG) instrument used in the trials (BILAG classic in belimumab trials and BILAG 2004 in anifrolumab trials) that could not be reconciled.

Several other efficacy outcomes were considered for analysis at 52 weeks, including proportion of patients with ≥4-point reduction in SLEDAI, SLEDAI response on specific organ domains, flares, glucocorticoid reduction and anti–double-stranded DNA antibody (anti-dsDNA)

Lupus Sci Med: first published as 10.1136/lupus-2023-000907 on 5 May 2023. Downloaded from http://lupus.bmj.com/ on June 6, 2023 at Washington University School of Medicine Library & Protected by copyright.

levels. However, robust analyses were not feasible with the evidence base identified (explained in detail in online supplemental appendix 5). Of note, for the proportion of patients achieving ≥4-point reduction in SLEDAI, MUSE was the only anifrolumab trial identified from the SLR that reported this outcome.[12] However, the 4-point reduction in MUSE was based on the clinical components only without consideration of the immunological components. After the SLR was completed, a pooled analysis of TULIP-1 and TULIP-2 ≥4-point reduction in SLEDAI was reported.[18] To ensure the credibility of an ITC, results need to be available for each trial separately (ideally, results are available in terms of the proportion of responders for each arm). Thus, given these data limitations and their potential impact on the credibility of an ITC for ≥4-point reduction in SLEDAI, an ITC of this outcome could not be conducted to provide credible evidence on the comparative efficacy of belimumab versus anifrolumab at 52 weeks. However, exploratory analyses of ≥4-point reduction in SLEDAI were conducted (anifrolumab trial data from the ITC of Bruce *et al*) to aid our understanding of the results of Bruce *et al* for this outcome.[18]

Statistical analysis

Analysis scenarios

The feasibility assessment revealed that there were differences in EMs for SRI-4 between the belimumab and anifrolumab studies, meaning PAIC methods would be required to conduct an unbiased ITC of SRI-4. For the primary outcome of SRI-4, a fixed effects (FE) multilevel network meta-regression (ML-NMR) model that adjusted for all possible 'imbalanced EMs', but no prognostic variables, was selected as the base-case. The rationale for this decision was that the model was capable of adjusting for any meaningful bias introduced by EMs, without making any sacrifices in terms of simplifying the network structure. Adjusting for all possible EMs (specifically Black African ancestry) would have required pooling some belimumab trials together and treating them as a single trial. Four separate sensitivity analyses were conducted for SRI-4 to assess the robustness of our results to alternative sets of variables for adjustment and alternative PAIC methods (simulated treatment comparison (STC) and matching adjusted indirect comparison (MAIC)) methods. Of note, ML-NMR and STC are similar in that, in addition to including EMs in the model (to account for bias), prognostic variables can also be included (no interaction with treatment) to obtain more precise estimates. This is in contrast to MAIC, where only EMs should be included.

Additional sensitivity analyses were also conducted to understand the impact of using the modified SRI-4 definition for the belimumab trials (sensitivity analysis used the SRI-4 definition for the belimumab trials based on SELENA-SLEDAI) and to understand the impact of treating belimumab intravenous and subcutaneous formulations as equivalent treatments (sensitivity with altered network structure, so belimumab intravenous and subcutaneous were individual treatment nodes and compared with anifrolumab separately). Additional details on the analyses can be found in online supplemental appendix 6. Of note, Sensitivity 3 (described in online supplemental appendix 6) was the preplanned base-case analysis but was moved to a sensitivity because it required treating the five available belimumab trials as three trials (BLISS-52 and BLISS-76 were pooled and the NEA study and EMBRACE were pooled). Relatedly, STC and MAIC methods suffer from a similar limitation in that they can only be applied to simple networks of evidence, and as a result, we had to pool all belimumab trials together and all anifrolumab trials together in STC and MAIC analyses.

Exploratory analyses were conducted to emulate the approach implemented in Bruce *et al* for the clinical response outcomes SRI-4 and ≥4-point reduction in SLEDAI.[18] Specifically, this meant conducting the analyses with the same evidence base and with the same results for the anifrolumab trials as reported in Bruce *et al*,[18] which in some cases differed from the results previously published for the trials. We also used the same methods (STC and MAIC), network structure and set of EMs as in Bruce *et al*.[18] All exploratory analyses were undertaken as in Bruce *et al*,[18] with the original SRI-4 and ≥4-point reduction in SLEDAI definitions for the belimumab trials based on SELENA-SLEDAI and the SRI-4 definition for the anifrolumab trials incorporating SLEDAI-2K. However, IPD from the belimumab trials were used to inform the population adjustments, instead of IPD from the anifrolumab trials as in Bruce *et al*.[18] Importantly, the IPD from the belimumab trials is a larger sample than that from the anifrolumab trials (1125 vs 710) and is representative of a broader SLE population (includes patients with and without BILAG ≥1 A or ≥2 B at baseline).

Model implementation

The steps in ML-NMR include deriving the aggregate level likelihood and then deriving the integral in the aggregate model. Deriving the aggregate-level model in ML-NMR requires using IPD from the trials to inform the covariate distributions and correlation structure of variables from the studies. While IPD was available (and used) for the belimumab trials, the IPD for the anifrolumab trials was not. Thus, the observed distributions and correlations from the belimumab trials were used to inform the distributions and correlations in the anifrolumab trials. The FE model used a non-informative normal prior distribution (location=0, scale=100) on each parameter of interest. Three chains (7000 iterations, out of which the first 4000 were the burn-in iterations) were run on each ML-NMR. A random effects (RE) ML-NMR (half-normal (location=0, scale=0.5) prior distribution for the between-study SD) was also conducted for each FE ML-NMR as a check for residual heterogeneity remaining after adjusting for the selected EMs. ML-NMR was implemented in a Bayesian framework by using Markov chain

Lupus Sci Med: first published as 10.1136/lupus-2023-000907 on 5 May 2023. Downloaded from http://lupus.bmj.com/ on June 6, 2023 at Washington University School of Medicine Library & Protected by copyright.

Monte Carlo sampling and with the 'multinma' package in R.[27] Median ORs and 95% credible intervals (CrI) were reported. Treatment-rank probabilities were produced, as well as surface under the cumulative ranking curve (SUCRA) values. The relative effects, ranking probabilities and SUCRA values were estimated for each study population of interest (each individual trial population included in the network, as well as the combined anifrolumab and belimumab populations).

As noted above, for our MAIC and STC analyses, we had to pool all belimumab trials together and all anifrolumab trials together and treat them as two large pseudo-trials. The MAIC and STC analyses were then conducted following the methods described by Signorovitch *et al* and National Institute for Health and Care Excellence guidelines.[28 29] See online supplemental appendix 2 for full model implementation details.

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting or dissemination plans of our research.

RESULTS

Clinical response of belimumab versus anifrolumab at 52 weeks

Evidence base

Nineteen unique trials were identified by the SLR. The detailed findings of the SLR and feasibility assessment are included in online supplemental appendix 5. Eight of the 19 trials that were identified were ultimately eligible for the SRI-4 analysis at 52 weeks comparing the approved doses of belimumab (10 mg/kg intravenous and 200 mg subcutaneous) and anifrolumab (300 mg intravenous): BLISS-52 (NCT00424476); BLISS-76 (NCT00410384); BLISS-SC (NCT01484496); NEA study (NCT01345253); EMBRACE (NCT01632241); TULIP-1 (NCT02446912); TULIP-2 (NCT02446899); MUSE (NCT01438489).[6–9 11–13 26] The trial-level SRI-4 results are presented in figure 1. More detailed information on the inclusion criteria, intervention, baseline characteristics and outcome definitions for these trials is included in online supplemental appendix 5.

ITC SRI-4 at 52 weeks

Eight characteristics were identified as likely EMs for SRI-4 (table 1). Accordingly, the trials would need to be balanced in terms of these characteristics to conduct an unbiased ITC. However, data limitations precluded the possibility of evaluating (and potentially adjusting) the level of balance for two of the variables; body mass index (BMI) was not available in MUSE and none of the trials reported smoking status. Thus, it was possible to adjust for six (SLEDAI-2K, Black African ancestry, low C3, low C4, anti-dsDNA and any glucocorticoid use) of the potential eight EMs. For two (Black African ancestry and any glucocorticoid use) of these six EMs, the level of imbalance was negligible (table 1). Of the remaining four, if no population adjustment was made, one of the variables would be expected to introduce bias in favour of anifrolumab (SLEDAI-2K) and three would be expected to introduce bias in favour of belimumab (low C3, low C4, anti-dsDNA).

In the base-case ML-NMR analysis of the SRI-4 outcome that adjusted for the four imbalanced EMs, belimumab and anifrolumab were generally comparable, with the direction of the point estimate slightly favouring belimumab (OR (95% CrI) 1.04 (0.74–1.45)). There was a 0.58 probability that belimumab was the more effective treatment and a 0.42 respective probability for anifrolumab. Of note, while the model predictions were in line with the observed SRI-4 results in the belimumab trials, the

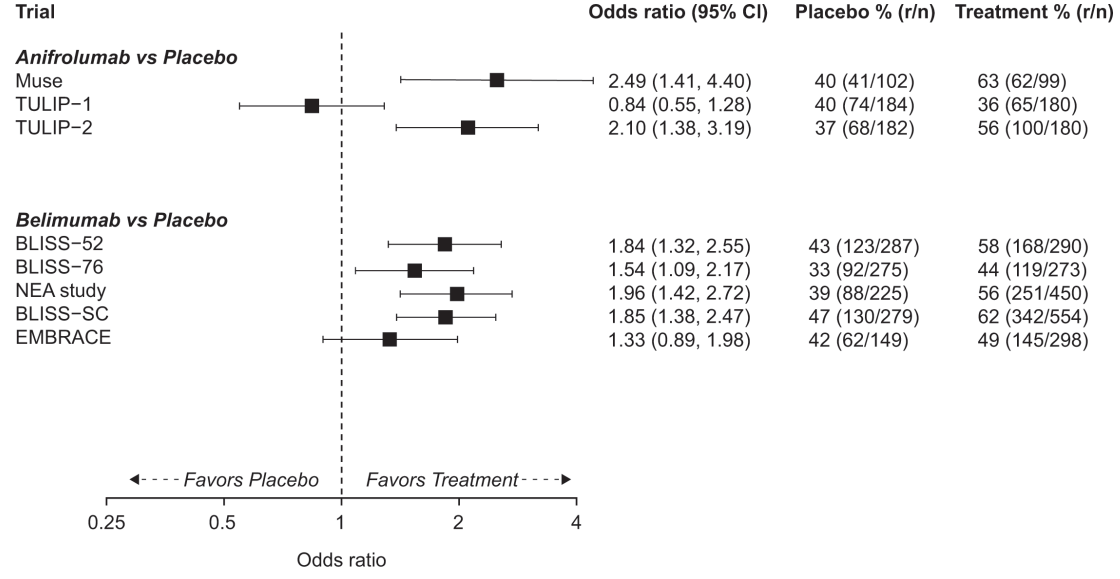


Figure 1 Trial level results that contributed to the ITC for SRI-4 at 52 weeks. ITC, indirect treatment comparison; n, sample size; r, number of responders; SRI-4, SLE Responder Index-4.

Lupus Sci Med: first published as 10.1136/lupus-2023-000907 on 5 May 2023. Downloaded from http://lupus.bmj.com/ on June 6, 2023 at Washington University School of Medicine Library & Protected by copyright.

Table 1 Potential treatment EMs for SRI-4: characteristics that need to be balanced across trials

Baseline characteristic	Expected effect-modifying relationship (relative to placebo) with outcome	Anticipated bias (direction and magnitude) for indirect comparison*		Rationale for anticipated bias (values are belimumab trials vs anifrolumab trials)	Adjustment in model
SLEDAI-2K	Treatment effect is larger in populations with higher SLEDAI-2K values		Moderate/large bias in favour of anifrolumab	Strong treatment EM with moderate difference (mean of 10.4 vs 11.3)	Balance mean SLEDAI-2K across trials
Race	Treatment effect is smaller in Black African ancestry race than other races		Negligible bias in favour of anifrolumab	Moderate treatment EM with small difference (20% vs 14% Black African Ancestry)	Not adjusted in the base-case analysis†
C3	Treatment effect is larger among patients with low C3 concentration		Moderate bias in favour of belimumab	Moderate treatment EM with moderate difference (49% vs 36% low C3)	Balance proportion with low C3 across trials
C4	Treatment effect is larger among patients with low C4 concentration		Moderate bias in favour of belimumab	Moderate treatment EM with moderate difference (37% vs 23% low C4)	Balance proportion with low C4 across trials
Anti-dsDNA	Treatment effect is larger among anti-dsDNA positive patients		Small bias in favour of belimumab	Small treatment EM with large difference (71% vs 40% positive)	Balance proportion anti-dsDNA positive‡
Glucocorticoid use	Treatment effect is larger among patients with any glucocorticoid use		Negligible bias in favour of belimumab	Moderate treatment EM with small difference (88% vs 82% with any use)	Not adjusted in the base-case analysis†
Smoking status	Treatment effect is smaller in smokers	?	Bias for characteristic is unknown	Smoking status is not reported for any trial in the evidence base	None
BMI	Treatment effect is smaller in patients with high BMI	?	Bias for characteristic is unknown	BMI is only available for some trials (but appears similar across trials that report it)§	None

*Arrow thickness indicates the strength of the effect.
†Characteristic was adjusted for in sensitivity analyses.
‡Positivity in belimumab trials defined based on 30 IU/mL threshold, while positivity in anifrolumab trials was defined based on 15 IU/mL threshold. While it was possible to alter the definition in the belimumab trials to match the definition in the anifrolumab trials, a 15 IU/mL threshold was not clinically meaningful for the belimumab trials.
§BMI of 27.6 kg/m^2 for pooled TULIP trials (reported in Bruce *et al*[18]) and BMI of 25.4 kg/m^2 for pooled belimumab trials.
Anti-dsDNA, anti–double-stranded DNA antibody; BMI, body mass index; C3/C4, complement component 3/4; EM, effect modifier; SLEDAI-2K, SLE Disease Activity Index 2000; SRI-4, SLE Responder Index-4.

predictions for the anifrolumab trials did not follow the observed study-level SRI-4 results for the three anifrolumab trials (based on visual comparison of observed trial-level results in figure 1 and model predictions in figure 2). To this point, the deviance information criterion from the RE model was only marginally lower than the base-case FE model (4076 vs 4078), indicating similar model fit. However, the estimate for the heterogeneity parameter was relatively large (tau=0.26) and was accompanied by a relatively large amount of uncertainty (SD of tau=0.15).

The ORs of belimumab versus anifrolumab were highly consistent between the base-case and all sensitivity analyses (sensitivity analyses that employed alternative sets of variables for adjustment and alternative PAIC methods in figure 3; additional analysis results can be found in online supplemental appendix 6). The base-case and sensitivity analysis results were also in line with the results of the standard FE Bayesian network meta-analysis (NMA; OR (95% CrI) 1.13 (0.83–1.53)). Convergence to the posterior distribution was achieved in all Bayesian (NMA and ML-NMR) analyses.

Emulating the approach of Bruce *et al*

The results obtained for SRI-4 when emulating the Bruce *et al*[18] approach suggested that belimumab and anifrolumab were generally comparable, with the direction of the point estimate slightly favouring belimumab (STC OR (95% CI) 1.06 (0.65 to 1.72); MAIC OR (95% CI) 1.11 (0.66 to 1.86)).

The results from the two exploratory analyses with ≥4-point reduction in SLEDAI also suggested that belimumab and anifrolumab were generally comparable, with

Lupus Sci Med: first published as 10.1136/lupus-2023-000907 on 5 May 2023. Downloaded from http://lupus.bmj.com/ on June 6, 2023 at Washington University School of Medicine Library & Protected by copyright.

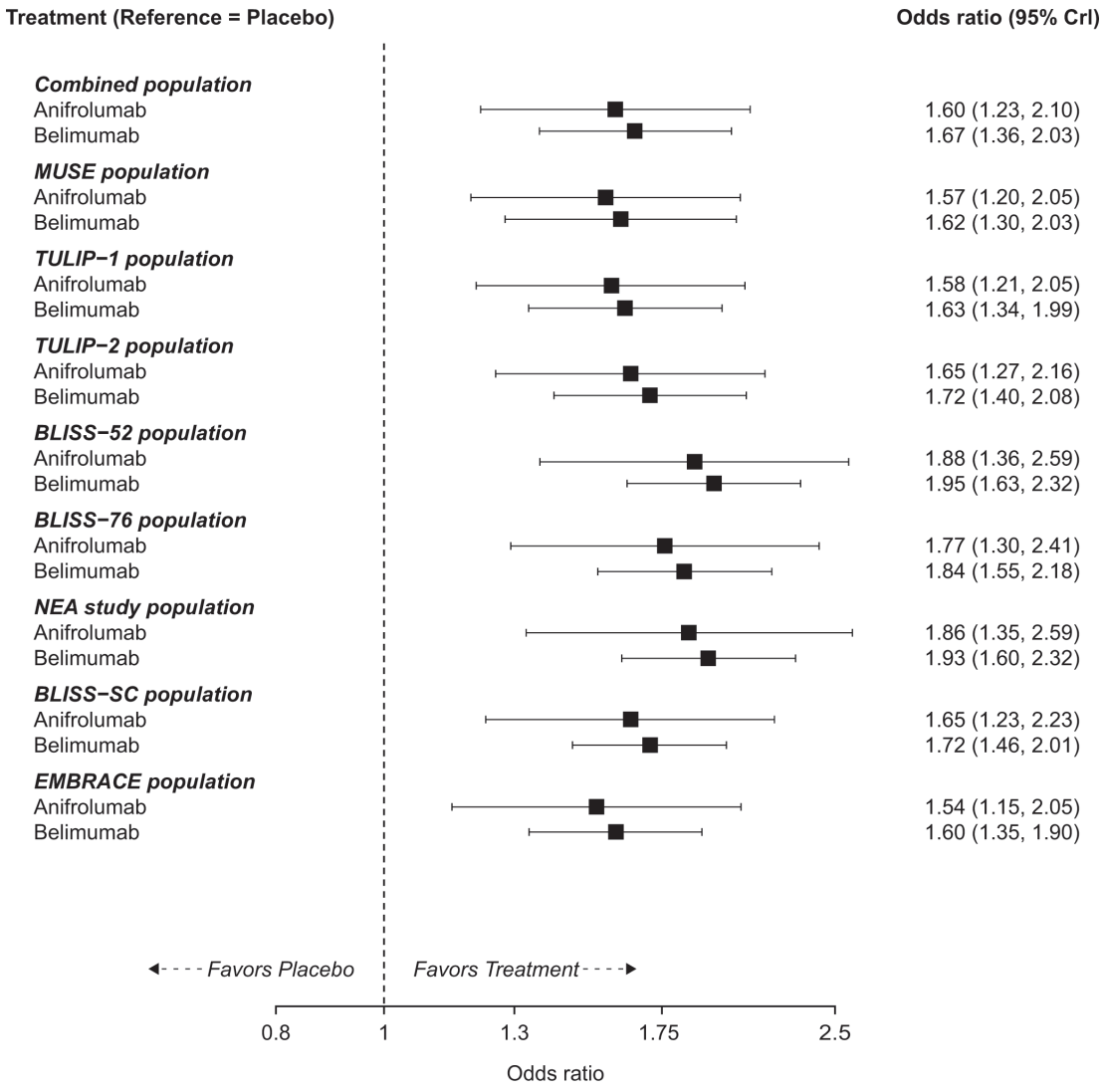


Figure 2 Predicted ORs for belimumab plus standard therapy and anifrolumab plus standard therapy versus placebo plus standard therapy for the base-case ML-NMR analysis of SRI-4 at 52 weeks in each population. Combined population is the pooled population across the three anifrolumab trials. CrI, credible interval; ML-NMR, multilevel network meta-regression; SRI-4, SLE Responder Index-4.

the direction of the point estimate slightly favouring belimumab (STC OR (95% CI) 1.15 (0.71 to 1.86); MAIC OR (95% CI) 1.14 (0.68 to 1.92)).

DISCUSSION

This study implemented a PAIC of RCT data to evaluate the efficacy of belimumab versus anifrolumab at 52 weeks in adults with SLE. The results of our analysis suggest that belimumab and anifrolumab are generally comparable in terms of SRI-4 at 52 weeks, but we cannot rule out the possibility of a clinically meaningful benefit for either treatment. Our results were consistent across the host of sensitivity analyses conducted. Given the differences identified in potential EMs, the ML-NMR results (and results from our other population adjustment models) are assumed to be less biased than the results using standard Bayesian NMA. Nonetheless, much of the bias in a standard Bayesian NMA appears to cancel out (some in favour of anifrolumab and some in favour of belimumab), so the results of the Bayesian NMA and PAIC analyses are largely consistent.

A key requirement of ITCs is that either the populations are inherently similar in terms of EMs (in the case of a standard ITC), or in the case of a PAIC, that they are appropriately adjusted to remove any inherent differences so that unbiased estimates can be obtained. When population adjustments are necessary, the population sample contributing the IPD must be large enough and broad enough to accurately estimate the treatment effects in the comparator population.[16 20] Our primary analyses with SRI-4 clearly met this requirement, with the IPD population sample (the five belimumab trials) consisting of >3000 patients, which was broad enough to accurately estimate the treatment effects in the anifrolumab population. The total sample size of our IPD in the MAIC of SRI-4 was 3080, with an effective sample size (ESS)

Lupus Sci Med: first published as 10.1136/lupus-2023-000907 on 5 May 2023. Downloaded from http://lupus.bmj.com/ on June 6, 2023 at Washington University School of Medicine Library &. Protected by copyright.

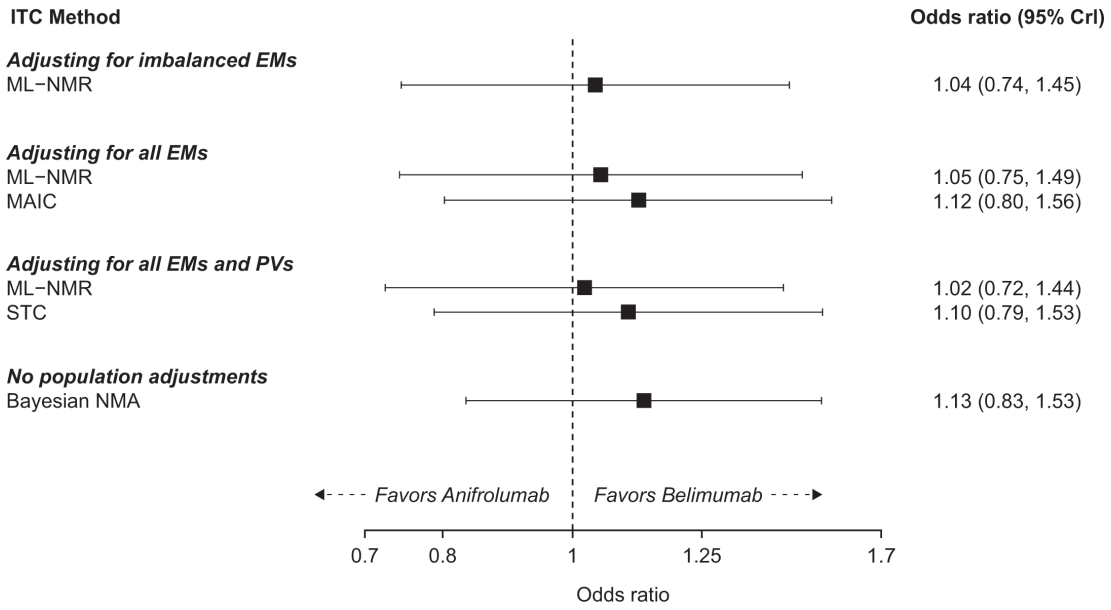


Figure 3 SRI-4 results at 52 weeks of belimumab plus standard therapy versus anifrolumab plus standard therapy for the base-case and sensitivity analyses. CrI, credible interval; EM, effect modifier; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; ML-NMR, multilevel network meta-regression; NMA, network meta-analysis; PV, prognostic variable; SRI-4, SLE Responder Index-4; STC, simulated treatment comparison.

post-adjustment of 1531. While this represents a sizeable reduction from the full sample size used to estimate the OR of belimumab versus placebo across the five belimumab trials, it is still a robust sample size to use for a PAIC.

An unexpected finding from our extensive set of PAIC analyses was that the high level of heterogeneity in SRI-4 results at 52 weeks across the three anifrolumab trials appears largely unrelated to any population differences across these three trials in EMs. This is true not only for our set of EMs but also appears to be true for the more extensive set that was identified in Bruce *et al*.[18] Consequently, our population-adjusted analyses can successfully explain the observed variation in SRI-4 for the belimumab trials, but not the differences in the trial-level SRI-4 results for the three anifrolumab trials. This finding highlights the fact that more research on anifrolumab is needed (which is beyond the scope of this study) to fully understand the effect of anifrolumab on SRI-4 at 52 weeks in the general SLE population. In the context of the current study, this finding means that the level of uncertainty around the placebo versus anifrolumab comparison and around the belimumab versus anifrolumab comparison may be even larger than what is estimated in our population-adjusted models.

When emulating the Bruce *et al*[18] approach, we obtained estimates in line with our primary analyses. These results are significantly different from those reported in Bruce *at al*.[18] For example, whereas we obtained an SRI-4 OR of 1.11 with the MAIC, Bruce *et al*[18] obtained an OR of 0.34 (reported as 2.91 in their publication as belimumab was used as the reference treatment). The key difference between our emulation of the Bruce *et al* approach and the actual approach in Bruce *et al* is that we had access to different IPD (we used IPD from the belimumab trials and Bruce *et al* used IPD from anifrolumab trials).[18] Consequently, our comparison was made in the combined anifrolumab trial population and the Bruce *et al*[18] comparison was made in the combined belimumab trial population (MAICs and STCs estimates can only be produced within the population that does not have IPD). If belimumab and anifrolumab treatment effects were modified in entirely different ways by the EMs, then it would be theoretically possible for both results to be correct. However, this is not considered clinically plausible, and therefore, other explanations are more likely.

It is likely that most or all of the differences between our results when emulating the Bruce *et al* approach and results in Bruce *et al* can be explained by the fact that Bruce *et al* did not have sufficient IPD available to undertake their approach.[18] As reported in Bruce *et al*,[18] the total sample size from the two anifrolumab trials (TULIP-1 and TULIP-2) in the MAIC of SRI-4 was 710 and the ESS post-adjustment of these two trials was only 71 patients (a 90% reduction). This can be loosely interpreted to mean that only 71 patients were used to inform the anifrolumab versus placebo comparison that was indirectly compared with belimumab. In contrast, the total sample size of our IPD in the MAIC of SRI-4 when emulating the Bruce *et al*[18] approach was 1125 (BLISS-52 and BLISS-76), with an ESS post-adjustment of 351 (approximately a 69% reduction). Thus, when emulating the Bruce *et al* approach,[18] we had an ESS approximately five times the size of what was available to inform the population adjustment produced in Bruce *et al*. When comparing the ESS from our primary analysis (n=1531) to that of Bruce *et al* (n=71),[18] our ESS is over 20 times larger.

It is also important to note that, beyond having limited IPD, there are further limitations to the Bruce *et al*

Lupus Sci Med: first published as 10.1136/lupus-2023-000907 on 5 May 2023. Downloaded from http://lupus.bmj.com/ on June 6, 2023 at Washington University School of Medicine Library &. Protected by copyright.

approach.[18 30] First, not all eligible trials in the evidence base were included in the analysis. Bruce *et al* contend that this was a necessary limitation due to issues with how STC and MAIC methods must be implemented.[31] However, the ML-NMR method we used does not suffer from the issues they allude to.[31] ML-NMR can be incorporated for any connected network of evidence and also provides a way to check assumptions (via a RE model) and evaluate model performance.[17 32] Thus, as we have demonstrated here, there is no need to remove eligible trials from the evidence base. Second, Bruce *et al*[18] employed SRI-4 values for the TULIP trials (≥4-point reduction in SLEDAI has not been reported elsewhere so could not be verified) that were higher than previously reported in the primary TULIP publications: OR of 1.63 reported in Bruce *et al*, while an OR of 1.33 would be expected based on a pooling of the prespecified results in the primary publications.[11 13] One possible explanation for the discrepancy could be that Bruce *et al* employed results from a post hoc analysis of the TULIP-1 SRI-4 results.[18] However, even if the revised post hoc definition for TULIP-1 was used when pooling the trials, the OR would be 1.56. Third, Bruce *et al*[18] adjusted for the proportion of patients with BILAG≥1 A or ≥2 B at baseline in the trials, despite the belimumab and anifrolumab trials using different versions of the BILAG (belimumab trials used the BILAG Classic; the anifrolumab trials used the BILAG 2004). In particular, the BILAG 2004 added two new organs/systems, removed the vasculitis section and rearranged other organ systems.[33] Thus, the apparent differences in BILAG across the trial populations may just be an artefact of the different instruments. This issue is further compounded because the apparent difference in proportion of patients with BILAG≥1 A or ≥2 B at baseline in the belimumab and anifrolumab trials appears to be the primary driver of why the IPD sample of Bruce *et al*[18] had poor overlap with the belimumab trial population. There were only approximately 40 patients (5.6% of the sample) in the anifrolumab trials that had no BILAG≥1 A or ≥2 B, and yet these 40 patients would have needed to account for 39% of the sample in order to align with the belimumab trial population. With such a small group of patients, even altering the results of just two or three patients (eg, observing 4 of 20 responses vs 7 out of 20 responses in the placebo arm) could have a dramatic impact on the overall results.

Our study also had limitations, mainly that our efficacy analyses were limited to a single outcome (SRI-4) and could only be conducted at 52 weeks. While SRI-4 has been associated with improvements in clinical, laboratory and patient-reported outcome measures,[34 35] no single outcome provides a comprehensive view of efficacy. With SRI-4, SLEDAI is used to assess improvement, while BILAG and Physician Global Assessment are incorporated to capture worsening. Thus, analyses of SRI-4 alongside outcomes that assess improvement in terms of BILAG (such as BILAG-based Composite Lupus Assessment) would provide a more nuanced picture of belimumab's potential to improve disease activity relative to anifrolumab. Similarly, Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index, which is a key measure for disease modification in SLE, represents another important dimension of efficacy not considered in this study.[36] Further analyses at other timepoints would be valuable to better understand how quickly both treatments become effective and how well efficacy is maintained. Although we assessed the feasibility of numerous other efficacy analyses, it was not possible to undertake analyses with any other endpoints. Even within our analysis of SRI-4, it must be noted that there were differences across trials in the precise definition of SRI-4 that was employed. Specifically, there were potential differences in joint scoring for the SLEDAI component of SRI-4 and there were also differences in terms of the BILAG instrument incorporated in the trials.

Our SRI-4 analyses were also unable to adjust for all eight of the EMs identified (adjusted for four in the base-case and six in the sensitivity). Specifically, none of our analyses adjusted for BMI or smoking status. Thus, it is possible that there was residual confounding in our analysis due to differences in BMI and smoking status across the belimumab and anifrolumab trials. We believe this is unlikely for BMI based on the limited BMI information that is available. However, the magnitude of the difference in the proportion of smokers is unknown. Beyond the EMs that could not be accounted for, there are also differences in time periods that the trials span (the anifrolumab trials were conducted in a post-belimumab world), which may translate into important differences in ST and prior therapies received at baseline. While the methodology we have used (only comparing the ORs across trials as opposed to the absolute proportion of responders) should mostly protect our results from being affected by this issue, we acknowledge the potential that some differences could still modify the treatment effects. A further limitation is that we only had access to the belimumab trial IPD and consequently had to make the 'shared EMs' assumption (that anifrolumab vs placebo relative effects are modified in the same way as belimumab vs placebo) to conduct the ML-NMR. If this assumption is violated, the results of the ML-NMR may be called into question.[32] However, the results of the ML-NMR, STC and MAIC are all very consistent and the latter two methods do not explicitly require the shared EM assumption (even when the shared EM assumption is violated, STC and MAIC are still unbiased in the specific population in which the analysis was undertaken). Thus, at worst case, the results may not be generalisable to other SLE populations. The fact that we did not have access to the anifrolumab IPD also meant we had to assume the type of marginal distribution of covariates and the correlation structure for the anifrolumab trials (not reported in the anifrolumab trials) based on what was observed in the belimumab trials.

In conclusion, we performed a robust PAIC analysis that suggests belimumab and anifrolumab are generally

Lupus Sci Med: first published as 10.1136/lupus-2023-000907 on 5 May 2023. Downloaded from http://lupus.bmj.com/ on June 6, 2023 at Washington University School of Medicine Library & Protected by copyright.

comparable in terms of SRI-4 response at 52 weeks. Future comparisons of belimumab versus anifrolumab may be valuable as more data for anifrolumab become available. It remains to be seen if specific groups of patients could derive a greater benefit from anifrolumab or from belimumab, and there is certainly an unmet need to identify robust predictors towards more personalised selection of available biological agents in SLE. However, our study did not find evidence to support that patients with SLE as a group would benefit from a change in treatment practices from belimumab to anifrolumab or vice versa.

Author affiliations
[1]Evidence Synthesis, Modeling & Simulation, Evidera, St-Laurent, Quebec, Canada
[2]Institute of Neurosciences "Federico Olóriz", University of Granada, Granada, Spain
[3]Evidence Synthesis, Modeling & Communication, Evidera, Waltham, Massachusetts, USA
[4]Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
[5]Department of Rheumatology, Clinical Immunology and Allergy, University of Crete School of Medicine, Crete, Greece
[6]Division of Rheumatology, Department of Medicine, Washington University School of Medicine, Saint Louis, Missouri, USA
[7]First Department of Propaedeutic Internal Medicine, "Laikon" General Hospital, National Kapodistrian University of Athens Medical School, Athens, Greece
[8]Specialty Care, Global Medical Affairs, GSK, Collegeville, Pennsylvania, USA
[9]Value Evidence and Outcomes, GSK, Brentford, UK
[10]Value Evidence and Outcomes, GSK, Collegeville, Pennsylvania, USA

Acknowledgements The authors would like to thank Kerry Gairy, previously an employee of GSK, and Kyle Fahrbach, Evidera, for their contributions to the development of the study's protocol and data interpretation. Medical writing support was provided by Helen Taylor, PhD, Fishawack Indicia Ltd, UK, part of Fishawack Health, and was funded by GSK.

Contributors BN, PS, MS and AM contributed to the acquisition of the data. BN, PS, MS, AM, RAL, DC and NB contributed to the conception or design of the study. All authors contributed to the analysis or interpretation of the data. NB is the guarantor of the manuscript.

Funding This study was funded by GSK (GSK Study 214174).

Competing interests BN, PS, MS and AM are employees of Evidera, a part of Thermo Fisher Scientific. MP has received grant/research support from AstraZeneca, Aurinia, Eli Lilly, Exagen, GSK, Janssen and Thermo Fisher. MP has also received consulting fees from the BPR Scientific Advisory Committee, Alexion, Amgen, AnaptysBio, Argenx, AstraZeneca, Aurinia, AxDev, Biogen, Boston Pharmaceuticals, Caribou Biosciences, CVS Health, Eli Lilly, Gilead Biosciences, GSK, Idorsia Pharmaceuticals, Janssen, Kezar Life Sciences, Kira Pharmaceuticals, Momenta Pharmaceuticals, Nimbus Lakshmi, Proviant, Sanofi, SinoMab and UCB. MP has received speakers' fees from Aurinia, MedShr and Arthros-FocusMedEd, and has received consulting fees for participation in a data safety monitoring board or advisory board for EMD Serono, Emergent Biosolutions, IQVIA and PPD Development. GKB has received consulting fees from Pfizer, Lilly and Novartis, and has received honorary fees from GSK, AstraZeneca, Pfizer, Novartis, Aenorasis, AbbVie and Lilly. GKB has also received a research grant from Pfizer. AHJK has received research support to Washington University from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant number P30 AR073752), National Center for Advancing Translational Sciences (grant number UL1 TR002345), Leona M. and Harry B. Helmsley Charitable Trust, Rheumatology Research Foundation, and National Multiple Sclerosis Society, GSK, and Foghorn Therapeutics. AHJK has performed consultancy for Alexion Pharmaceuticals, ANI Pharmaceuticals, AstraZeneca, Aurinia Pharmaceuticals, Exagen Diagnostics, GSK, Kypha and Pfizer unrelated to this work. AHJK has received payment or honoraria (for lectures, presentations, speakers bureaus, manuscript writing or educational events) from AstraZeneca, Aurinia Pharmaceuticals, Exagen Diagnostics and GSK. AHJK has participated on a data safety monitoring board or advisory board for National Institutes of Health/National Institute of Arthritis and Musculoskeletal and Skin Diseases. AHJK has been a board member for the Rheumatology Research Foundation Scientific Advisory Board and the Lupus Foundation of America-Heartland Chapter, and president of the St Louis Rheumatology Association. AHJK is the inventor of patent number 11029318 with Kypha unrelated to this work. The funders had no role in the decision to publish or preparation of this manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of Washington University, its affiliated academic health care centers, or the National Institutes of Health. AF has received honoraria and consulting fees from GSK, Aenorasis and AstraZeneca. AF has been a paid speaker for AbbVie, Amgen, Pfizer, Lilly, Genesis-Pharma, Novartis, UCB and Boehringer-Ingelheim. RAL, DC and NB are employees of GSK and hold stocks and shares in GSK.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Anonymised individual participant data for belimumab trials and the study documents for this analysis can be requested for further research from www.clinicalstudydatarequest.com.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Michelle Petri http://orcid.org/0000-0003-1441-5373
George K Bertsias http://orcid.org/0000-0001-5299-1406
Antonis Fanouriakis http://orcid.org/0000-0003-2696-031X
Roger A Levy http://orcid.org/0000-0001-6393-6031
Nick Ballew http://orcid.org/0000-0003-3507-0028

REFERENCES

1 Kaul A, Gordon C, Crow MK, *et al*. Systemic lupus erythematosus. *Nat Rev Dis Primers* 2016;2:16039.
2 Fanouriakis A, Kostopoulou M, Alunno A, *et al*. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. *Ann Rheum Dis* 2019;78:736–45.
3 Hill DD, Eudy AM, Egger PJ, *et al*. Impact of systemic lupus erythematosus disease activity, hydroxychloroquine and NSAID on the risk of subsequent organ system damage and death: analysis in a single US medical centre. *Lupus Sci Med* 2021;8:e000446.
4 GSK. Benlysta us prescribing information. Available: https://gskpro.com/content/dam/global/hcpportal/en_US/Prescribing_Information/Benlysta/pdf/BENLYSTA-PI-MG-IFU.PDF [Accessed Jul 2022].
5 Levy RA, Gonzalez-Rivera T, Khamashta M, *et al*. 10 years of belimumab experience: what have we learnt? *Lupus* 2021;30:1705–21.
6 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 3 trial. *Lancet* 2011;377:721–31.
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* 2011;63:3918–30.
8 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* 2017;69:1016–27.
9 Zhang F, Bae S-C, Bass D, *et al*. A pivotal phase III, randomised, placebo-controlled study of belimumab in patients with systemic

Lupus Sci Med: first published as 10.1136/lupus-2023-000907 on 5 May 2023. Downloaded from http://lupus.bmj.com/ on June 6, 2023 at Washington University School of Medicine Library & Protected by copyright.

lupus erythematosus located in China, Japan and South Korea. *Ann Rheum Dis* 2018;77:355–63.
10 Riggs JM, Hanna RN, Rajan B, *et al*. Characterisation of anifrolumab, a fully human anti-interferon receptor antagonist antibody for the treatment of systemic lupus erythematosus. *Lupus Sci Med* 2018;5:e000261.
11 Morand EF, Furie R, Tanaka Y, *et al*. Trial of anifrolumab in active systemic lupus erythematosus. *N Engl J Med* 2020;382:211–21.
12 Furie R, Khamashta M, Merrill JT, *et al*. Anifrolumab, an anti-interferon-α receptor monoclonal antibody, in moderate-to-severe systemic lupus erythematosus. *Arthritis & Rheumatology* 2017;69:376–86.
13 Furie RA, Morand EF, Bruce IN, *et al*. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. *The Lancet Rheumatology* 2019;1:e208–19.
14 Dias S, Sutton AJ, Ades AE, *et al*. 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.
15 Salanti G, Higgins JPT, Ades AE, *et al*. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17:279–301.
16 Phillippo DM, Ades AE, Dias S, *et al*. Methods for population-adjusted indirect comparisons in health technology appraisal. *Med Decis Making* 2018;38:200–11.
17 Phillippo DM, Dias S, Ades AE, *et al*. Multilevel network meta-regression for population-adjusted treatment comparisons. *J R Stat Soc Ser A Stat Soc* 2020;183:1189–210.
18 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* 2022;11:765–77.
19 Ballew N, Mian A, Levy RA, *et al*. Letter to the editor: indirect treatment comparison of anifrolumab efficacy versus belimumab in adults with systemic lupus erythematosus. *J Comp Eff Res* 2023;12:e220106.
20 Phillippo DM, Dias S, Ades AE, *et al*. Assessing the performance of population adjustment methods for anchored indirect comparisons: a simulation study. *Stat Med* 2020;39:4885–911.
21 Remiro-Azócar A, Heath A, Baio G. Methods for population adjustment with limited access to individual patient data: a review and simulation study. *Res Synth Methods* 2021;12:750–75.
22 Remiro-Azócar A, Heath A, Baio G. Parametric G-computation for compatible indirect treatment comparisons with limited individual patient data. *Res Synth Methods* 2022;13:716–44.
23 Cochrane. Cochrane Handbook for systematic reviews of interventions. Available: https://training.cochrane.org/handbook/current [Accessed Oct 2022].
24 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
25 Cope S, Zhang J, Saletan S, *et al*. A process for assessing the feasibility of a network meta-analysis: a case study of everolimus in combination with hormonal therapy versus chemotherapy for advanced breast cancer. *BMC Med* 2014;12:93.
26 Ginzler E, Guedes Barbosa LS, D'Cruz D, *et al*. Phase III/IV, randomized, fifty-two-week study of the efficacy and safety of belimumab in patients of black African ancestry with systemic lupus erythematosus. *Arthritis Rheumatol* 2022;74:112–23.
27 Phillippo DM. Multinma: Bayesian network meta-analysis of individual and aggregate data. R package version 0.5.0. Available: https://zenodo.org/record/7033094#.Y8bhAHb7Q2w [Accessed Oct 2022].
28 Signorovitch JE, Sikirica V, Erder MH, *et al*. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health* 2012;15:940–7.
29 Phillippo DM, Ades T, Dias S, *et al*. Nice dsu technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE. Available: https://research-information.bris.ac.uk/ws/portalfiles/portal/94868463/Population_adjustment_TSD_FINAL.pdf [Accessed Oct 2022].
30 Ballew N, Mian A, Levy RA, *et al*. Letter to the editor: indirect treatment comparison of anifrolumab efficacy versus belimumab in adults with systemic lupus erythematosus. *J Comp Eff Res* 2023;12.
31 Bruce IN, Golam S, Steenkamp J, *et al*. Letter in reply: indirect treatment comparison of anifrolumab efficacy versus belimumab in adults with systemic lupus erythematosus. *J Comp Eff Res* 2023;12:e220192.
32 Phillippo DM, Dias S, Ades AE, *et al*. Validating the assumptions of population adjustment: application of multilevel network meta-regression to a network of treatments for plaque psoriasis. *Med Decis Making* 2023;43:53–67.
33 Murphy C-L, Yee C-S, Gordon C, *et al*. From BILAG to BILAG-based combined lupus assessment-30 years on. *Rheumatology (Oxford)* 2016;55:1357–63.
34 Furie R, Wang L, Illei G, *et al*. Systemic Lupus Erythematosus (SLE) Responder Index response is associated with global benefit for patients with SLE. *Lupus* 2018;27:955–62.
35 Thanou A, Chakravarty E, James JA, *et al*. Which outcome measures in SLE clinical trials best reflect medical judgment? *Lupus Sci Med* 2014;1:e000005.
36 van Vollenhoven R, Askanase AD, Bomback AS, *et al*. Conceptual framework for defining disease modification in systemic lupus erythematosus: a call for formal criteria. *Lupus Sci Med* 2022;9:e000634.

Lupus Sci Med: first published as 10.1136/lupus-2023-000907 on 5 May 2023. Downloaded from http://lupus.bmj.com/ on June 6, 2023 at Washington University School of Medicine Library & Protected by copyright.

Belimumab versus anifrolumab in adults with systemic lupus erythematosus: an indirect comparison of clinical response at 52 weeks (18/50 words max)

Dr. Binod Neupane[1], Ms. Pragya Shukla[1], Dr. Mahmoud Slim[1,2], Ms. Amber Martin[3], Dr. Michelle Petri[4], Dr. George K. Bertsias[5], Dr. Alfred H. J. Kim[6], Dr. Antonis Fanouriakis[7], Dr. Roger Abramino Levy[8], Mr. Deven Chauhan[9], Dr. Nick Ballew[10]

[1]Evidence Synthesis, Modeling & Simulation, Evidera, St-Laurent, QC, Canada; [2]Institute of Neurosciences "Federico Olóriz", University of Granada, Granada, Spain; [3]Evidence Synthesis, Modeling & Communication, Evidera, Waltham, MA, USA; [4]Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [5]Department of Rheumatology, Clinical Immunology and Allergy, University of Crete School of Medicine, Crete, Greece; [6]Division of Rheumatology, Department of Medicine, Washington University School of Medicine, Saint Louis, MO, USA; [7] First *Department of Propaedeutic Internal Medicine, "Laikon" General Hospital, National Kapodistrian University of Athens Medical School, Athens, Greece; [8]Specialty Care, Global Medical Affairs, GSK, Collegeville, PA, USA; [9]Value Evidence and Outcomes, GSK, Brentford, UK; [10]Value Evidence and Outcomes, GSK, Collegeville, PA, USA*

Running head: Indirect comparison of clinical response with belimumab versus anifrolumab

Neupane B, *et al. Lupus Sci Med* 2023; 10:e000907. doi: 10.1136/lupus-2023-000907

Corresponding author:

Name: Nick Ballew

Address: GSK, Value Evidence and Outcomes, 1250 S. Collegeville Road, Collegeville, PA 19426, USA

Tel: +1 610-917-7000

Email: nick.g.ballew@gsk.com

ORCID ID: 0000-0003-3507-0028

SUPPLEMENTARY MATERIAL

Appendix 1: Background on indirect treatment comparisons (ITCs), prognostic factors and treatment effect modifiers (EMs)

Randomized controlled trials (RCTs) are the gold standard for estimating the relative efficacy and safety estimates between treatments of interest. However, in the absence of head-to-head comparisons, ITC techniques can generate valid comparative estimates when the corresponding assumptions are met.[1, 2] The first step towards ensuring the credibility of any ITC is to establish the validity of the evidence base.[3] In other words, the analysis needs to accurately reflect the complete evidence base available to make a like-for-like ITC. This involves a comprehensive search of the literature, the accurate extraction of information from each trial, and a comprehensive feasibility assessment to evaluate if the different assumptions for conducting standard ITC are adequately fulfilled.

Once these steps have been fulfilled, the standard approach to indirectly compare treatments from RCTs that share a common comparator (placebo) is to respect randomization within trials and compare relative treatment effects across the studies.[4] In other words, instead of comparing the proportion of responders on anifrolumab from study 1 to the proportion of responders on belimumab from study 2, the odds ratio (OR) of anifrolumab versus placebo from study 1 is compared with the OR of belimumab versus placebo in study 2. The rationale behind this approach is that it can produce unbiased estimates across a greater number of scenarios when the distributions of the treatment effects modifiers in study 1 and study 2 are similar.

Whereas an ITC of arm-level outcomes (proportions) can be biased by differences across trials that have either a prognostic effect or treatment-modifying effect on a given outcome, the standard approach using relative effects (ORs) remains unbiased in the face of differences in prognostic characteristics.[5] Prognostic characteristics are those that have an impact on the arm-level effect of a treatment without altering the relative effect (impact proportion of responders across treatments in a similar way, so that the OR remains unchanged). In contrast, EMs are characteristics that alter the relative effect of a treatment, so that the treatment is more or less effective than an alternative treatment, depending on the level of the EM. For example, if the OR of systemic lupus erythematosus (SLE) Responder Index-4 (SRI-4) response in belimumab versus placebo is 0.8 in a low disease activity population and 2.0 in a high disease activity population, disease activity is an EM. When there are differences across trials in EMs, the transitivity assumption of standard ITC is violated, which will subsequently result in the generation of biased estimates.

Appendix 2: Population-adjusted indirect comparisons (PAICs)

Conventional ITC techniques, such as network meta-analysis (NMA), assume transitivity in the network, i.e., the distributions of EMs are balanced across the different sets of trials included in the network.[6] In the event of intransitivity, adjustment techniques, referred to as PAICs, can account for the differences in EMs by leveraging individual patient data (IPD) of the index trial.[7, 8] PAICs include matching-adjusted indirect comparison (MAIC), simulated treatment comparison (STC), and multi-level network meta-regression (ML-NMR).

Key assumptions

While conventional ITC techniques assume the "constancy of relative effects", i.e., transitivity, PAICs including MAIC, STC, and ML-NMR relax this assumption to "conditional constancy of relative effects" following the adjustment of the imbalanced EMs with respect to the chosen comparison scale.

The population-adjusted estimates generated using the MAIC and STC are only applicable to the population of the comparator trial (i.e., anifrolumab trials). To generalize such estimates to other target populations, the "shared EM" assumption needs to be met. Such an assumption might be needed in the context of ML-NMR when conducted in smaller networks (like the current network in our study).

While MAIC and STC disregard the correlation between covariates, assumptions regarding the marginal distribution and the correlation structure of covariates are required in the ML-NMR to construct the covariate joint distribution in trials with aggregate data.[8, 9]

ML-NMR

The ML-NMR, an extension of the conventional NMA framework, synthesizes the evidence from a connected network of studies where IPD from certain trials (i.e., belimumab trials) and aggregate data from other trials (i.e., anifrolumab trials) are available. In contrast to the other PAIC methods, ML-NMR allows the inclusion of more than just two trials and enables the conduct of comparisons in any target population within a given covariate distribution.[10] Furthermore, to avoid aggregation bias, the ML-NMR integrates an individual-level model over the covariate distribution from each study with aggregate data instead of using the mean covariate values.

In the base-case ML-NMR analysis adjusting for only imbalanced covariates, all five belimumab trials and all three anifrolumab trials were included in the network as unique trials. However, in the models adjusting all identified and feasible treatment EMs and prognostic factors, the belimumab trials were pooled into three trials as follows: (1) BLISS-52/BLISS-76, (2) EMBRACE and NEA study, and (3) BLISS-SC. We undertook this approach because the percentage of non-Black African ancestry patients (one of the EMs that was relatively balanced across belimumab and anifrolumab trials) was <2% in the EMBRACE trial.

The ML-NMR analyses were run under the Bayesian framework using multinma package in R.[11] Vague, Normal(0,100), prior was assumed for effects parameters (i.e., for the log-odds ratios and baseline effects). In a sensitivity analysis with random-effect model to assess the residual heterogeneity, half-normal(0.5) prior was used for the between-study standard deviation (SD) parameter. Markov Chain Monte-Carlo (MCMC) simulations were run in three chains where MCMC samples from the first 7000 iterations were discarded and samples

from another 4000 iterations were saved in each chain for posterior estimation. Convergence to the posterior distributions was achieved in all analyses. The median and (2.5^{th}, 97.5^{th}) quantiles of the saved (posterior) samples of a parameter were used as the estimate and 95 credible limits for the parameter.

MAIC

The anchored MAIC employs propensity score re-weighting to balance the differences in key EMs between included trials. The weights are derived in such a way that the re-weighted population profile of the index trial (i.e., belimumab trials) with respect to the EMs matches that of the comparator trials (i.e., anifrolumab trials). Subsequently, the treatment effects on the outcome of interest can be compared between balanced trial populations.[10, 12]

The weights are derived using a propensity score-type logistic regression model which predicts the enrollment in the anifrolumab trials versus the belimumab trials, as a function of the treatment EMs. Specifically, weights are estimated as $w_i = exp\ (\alpha + x_i'\beta)$, where

x_i' consists of the list of EMs

The β coefficients are estimated by the method of moments rather than maximum likelihood, as only aggregate data from the comparator trial (anifrolumab trials) are available[12]

Once the coefficients are estimated, the individual patient weights using IPD in the belimumab trials are estimated. The weights can then be used to calculate the effective sample size (ESS) achieved after weighting as $ESS = (\sum w_i)^2/(\sum w_i^2)$. Small ESS is indicative

of poor population overlap between the index and comparator trials and can subsequently lead to unstable model estimates.[10]

In our study, the MAIC was undertaken while pooling the IPD from the five belimumab trials representing a single index trial. Participants weights were calculated while centering the EMs from the belimumab trials to the corresponding covariates values in the anifrolumab trials. The calculated weights were used in a simple (weighted) logistic regression analysis model using belimumab IPD with the outcome (i.e., SRI-4) regressed against the treatment, which was the unique covariate in the model, to estimate the OR of achieving SRI-4 response for belimumab versus placebo in the average anifrolumab population. The (raw/unadjusted) estimate of OR of anifrolumab versus placebo was computed by pooling the results from the three eligible anifrolumab trials. Finally, the OR for belimumab versus anifrolumab was computed using the Bucher et al. ITC method.[13]

In the MAIC analyses that was undertaken to emulate Bruce et al. methods (for the SRI-4 and SLE Disease Activity Index [SLEDAI]-2000 [2K] 4-point reduction outcomes),[14] the methodological approach outlined above was followed using the data from two belimumab trials (BLISS-52/BLISS-76; pooled as a single trial) and two anifrolumab trials (TULIP-1/TULIP-2; pooled as a single trial). In addition, the same set of EMs identified by Bruce et al. 2022 was used.[14]

STC

The anchored STC is based on regression-based adjustment. The STC fits logistic regression using the IPD from the index trials (belimumab trials) to create a predictive equation. The covariates included in the model are centered at the published mean estimates from the anifrolumab trials. As per the National Institute for Health and Care Excellence guidance,[10] it is recommended to include all the EMs that are imbalanced between trials as well as prognostic variables as this will improve model fit.

The predictive equation is then used to estimate the effects of belimumab in the comparator trial population (i.e., anifrolumab trials' population). These results can then be used to estimate the relative effects of belimumab versus anifrolumab in the comparator trial population.

As in the MAIC, first the EMs and the prognostic variables (not applicable in MAIC) from belimumab trials were centered using the IPD at the weighted average of the means of the corresponding covariate values in the anifrolumab trials. Then, a logistic regression model of SRI-4 was directly run (unlike in MAIC where the weights are first derived before running the regression analysis) to estimate the effect of belimumab versus placebo in the average anifrolumab trial population. In the model, the outcome of interest (i.e., SRI-4) was regressed against the treatment (belimumab vs placebo), all treatment EM variables (centered) and their interactions with the treatment and all prognostic variables (centered). Since the covariates were already centered at average anifrolumab population, the estimate of the treatment (belimumab) effect is its estimate in average anifrolumab population. Then the estimate of OR of belimumab versus placebo was compared with that of anifrolumab

versus placebo using the Bucher et al. ITC method.[13] In our analyses, the data from all five belimumab trials and all three anifrolumab trials were included.

In the STC analyses that were undertaken to emulate Bruce et al. methods (for the SRI-4 and SLEDAI-2K 4-point reduction outcomes),[14] the methodological approach outlined above was followed using the data from two belimumab trials (BLISS-52/BLISS-76; pooled as a single trial) and two anifrolumab trials (TULIP-2/TULIP-2; pooled as a single trial). In addition, the same set of EMs identified by Bruce et al. was used.[14]

Appendix 3: Details on the systemic literature review (SLR) methodology

Data sources

We conducted an SLR according to the rigorous methodology outlined by the Cochrane Collaboration and Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.[6, 15] Literature searches were conducted in Embase, MEDLINE, and The Cochrane Central Register of Controlled Clinical Trials to identify English-language studies conducted on humans and published by April 12, 2022. The strategies for each electronic literature database included a combination of free-text and medical subject headings, grouped into the following categories: population, interventions, study design, and limits (including timeframe, language, and publication type). In addition, the searches were supplemented by the review of records from 10 key conferences from 2019–2021 meetings and the clinicaltrials.gov trial registry.

Study selection

Studies were screened against predefined inclusion and exclusion criteria based on the population, intervention, comparison, outcome, and study design and timeframe described in **Table S1**. Studies were eligible for inclusion if they reported on randomized controlled or single-arm trials investigating the efficacy and/or safety of belimumab or anifrolumab in adult patients diagnosed with SLE. Title and abstract screening, as well as full-text screening, were undertaken by two independent investigators and any discrepancies were resolved by a third more senior investigator.

Data extraction

Data from eligible studies were extracted by a single investigator using standardized data extraction tables. All extractions were independently validated by a senior investigator. For each of the included studies, we extracted data elements corresponding to the study design characteristics (study phase, duration, and eligibility criteria), treatment characteristics (dose strength, frequency, route of administration), baseline patient characteristics, and efficacy and safety outcomes of interest. The methodological quality of the RCTs was assessed using Cochrane Risk of Bias Assessment Tool v1.0.

Table S1. Eligibility criteria

Domain	Inclusion Criteria	Exclusion Criteria
Population	• Adult patients (≥18 years) diagnosed with SLE	• Patients without SLE • Patients with only active LN; included if only kidney involvement in SLE • ≥15% of patients have LN[a] • Pediatric patients <18 years • Patients with comorbid SLE and rheumatoid arthritis
Interventions	• Belimumab plus standard therapy • Anifrolumab plus standard therapy	• Study evaluates treatment other than interventions of interest
Comparators	• Placebo plus standard therapy • Standard therapy alone	• NA
Outcomes	Efficacy: • SELENA-SLEDAI score: change in score, % with response[b]	• Studies that did not report at least one of the outcomes of interest

Domain	Inclusion Criteria	Exclusion Criteria
	○ Response rates for the specific SELENA-SLEDAI organ domains involved at baseline • SLEDAI-2K score: change in score, % with response[b] ○ Response rates for the specific SLEDAI-2K organ domains involved at baseline • BILAG score: change in score, % with response[b] ○ Response rates for the specific BILAG organ domains involved at baseline • BICLA: % with response[b] • PGA scale: change in score • SDI score: change in score • SRI-4: % with response[b] • CLASI: change in score, % with response[b] • Flares ○ Annual flare rate ○ Time to first flare ○ Proportion of patients with flares • ≥50% reduction in both swollen and tender joints	

Domain	Inclusion Criteria	Exclusion Criteria
	• Reduction in glucocorticoids use PROs: • SF-36 • FACIT Safety: • Incidence and severity of AEs • Incidence of SAEs • Mortality • Any discontinuations • Discontinuations due to AEs	
Study design	• RCTs • Single-arm clinical trials • Pooled studies[c]	• Crossover designs that did not include adequate washout period (≥7 days) and did not have statistical analysis taking paired design into account

Domain	Inclusion Criteria	Exclusion Criteria
		• Letters, case reports, editorials, reviews • Observational designs: prospective and retrospective cohorts, cross-sectional, and case-control studies
Time period	• January 1, 1946 – April 12, 2022	• Studies published after April 2022
Language	• English	• Languages other than English

[a]Includes patients with a diagnosis of LN and baseline grade A scores in the renal domain of BILAG or any indication of renal involvement at baseline. Trials with a mixed patient population (i.e., including patients with LN) were included as long as the proportion of patients with LN ≤15%; [b]Definition of response or remission to be captured; [c]For the purposes of quantitative evidence synthesis, results of pooled studies were not included if the individual trial findings were included, to avoid data duplication.

AE, adverse event; BICLA, BILAG-Based Composite Lupus Assessment; BILAG, British Isles Lupus Activity Group; CLASI, Cutaneous Lupus Erythematosus Disease Area and Severity Index; FACIT, Functional Assessment of Chronic Illness Therapy measurement system; LN, lupus nephritis; NA, not applicable; PGA, Physician Global Assessment; PRO, patient-reported outcome; RCT, randomized controlled trial; SAE,

serious adverse event; SDI, Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index; SELENA, Safety of Estrogens in Lupus National Assessment; SF-36, 36-item Short-form health survey; SLE, systemic lupus erythematosus; SLEDAI-2K, SLE Disease Activity Index 2000; SRI-4, SLE Responder Index-4.

Appendix 4: Process for identifying and selecting treatment EMs

Details on the exploratory analyses and results for SRI-4

A 3-step approach was used to identify treatment EMs and prognostic variables:

1. The literature on belimumab and anifrolumab in SLE was reviewed for well-established EMs and potential differences based on subgroup results reported in RCTs were examined
2. For each baseline covariate, regression models were run in which the outcome of interest was regressed against treatment and the covariate (prognostic effect testing) and against treatment, covariate and their interaction term (effect modification testing)
3. Clinical input was obtained on the relevance of the EMs and prognostic factors identified in Step 1, or on other factors not picked by regression models

- Predictive equations in the regression models were developed to identify potential EMs and prognostic variables using IPD from the belimumab trials only (due to unavailability of anifrolumab IPD). The strength of the effect was assessed, and significance level of the effects were used to identify potential EMs ($p<0.1$)
- The results of the logistic regression analyses assessing the interaction effects of pooled belimumab doses with the different covariates on SRI-4 gave the expected difference in the relative treatment effect of treatment versus placebo for the level

of variable versus the reference (for categorical variables) or versus one unit increase (for continuous variables) (**Table S2**)

- To identify potential prognostic factors, we performed logistic regression analysis including treatment and baseline characteristics as independent factors on SRI-4 (**Table S2**)

- Analyses on SRI-4 were derived from Safety of Estrogens in Lupus National Assessment (SELENA)-SLEDAI and were expected to align with outcomes derived from SLEDAI-2K

- A total of six treatment EMs were identified: Black African ancestry [binary], SLEDAI-2K [continuous], complement (C)3 [binary], C4 [binary], anti–double-stranded DNA (dsDNA) antibody positive [binary], any oral corticosteroid (OCS) use [binary]

    - Black African ancestry has the following two categories: all others; Black African ancestry

    - For anti-dsDNA positive: positivity in belimumab trials was based on 30 IU/mL threshold while positivity in anifrolumab trials was based on 15 IU/mL threshold. The numerical values may not be equivalent across test types so instead of applying the same numerical threshold, the decision was made to apply the original threshold used in each trial

- Two potential prognostic variables were identified: Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index (SDI) [continuous] and immunosuppressant use [binary]

  - For any immunosuppressant use: Only the pooled results were reported for anifrolumab trials. Thus, it was assumed that immunosuppressant use at baseline was 48% for all three anifrolumab trials

- SLEDAI, British Isles Lupus Activity Group (BILAG), and Physician Global Assessment (PGA) are all measures of disease severity. Therefore, only SLEDAI was selected for adjustment. SDI, which demonstrated prognostic capacity, is specific to organ damage and therefore can be adjusted for simultaneously with SLEDAI

- Two variables, smoking status and body-mass index (BMI), that were considered EMs based on the feedback obtained from the lupus experts were not adjusted for. While none of the included studies reported the smoking status at baseline, BMI was missing in the MUSE trial.[16] In general, the BMI was balanced across the belimumab (25.4 kg/m^2) and anifrolumab trials (27.6 kg/m^2)

- We did not adjust for the disease duration since the means/SDs from TULIP-1 and TULIP-2 were not reported. In addition, assuming that the reported medians were equal to the mean disease duration was not possible because the data were highly skewed. Furthermore, using the method proposed by Wan et al.[17] to estimate the mean from the median produced much larger estimates (around 13 years more compared with the mean duration available from MUSE), which indicated that these values were overestimates

Table S2. Effect modifiers and prognostic factors tested using logistic regression

		Logistic regression models assessing effect modification[a]			**Logistic regression models assessing prognostic effects[b]**	
Variable (Reference population)	**Level**	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC)** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC) x Covariate** OR [p-value]	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/ 200 mg SC)** OR [p-value]
Age, years	Continuous[c]	1.00 [0.9112]	1.69 [0.0001]	0.99 [0.4056]	1.00 [0.4129]	1.68 [0.0001]
Sex ('Male')	Female	1.11 [0.6769]	1.84 [0.0513]	0.91 [0.7692]	1.05 [0.7638]	1.69 [0.0001]
Race ('Others')	White	0.70 [0.0609]	1.94 [0.0030]	1.04 [0.8672]	0.73 [0.0120]	1.73 [0.0001]
	Asian	0.69 [0.0485]		0.89 [0.6408]	0.64 [0.0006]	
	Black African ancestry	0.75 [0.1795]		0.63 [0.1004]	0.57 [0.0001]	
Race ('White/Others')	Asian	0.89 [0.3753]	1.98 [0.0001]	0.87 [0.4061]	0.82 [0.0169]	1.72 [0.0001]
	Black African ancestry	0.98 [0.8831]		0.62 [0.0174]	0.72 [0.0010]	
Race ('All Others')	Black African ancestry	1.02 [0.8748]	1.84 [0.0001]	0.67 [0.0313]	0.79 [0.0096]	1.70 [0.0001]

		Logistic regression models assessing effect modification[a]			**Logistic regression models assessing prognostic effects[b]**	
Variable (Reference population)	**Level**	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC)** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC) x Covariate** OR [p-value]	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/ 200 mg SC)** OR [p-value]
SELENA-SLEDAI	Continuous[c]	1.06 [0.0002]	1.72 [0.0001]	1.08 [0.0002]	1.11 [0.0001]	1.72 [0.0001]
SELENA-SLEDAI ('≤9')	≥10	1.63 [0.0001]	1.28 [0.0299]	1.69 [0.0006]	2.24 [0.0001]	1.73 [0.0001]
SDI	Continuous[c]	0.82 [0.0005]	1.60 [0.0001]	1.07 [0.3429]	0.85 [0.0001]	1.67 [0.0001]
BILAG No A or B ('No')	Yes	0.79 [0.2580]	1.77 [0.0001]	0.62 [0.0729]	0.58 [0.0001]	1.70 [0.0001]
BILAG 1A/2B ('No')	Yes	1.20 [0.1388]	1.42 [0.0045]	1.35 [0.0551]	1.44 [0.0001]	1.71 [0.0001]
Cardiovascular & Respiratory ('No')	Yes	1.08 [0.7565]	1.69 [0.0001]	0.95 [0.8849]	1.05 [0.7570]	1.69 [0.0001]
CNS ('No')	Yes	0.32 [0.0804]	1.67 [0.0001]	3.25 [0.1169]	0.73 [0.3185]	1.69 [0.0001]
Hematology ('No')	Yes	0.98 [0.9251]	1.71 [0.0001]	0.85 [0.5030]	0.89 [0.3272]	1.69 [0.0001]
Immunologic ('No')	Yes	0.72 [0.0182]	1.34 [0.0694]	1.34 [0.1078]	0.85 [0.0787]	1.69 [0.0001]

		Logistic regression models assessing effect modification[a]			**Logistic regression models assessing prognostic effects[b]**	
Variable (Reference population)	**Level**	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC)** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC) x Covariate** OR [p-value]	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/ 200 mg SC)** OR [p-value]
Mucocutaneous ('No')	Yes	1.79 [0.0009]	1.88 [0.0019]	0.89 [0.5938]	1.66 [0.0001]	1.70 [0.0001]
Musculoskeletal ('No')	Yes	1.83 [<0.0001]	1.61 [0.0002]	1.10 [0.5427]	1.94 [0.0001]	1.72 [0.0001]
Renal ('No')	Yes	0.87 [0.3075]	1.69 [0.0001]	0.99 [0.9646]	0.87 [0.0881]	1.69 [0.0001]
Vascular ('No')	Yes	1.31 [0.1979]	1.64 [0.0001]	1.39 [0.2288]	1.60 [0.0004]	1.68 [0.0001]
PGA	Continuous[c]	1.19 [0.1663]	1.69 [0.0001]	0.93 [0.6680]	1.14 [0.0908]	1.69 [0.0001]
PGA ('≥1')	<1	0.74 [0.1330]	1.69 [0.0001]	1.01 [0.9667]	0.74 [0.0150]	1.69 [0.0001]
C3	Continuous[c]	1.01 [<0.0001]	1.71 [0.0001]	0.99 [0.0005]	1.01 [0.0001]	1.70 [0.0001]
Low C3 ('No')	Yes	0.63 [0.0001]	1.35 [0.0037]	1.64 [0.0010]	0.85 [0.0298]	1.70 [0.0001]
C4	Continuous[c]	1.04 [<0.0001]	1.70 [0.0001]	0.96 [0.0001]	1.02 [0.0001]	1.69 [0.0001]
Low C4 ('No')	Yes	0.53 [<0.0001]	1.36 [0.0009]	1.74 [0.0004]	0.74 [0.0001]	1.67 [0.0001]

		Logistic regression models assessing effect modification[a]			Logistic regression models assessing prognostic effects[b]	
Variable (Reference population)	**Level**	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC)** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC) x Covariate** OR [p-value]	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/ 200 mg SC)** OR [p-value]
SLE duration, years	Continuous[c]	0.97 [0.0007]	1.69 [0.0001]	1.02 [0.0442]	0.98 [0.0027]	1.69 [0.0001]
Azathioprine use ('No')	Yes	0.85 [0.2687]	1.68 [0.0001]	1.04 [0.8196]	0.87 [0.1292]	1.69 [0.0001]
Methotrexate use ('No')	Yes	1.00 [0.9898]	1.73 [0.0001]	0.76 [0.2300]	0.86 [0.2021]	1.68 [0.0001]
Steroid use ('No')	Yes	0.93 [0.7003]	1.12 [0.6167]	1.60 [0.0441]	1.25 [0.0520]	1.69 [0.0001]
Anti-dsDNA (Original)	Continuous[c]	1.00 [0.1154]	1.61 [0.0001]	1.00 [0.0932]	1.00 [0.9367]	1.69 [0.0001]
Anti-dsDNA (log-transformed)	Continuous[c]	0.85 [0.0001]	1.71 [0.0001]	1.13 [0.0148]	0.93 [0.0016]	1.70 [0.0001]
Anti-dsDNA ('<30 IU/mL')	≥30 IU/mL	0.75 [0.0251]	1.43 [0.0085]	1.27 [0.1483]	0.87 [0.0777]	1.69 [0.0001]
OCS dose ≥7.5 mg/day ('≤7.5 mg/day')	>7.5 mg/day	1.19 [0.1490]	1.66 [0.0001]	1.02 [0.8983]	1.21 [0.0133]	1.69 [0.0001]

		Logistic regression models assessing effect modification[a]			**Logistic regression models assessing prognostic effects[b]**	
Variable (Reference population)	**Level**	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC)** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC) x Covariate** OR [p-value]	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/ 200 mg SC)** OR [p-value]
OCS dose ≥10 mg/day ('≤10 mg/day')	>10 mg/day	1.09 [0.4924]	1.58 [0.0001]	1.20 [0.2463]	1.21 [0.0104]	1.68 [0.0001]
Immuno-suppressants ('No')	Yes	0.73 [0.0074]	1.62 [0.0001]	1.08 [0.6148]	0.77 [0.0002]	1.69 [0.0001]
Antimalarials ('No')	Yes	1.31 [0.0361]	2.01 [0.0001]	0.78 [0.1259]	1.12 [0.1455]	1.69 [0.0001]
MMF use ('No')	Yes	0.63 [0.0041]	1.60 [0.0001]	1.41 [0.0963]	0.78 [0.0096]	1.69 [0.0001]
BMI	Continuous[c]	1.00 [0.7419]	1.84 [0.0589]	1.00 [0.8519]	1.00 [0.4405]	1.73 [0.0001]
BMI categorical ('Normal weight')	Underweigh t	1.38 [0.1794]	1.69 [0.0001]	0.90 [0.7258]	1.30 [0.0938]	1.74 [0.0001]
	Overweight	1.02 [0.8968]		1.23 [0.2735]	1.15 [0.1281]	
	Obese	1.11 [0.5219]		0.93 [0.7392]	1.06 [0.5390]	

		Logistic regression models assessing effect modification[a]			Logistic regression models assessing prognostic effects[b]	
Variable (Reference population)	**Level**	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC)** OR [p-value]	**Belimumab (10 mg/kg IV/200 mg SC) x Covariate** OR [p-value]	**Covariate** OR [p-value]	**Belimumab (10 mg/kg IV/ 200 mg SC)** OR [p-value]
Obese ('Obese')	Non-Obese	0.93 [0.6445]	1.58 [0.0091]	1.12 [0.5597]	1.00 [0.9927]	1.73 [0.0001]

[a]Effect modification was examined by fitting logistic regression models in which the outcome, SRI-4, was regressed against treatment, covariate and their interaction term (effect modification testing); [b]prognostic effects were examined by fitting logistic regression models in which the outcome, SRI-4, was regressed against treatment and the covariate (prognostic effect testing); [c]continuous variables were centered at the mean values.

Anti-dsDNA, anti–double-stranded DNA; BMI, body mass index; BILAG, British Isles Lupus Assessment Group; C3/C4, complement 3/4; CNS, central nervous system; IV, intravenous; MMF, mycophenolate mofetil; OCS, oral corticosteroid; OR, odds ratio; PGA, Physician Global Assessment; SC, subcutaneous; SDI, Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR)

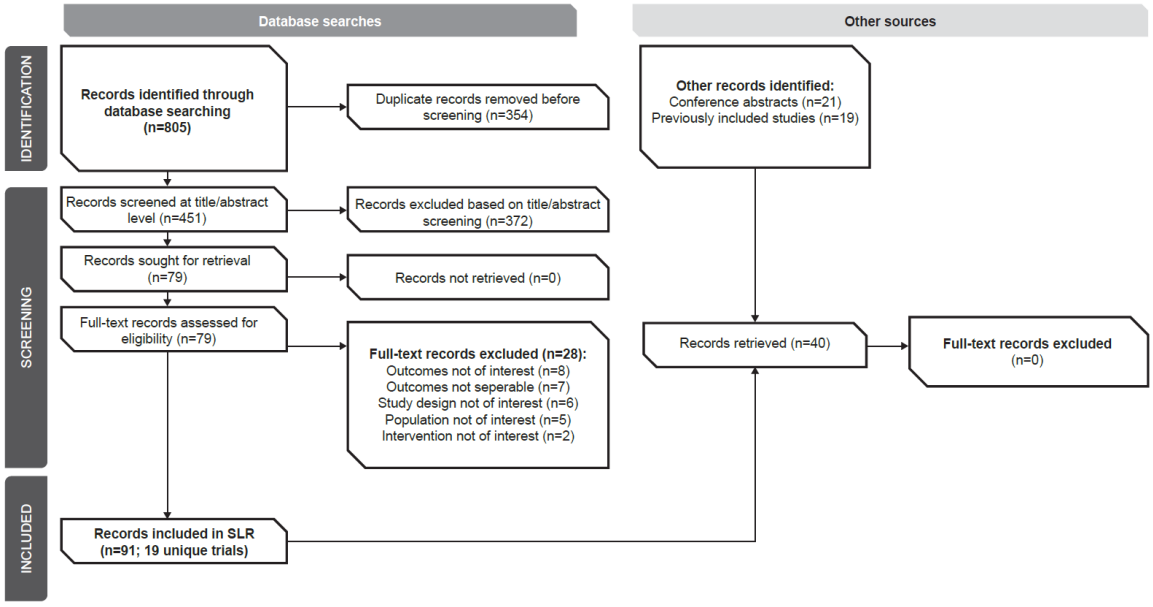
Damage Index; SELENA, Safety of Estrogens in Lupus National Assessment; SLE, systemic lupus erythematosus; SLEDAI, SLE Disease Activity Index; SRI-4, SLE Responder Index-4.

Appendix 5: SLR results and feasibility assessment findings

SLR results

The SLR searches identified 451 unique publications from electronic databases and 40 from other sources. Overall, 91 publications reporting on 19 unique trials were eligible for inclusion in the SLR. **Figure S1** summarizes the flow of included studies in the SLR.

Figure S1. PRISMA flow diagram



PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SLR, systematic literature review.

Ten trials were potentially eligible for quantitative synthesis.[16, 18-26] These included BLISS-52 [NCT00424476]; BLISS-76 [NCT00410384]; BLISS-SC [NCT01484496]; NEA study [NCT01345253]; EMBRACE [NCT01632241]; BASE [NCT01705977]; Wallace et al.[26] [NCT00071487]; TULIP-1 [NCT02446912]; TULIP-2 [NCT02446899]; MUSE [NCT01438489].

The nine trials not considered for quantitative synthesis were excluded for the following reasons:

- Study design not of interest (i.e., non-randomized or phase I randomized trials) (n=6)[27-32]
- Did not connect to any of the NMA networks (n=1)[33]
- Unapproved anifrolumab formulation (n=1)[34]
- Mandatory exposure to prior biologic therapy (rituximab) prior to randomization (n=1)[35]

Feasibility assessment

A feasibility assessment was undertaken to assess the two main assumptions of conducting NMA (i.e., homogeneity of included trials and transitivity). To this end, we comprehensively compared the included trials in terms of study design including inclusion/exclusion criteria and treatment implementation, outcomes' definition, and baseline patient characteristics (including potential imbalances in prognostic factors or treatment EMs).

Study design

All ten included studies were randomized, placebo-controlled, double-blind, multicenter trials. MUSE[36] and Wallace et al.[26] were phase II/IIb trials, BASE[25] was a phase IV trial, EMBRACE[24] was a phase III/IV trial while the remaining were phase III trials. Sample sizes

(i.e., overall number of randomized patients) ranged from 307 in MUSE to 4018 patients in BASE. The duration of follow-up ranged between 48 and 72 weeks. In terms of geographic location, all trials were multicenter with study sites in North America, Europe, Asia Pacific (including Australia), and Latin America. One trial was exclusively conducted in Asian centers with study sites in China, Japan, and South Korea[21] and another was exclusively conducted in North American centers (US and Canada).[26]

Important differences were identified in terms of the study design and eligibility criteria of the included trials:

- BASE[25] did not specify a minimum requirement for SELENA-SLEDAI score at enrollment. BLISS-SC[20], NEA study[21] and EMBRACE[24] enrolled patients with SELENA-SLEDAI score ≥8 at screening, Wallace et al.[26] enrolled patients with SELENA-SLEDAI score ≥4, while the remaining trials enrolled patients with SELENA–SLEDAI/SLEDAI-2K score of ≥6 at screening
- EMBRACE[24] was conducted only in patients of self-identified Black African ancestry race (US, Brazil, Columbia, France, South Africa, and UK)
- While all belimumab trials required patients to be seropositive (antinuclear antibody [ANA] titer ≥1:50 or anti-dsDNA ≥30 IU/ml), Wallace et al.[26] only required patients to have a history of measurable autoantibodies and patients were not required to be seropositive at screening

- The anifrolumab trials (TULIP-1, TULIP-2, and MUSE) included patients with BILAG-2004 organ domain scores of ≥1 A item or two B items and a PGA of disease activity score of ≥1. These criteria were not required in the belimumab trials
- While disease activity in the belimumab trials was measured using SELENA-SLEDAI, it was measured using SLEDAI-2K in anifrolumab trials. As per Gladman et al.,[37] the two definitions were considered comparable and the outcomes can be compared directly; however, a mapped version of SLEDAI-2K was used in the ITCs
- The anifrolumab trials (TULIP-1 and TULIP-2) required an attempt to taper OCS use between Weeks 8 or 12 and 40. This was not required in the belimumab trials

Given the requirement of seropositivity as part of belimumab label indication, the study by Wallace et al. was not eligible for inclusion in the ITCs.[26]

In addition to the approved doses of anifrolumab (i.e., 300 mg administered intravenously [IV]) and belimumab (i.e., 10 mg/kg IV and 200 mg administered subcutaneously), several trials assessed the efficacy and/or safety of unapproved doses of anifrolumab including MUSE (anifrolumab 1000 mg IV), BLISS-52 (belimumab 1 mg/kg IV), BLISS-76 (belimumab 1 mg/kg IV), and Wallace et al. 2009 (belimumab 1 mg/kg and 4 mg/kg IV). Arms of unapproved dose strengths of both agents were excluded.

Outcome definitions

The efficacy outcome definitions were similar across the trials, except for the definitions of steroid reduction from baseline, 4-point reduction in SLEDAI-2K, anti-dsDNA, and flares. While

the belimumab trials did not mandate or encourage steroid tapering, the TULIP-1 and TULIP-2 trials included a forced taper, where a steroid tapering attempt was required between Weeks 8 and 40. In addition, in the TULIP trials, the reduction in OCS was examined in the subgroup of patients who were receiving ≥10 mg/day OCS at baseline. Given the substantial methodological differences, the ITC of OCS reduction was deemed infeasible.

In the one anifrolumab trial (MUSE) that reported SLEDAI-2K (Clinical-SLEDAI) 4-point reduction, the outcome was calculated using the clinical components of the SLEDAI only (i.e., excluding the laboratory components for the immunologic domain variables of low complement and increased DNA binding). Thus, the ITC of this outcome was deemed infeasible. For the improvement in the specific organ domains, these were examined using the SELENA-SLEDAI in the belimumab trials and using the SLEDAI-2K in the anifrolumab trials. In addition, these analyses were conducted in the subgroup of patients with specific involvement of the corresponding organ domain at baseline; therefore, the distribution of baseline covariates in these subgroups is distinct from that of the overall intention-to-treat population. Hence, the analysis for SLEDAI organ domains was also deemed infeasible.

In the belimumab trials, flares were assessed using the SELENA-SLEDAI Flare Index, whereas the BILAG was used to examine flares in the anifrolumab trials. While both instruments are validated, the differences in definitions limit the comparability of flares incidence across trials. Reassessment of flares using BILAG in the belimumab trials could be undertaken to improve comparability in flare definition. However, the anifrolumab and belimumab studies used

different versions of the BILAG instrument (BILAG-2004 and BILAG-Classic, respectively). Given the substantial differences between the two versions, the ITC of flares was deemed infeasible.

Finally, the definition of anti-dsDNA positivity varied across trials. In the belimumab trials, the 30 IU/mL threshold was indicative of positive anti-dsDNA, whereas the 15 IU/mL threshold was used in the anifrolumab trials.

The BASE[25] trial only assessed safety endpoints and therefore was excluded from the quantitative synthesis due to the lack of efficacy outcomes of interest.

Therefore, the eight trials that were eligible for ITC of efficacy endpoints included the following:

- For belimumab: BLISS-52[18], BLISS-76[19], BLISS-SC[20], NEA study[21] and EMBRACE[24]
- For anifrolumab: TULIP-1[23, 38], TULIP-2[22], and MUSE[16]

Baseline patient characteristics

Table S3 summarizes the commonly reported patient baseline characteristics from the eight trials that were potentially eligible for inclusion in the ITCs. Across the eight trials eligible for ITCs, there were several differences that were noted. Focusing specifically on the likely treatment EMs, only small differences were noted for any OCS use and race. More substantial differences were noted for SLEDAI-2K, C3, C4, and anti-dsDNA. Data were not available for BMI or smoking status to allow for a comparison.

Lastly, while differences were identified in BILAG (proportion of patients with BILAG 1A/2B, proportion of patients with BILAG no A or B), it is difficult to decipher if this difference indicates a true difference in populations or it was just an artifact of the differences in instruments used across studies (the classic version of the BILAG was used in belimumab trials, whereas the 2004 version was used in anifrolumab trials).

Due to the differences in the baseline characteristics, particularly in those identified as EMs, it was concluded that conventional NMA was no longer feasible. Thus, PAICs were recommended.

Table S3. Baseline patient characteristics

Study	BLISS-52		BLISS-76		BLISS-SC		NEA		EMBRACE		Pooled BEL trials	TULIP-1		TULIP-2		MUSE		Pooled ANI trials
Treatment	PBO N=287	BEL 10 mg/kg IV N=290	PBO N=275	BEL 10 mg/kg IV N=273	PBO N=280	BEL 200 mg SC N=556	PBO N=226	BEL 10 mg/kg IV N=451	PBO N=149	BEL 10 mg/kg IV N=299	All arms N=3086	PBO N=184	ANI 300 mg IV N=180	PBO N=182	ANI 300 mg IV N=180	PBO N=102	ANI 300 mg IV N=99	All arms N=927
Age (years)																		
Mean (SD)	36.21 (11.8)	35.38 (10.8)	39.98 (11.9)	40.52 (11.2)	39.57 (12.6)	38.10 (12.1)	31.73 (9.2)	32.28 (9.7)	39.34 (12.2)	38.57 (11.1)	**36.97 (11.6)**	41.0 (12.3)	42.0 (12.0)	41.1 (11.5)	43.1 (12.0)	39.3 (12.9)	39.1 (11.9)	**41.2 (12.0)**
Sex, n (%)																		
Female	270 (94.1)	280 (96.6)	252 (91.6)	259 (94.9)	268 (95.7)	521 (93.7)	210 (92.9)	419 (92.9)	144 (96.6)	290 (97.0)	**2913 (94.4)**	171 (92.9)	165 (91.7)	170 (93.4)	168 (93.3)	93 (91.2)	93 (93.9)	**860 (92.7)**
Race, n (%)																		
White	82 (28.6)	71 (24.5)	188 (68.4)	189 (69.2)	166 (59.3)	335 (60.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	**1031 (33.4)**	137 (74.5)	125 (69.4)	107 (58.8)	110 (61.1)	41 (40.2)	35 (35.4)	**555 (59.9)**
Asian	105 (36.6)	116 (40.0)	11 (4.0)	11 (4.0)	63 (22.5)	119 (21.4)	225 (99.6)	450 (99.8)	0 (0.0)	0 (0.0)	**1100 (35.6)**	5 (2.7)	11 (6.1)	30 (16.5)	30 (16.7)	13 (12.7)	3 (3.0)	**92 (9.9)**
Black African ancestry	11 (4.0)	11 (4.0)	39 (14.2)	39 (14.3)	33 (11.8)	59 (10.6)	0 (0.0)	0 (0.0)	143 (96.0)	293 (98.0)	**619 (20.1)**	23 (13.0)	29 (16.0)	25 (13.7)	17 (9.4)	12 (11.8)	19 (19.2)	**125 (13.5)**
Others	89 (31.0)	92 (31.7)	37 (13.5)	36 (13.2)	21 (7.5)	47 (8.5)	1 (0.4)	1 (0.2)	6 (4.0)	6 (2.0)	**336 (10.9)**	19 (10.3)	15 (8.3)	20 (11.0)	23 (12.8)	36 (35.3)	42 (42.4)	**155 (16.7)**
BMI (kg/m^2)																		

Study	BLISS-52		BLISS-76		BLISS-SC		NEA		EMBRACE		Pooled BEL trials	TULIP-1		TULIP-2		MUSE		Pooled ANI trials
Treatment	**PBO N=287**	**BEL 10 mg/kg IV N=290**	**PBO N=275**	**BEL 10 mg/kg IV N=273**	**PBO N=280**	**BEL 200 mg SC N=556**	**PBO N=226**	**BEL 10 mg/kg IV N=451**	**PBO N=149**	**BEL 10 mg/kg IV N=299**	**All arms N=3086**	**PBO N=184**	**ANI 300 mg IV N=180**	**PBO N=182**	**ANI 300 mg IV N=180**	**PBO N=102**	**ANI 300 mg IV N=99**	**All arms N=927**
Mean (SD)	24.2 (4.6)	24.1 (4.8)	26.5 (5.9)	27.2 (7)	26.5 (7.2)	25.9 (6.3)	22.3 (4.0)	22.3 (3.4)	28.9 (6.9)	29.5 (7.4)	**25.4 (6.2)**	-	-	-	-	-	-	**27.6 (6.8)**
Disease duration (years)																		
Mean (SD)	5.93 (6.17)	5.03 (5.07)	7.42 (6.72)	7.20 (7.45)	6.80 (6.83)	6.37 (6.60)	5.97 (5.19)	6.07 (5.04)	6.86 (7.38)	7.26 (7.08)	**6.45 (6.36)**	-	-	-	-	7.55 (7.19)	7.99 (6.40)	-
Median (range)	3.9 (0.01-36.1)	3.6 (0.003-26.6)	5.8 (0.002-31.6)	4.7 (0.002-33.1)	4.6 (0.04-37.6)	4.3 (0.04-34.6)	4.7 (0.05-28.5)	5.0 (0.02-29.5)	3.8 (0.07-35.2)	5.02 (0.1-36.09)	**4.48 (0.0-37.6)**	6.6 (0.3-4)	7.3 (0.0-37.5)	6.5 (0.5-41.1)	7.9 (0.5-46.3)	-	-	**ANI = 7.1 (0-46.3) PBO = 6.3 (0.3-41.9)**
SELENA-SLEDAI, mean (SD)																		
Mean (SD)	9.70 (3.62)	9.97 (3.88)	9.78 (3.97)	9.51 (3.64)	10.33 (3.04)	10.47 (3.19)	10.15 (4.11)	9.85 (3.83)	10.17 (2.90)	9.94 (3.52)	**10.01 (3.60)**	-	-	-	-	-	-	-
SELENA-SLEDAI score ≥10, n (%)																		
Yes	158 (55.1)	160 (55.2)	140 (50.9)	136 (49.8)	168 (60.0)	352 (63.3)	124 (54.9)	233 (51.7)	90 (60.4)	153 (51.2)	**1714 (55.5)**	-	-	-	-	-	-	-
SLEDAI-2K, mean (SD)																		
Mean (SD)	10.0 (3.6)	10.4 (3.9)	10.0 (4.1)	9.7 (3.7)	10.5 (3.1)	10.9 (3.4)	10.8 (4.0)	10.6 (3.7)	10.5 (3.1)	10.2 (3.7)	**10.40 (3.7)**	11.5 (3.5)	11.3 (4.0)	11.5 (3.9)	11.4 (3.6)	11.1 (4.4)	10.7 (3.7)	**11.3 (3.8)**
SLEDAI-2K score ≥10, n (%)																		

Study	BLISS-52		BLISS-76		BLISS-SC		NEA		EMBRACE		Pooled BEL trials	TULIP-1		TULIP-2		MUSE		Pooled ANI trials
Treatment	PBO N=287	BEL 10 mg/kg IV N=290	PBO N=275	BEL 10 mg/kg IV N=273	PBO N=280	BEL 200 mg SC N=556	PBO N=226	BEL 10 mg/kg IV N=451	PBO N=149	BEL 10 mg/kg IV N=299	All arms N=3086	PBO N=184	ANI 300 mg IV N=180	PBO N=182	ANI 300 mg IV N=180	PBO N=102	ANI 300 mg IV N=99	All arms N=927
Yes	166 (57.8)	172 (59.3)	145 (52.7)	139 (50.9)	172 (61.4)	369 (66.4)	141 (62.4)	273 (60.5)	93 (62.4)	158 (52.8)	**1828 (59.2)**	135 (73.4)	125 (69.4)	131 (72.0)	129 (71.7)	-	-	-
PGA score																		
Mean (SD)	1.4 (0.5)	1.4 (0.5)	1.5 (0.5)	1.4 (0.5)	1.5 (0.5)	1.6 (0.4)	1.6 (0.4)	1.6 (0.5)	1.5 (0.5)	1.5 (0.5)	**1.5 (0.5)**	1.8 (0.4)	1.9 (0.4)	1.76 (0.40)	1.68 (0.41)	1.77 (0.44)	1.86 (0.39)	**1.79 (0.4)**
PGA score <1, n (%)[a]																		
Yes	43 (15.0)	32 (11.0)	33 (12.0)	51 (18.7)	19 (6.8)	40 (7.2)	8 (3.5)	26 (5.8)	15 (10.1)	40 (13.4)	**307 (9.9)**	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	**0 (0.0)**
BILAG Classic 1A/2B, n (%)																		
Yes	166 (57.8)	172 (59.3)	187 (68.0)	160 (58.6)	210 (75.0)	388 (69.8)	108 (47.8)	204 (45.2)	107 (71.8)	215 (71.9)	**1917 (62.1)**	-	-	-	-	-	-	-
BILAG 2004 1A/2B, n (%)																		
Yes[b]	-	-	-	-	-	-	-	-	-	-	-	184 (100.0)	180 (100.0)	182 (100.0)	180 (100.0)	102 (100.0)	99 (100.0)	**927 (100.0)**
BILAG No A or B, n (%)[a,c]																		
Yes	28 (9.8)	32 (11.0)	17 (6.2)	22 (8.1)	13 (4.6)	29 (5.2)	46 (20.4)	79 (17.5)	4 (2.7)	14 (4.7)	**284 (9.2)**	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	**0 (0.0)**
SDI, mean (SD)																		

Study	BLISS-52		BLISS-76		BLISS-SC		NEA		EMBRACE		Pooled BEL trials	TULIP-1		TULIP-2		MUSE		Pooled ANI trials
Treatment	PBO N=287	BEL 10 mg/kg IV N=290	PBO N=275	BEL 10 mg/kg IV N=273	PBO N=280	BEL 200 mg SC N=556	PBO N=226	BEL 10 mg/kg IV N=451	PBO N=149	BEL 10 mg/kg IV N=299	All arms N=3086	PBO N=184	ANI 300 mg IV N=180	PBO N=182	ANI 300 mg IV N=180	PBO N=102	ANI 300 mg IV N=99	All arms N=927
Mean (SD)	0.6 (0.9)	0.6 (1.0)	1.0 (1.5)	1.0 (1.4)	0.7 (1.2)	0.6 (1.0)	0.3 (0.6)	0.2 (0.6)	0.7 (1.0)	0.6 (1.0)	0.6 (1.05)	0.6 (1.0)	0.7 (1.2)	0.5 (0.8)	0.5 (0.9)	-	-	**0.6 (0.95)[d]**
SLEDAI organ domains,[e] n (%)																		
Musculoskeletal	165 (57.5)	174 (60.0)	207 (75.3)	194 (71.1)	218 (77.9)	438 (78.8)	75 (33.2)	139 (30.8)	115 (77.2)	235 (78.6)	**1960 (63.5)**	-	-	-	-	-	-	**684 (94.2)[f]**
Mucocutaneous	236 (82.2)	245 (84.5)	233 (84.7)	209 (76.6)	248 (88.6)	487 (87.6)	183 (81.0)	370 (82.0)	139 (93.3)	274 (91.6)	**2624 (85.0)**	-	-	-	-	-	-	**699 (96.3)[f]**
Immunological	234 (81.5)	248 (85.5)	205 (74.5)	206 (75.5)	211 (75.4)	427 (76.8)	202 (89.4)	410 (90.9)	106 (71.1)	197 (65.9)	**2446 (79.3)**	-	-	-	-	-	-	**467 (64.3)[f]**
Hematological	21 (7.3)	21 (7.2)	28 (10.2)	33 (12.1)	25 (8.9)	49 (8.8)	27 (11.9)	39 (8.6)	19 (12.8)	39 (13.0)	**301 (9.8)**	-	-	-	-	-	-	**73 (10.1)[f]**
Vascular	20 (7.0)	28 (9.7)	17 (6.2)	10 (3.7)	18 (6.4)	46 (8.3)	33 (14.6)	63 (14.0)	9 (6.0)	18 (6.0)	**262 (8.5)**	-	-	-	-	-	-	**79 (10.9)[f]**
Renal	81 (28.2)	76 (26.2)	46 (16.7)	46 (16.8)	50 (17.9)	102 (18.3)	102 (45.1)	206 (45.7)	34 (22.8)	55 (18.4)	**798 (25.9)**	-	-	-	-	-	-	**60 (8.3)[f]**
CNS	5 (1.7)	6 (2.1)	6 (2.2)	13 (4.8)	2 (0.7)	7 (1.3)	2 (0.9)	1 (0.2)	1 (0.7)	0 (0.0)	**43 (1.4)**	-	-	-	-	-	-	**4 (0.6)[f]**
Cardiovascular & Respiratory	14 (4.9)	10 (3.4)	18 (6.5)	27 (9.9)	18 (6.4)	29 (5.2)	3 (1.3)	2 (0.4)	12 (8.1)	23 (7.7)	**156 (5.1)**	-	-	-	-	-	-	**59 (8.1)[f]**

Study	BLISS-52		BLISS-76		BLISS-SC		NEA		EMBRACE		Pooled BEL trials	TULIP-1		TULIP-2		MUSE		Pooled ANI trials
Treatment	PBO N=287	BEL 10 mg/kg IV N=290	PBO N=275	BEL 10 mg/kg IV N=273	PBO N=280	BEL 200 mg SC N=556	PBO N=226	BEL 10 mg/kg IV N=451	PBO N=149	BEL 10 mg/kg IV N=299	All arms N=3086	PBO N=184	ANI 300 mg IV N=180	PBO N=182	ANI 300 mg IV N=180	PBO N=102	ANI 300 mg IV N=99	All arms N=927
Abnormal (low) complement concentration, n (%)																		
C3[g]	132 (46.0)	147 (50.7)	116 (42.2)	115 (42.1)	111 (39.6)	245 (44.1)	156 (69.0)	329 (72.9)	49 (32.9)	101 (33.8)	**1501 (48.6)**	65 (35.3)	58 (32.2)	72 (39.6)	72 (40.0)	43 (42.2)	28 (28.3)	**338 (36.5)**
C4[g]	160 (55.7)	180 (62.1)	143 (52.0)	147 (53.8)	71 (25.4)	146 (26.3)	73 (32.3)	131 (29.0)	31 (20.8)	53 (17.7)	**1135 (36.8)**	39 (21.2)	35 (19.4)	46 (25.3)	49 (27.2)	25 (24.5)	21 (21.2)	**215 (23.2)**
Abnormal anti-dsDNA, n (%)																		
Yes (≥30 IU/mL)	205 (71.4)	218 (75.1)	174 (63.2)	179 (65.6)	193 (68.9)	404 (72.7)	178 (75.8)	370 (82.0)	99 (66.4)	181 (60.5)	**2201 (71.3)**	-	-	-	-	-	-	**-**
Yes (>15 IU/mL)	287 (100.0)	290 (100.0)	275 (100.0)	273 (100.0)	211 (75.4)	435 (78.2)	195 (86.3)	397 (88.0)	105 (70.5)	200 (66.9)	**2668 (86.5)**	82 (44.6)	81 (45.0)	73 (40.1)	86 (47.8)	27 (26.5)	24 (24.2)	**373 (40.2)**
Mean (SD), U/mL	111.44 (75.0)	115.7 (73.4)	106.1 (46.6)	103.9 (73.9)	358.5 (843.9)	460.1 (1381.2)	303.24 (569.4)	430.4 (1420.8)	352.2 (921.1)	361 (1041.2)	**292.4 (945.7)**	-	-	-	-	-	-	**170.98 (431.5)[d]**
Immunosuppressive drug/immunomodulatory agents, n (%)																		
AZA	68 (23.7)	84 (29.0)	57 (20.7)	58 (21.2)	58 (20.7)	107 (19.2)	15 (6.6)	48 (10.6)	32 (21.5)	75 (25.1)	**602 (19.5)**	34 (18.5)	32 (17.8)	27 (14.8)	30 (16.7)	19 (18.6)	23 (23.2)	**165 (17.8)**
MTX/MTX sodium	35 (12.2)	20 (6.9)	60 (21.8)	39 (14.3)	39 (13.9)	52 (9.4)	15 (6.6)	29 (6.4)	23 (15.4)	45 (15.1)	**357 (11.6)**	38 (20.7)	22 (12.2)	35 (19.2)	34 (18.9)	16 (15.7)	19 (19.2)	**164 (17.7)**

Study	BLISS-52		BLISS-76		BLISS-SC		NEA		EMBRACE		Pooled BEL trials	TULIP-1		TULIP-2		MUSE		Pooled ANI trials
Treatment	PBO N=287	BEL 10 mg/kg IV N=290	PBO N=275	BEL 10 mg/kg IV N=273	PBO N=280	BEL 200 mg SC N=556	PBO N=226	BEL 10 mg/kg IV N=451	PBO N=149	BEL 10 mg/kg IV N=299	All arms N=3086	PBO N=184	ANI 300 mg IV N=180	PBO N=182	ANI 300 mg IV N=180	PBO N=102	ANI 300 mg IV N=99	All arms N=927
MMF	19 (6.6)	17 (5.9)	42 (15.3)	50 (18.3)	34 (12.1)	70 (12.6)	75 (33.2)	130 (28.8)	36 (24.2)	45 (15.1)	**518 (16.8)**	22 (12.0)	31 (17.2)	23 (12.6)	23 (12.8)	11 (10.8)	11 (11.1)	**121 (13.1)**
Any	122 (42.5)	123 (42.4)	154 (56)	148 (54.2)	137 (48.9)	244 (43.9)	146 (64.6)	292 (64.7)	88 (59.1)	167 (55.9)	**1621 (52.5)**	-	-	-	-	-	-	**445 (48.1)[d]**
Antimalarial (aminoquinoline) drug, n (%)																		
Yes	201 (70.0)	185 (63.8)	180 (65.5)	168 (61.5)	189 (67.5)	391 (70.3)	157 (69.5)	320 (71.0)	124 (83.2)	237 (79.3)	**2152 (69.7)**	134 (72.8)	124 (68.9)	133 (73.1)	119 (66.1)	75 (73.5)	76 (76.8)	**729 (70.7)**
Oral corticosteroid (prednisone or equivalent), n (%)																		
Yes	276 (96.2)	278 (95.9)	212 (77.1)	200 (73.3)	241 (86.1)	481 (86.5)	223 (98.7)	443 (98.2)	127 (85.2)	246 (82.3)	**2727 (88.4)**	153 (83.2)	150 (83.3)	151 (83.0)	141 (78.3)	88 (86.3)	79 (79.8)	**762 (82.2)**
Oral corticosteroid (≥10 mg/day)																		
Yes	190 (66.2)	201 (69.3)	125 (45.4)	117 (42.8)	164 (58.6)	330 (59.4)	181 (80.1)	344 (76.3)	94 (63.1)	183 (61.2)	**1929 (62.5)**	102 (55.4)	103 (57.2)	83 (45.6)	87 (48.3)	64 (63.4)	55 (55.6)	**494 (53.3)**

Cells filled in gray denote a large difference in baseline characteristics between BEL and ANI studies defined as >1 SD difference for continuous outcomes and >10% difference in any level of categorical outcomes.

[a]Based on inclusion criteria used in the anifrolumab trials; [b]the proportions of patients with BILAG 1A or 2B for the anifrolumab trials were based on the eligibility criteria of these trials, where patients were required to have severe disease activity in ≥1 domain or moderate activity in ≥2 domains (i.e., BILAG-2004 1A or 2B). However, in their review, Bruce et al.[39] reported that 94.4% of patients enrolled in the pooled TULIP-1 and TULIP-2 trials had BILAG ≥1A or ≥2B; [c]the classic version of the BILAG was used in BEL trials, whereas the 2004 version was used in ANI trials; [d]pooled results were derived from the Tummala et al.[40] pooled analysis of MUSE, TULIP-1, and TULIP-2 (N=925); [e]the organ domains were examined using the SELENA-SLEDAI in belimumab trials and the SLEDAI-2K in the anifrolumab trials; [f]results correspond to the pooled TULIP-1 and TULIP-2 trials only (N=726); [g]low C3: <90 mg/dL, low C4: <10 mg/dL for BLISS-SC, NEA, EMBRACE and ANI studies, and <16 mg/dL for BLISS-76 and BLISS-52.

ANI, anifrolumab; anti-dsDNA, anti–double-stranded DNA antibody; AZA, azathioprine; BEL, belimumab; BILAG, British Isles Lupus Assessment Group; BMI, body mass index; C3/C4, complement 3/4; CNS, central nervous system; IV, intravenous; MMF, mycophenolate mofetil; MTX, methotrexate; PBO, placebo; PGA, Physician Global Assessment; SC, subcutaneous; SD, standard deviation; SDI, Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index; SELENA-SLEDAI, Safety of Estrogens in Lupus Erythematosus National Assessment-SLE Disease Activity Index; SLE, systemic lupus erythematosus; SLEDAI-2K, SLE Disease Activity Index 2000.

Appendix 6: Detailed analysis results

Table S4. BEL versus ANI OR for all SRI-4 analyses[a]

Analysis	Method	Evidence base	Sample size of BEL trials (and ESS of IPD in MAICs)	Network structure (key provided below)	Adjustment approach (key provided below)	BEL versus ANI OR (95% CrI/CI)[b]	Probability best	SUCRA
Base-case	ML-NMR	All 8 trials	3080	1	1	1.04 (0.74 to 1.45)	ANI = 0.42 BEL = 0.58	ANI = 71% BEL = 79%
Sensitivity 1	ML-NMR	All 8 trials	3080	2	2	1.05 (0.75 to 1.49)	ANI = 0.39 BEL = 0.61	ANI = 69% BEL = 81%
Sensitivity 2	MAIC	All 8 trials	3080 (1531.3)	3	2	1.12 (0.80 to 1.56)	NA	NA

Sensitivity 3	ML-NMR	All 8 trials	3078	2	3	1.02 (0.72 to 1.44)	ANI = 0.45 BEL = 0.55	ANI = 72% BEL = 78%
Sensitivity 4	STC	All 8 trials	3078	3	3	1.10 (0.79 to 1.53)	NA	NA
Supplementary 1	NMA	All 8 trials	3080	1	No adjustments	1.13 (0.83 to 1.53)	NA	NA
Based on SELENA-SLEDAI in BEL trials	ML-NMR	All 8 trials	3078	1	3	0.97 (0.69 to 1.37)	ANI = 0.57 BEL = 0.43	ANI = 0.78 BEL = 0.72
BEL SC and BEL IV split	ML-NMR	All 8 trials	3078	4	3	IV = 0.99 (0.69 to 1.44) SC = 1.08 (0.71 to 1.66)	ANI = 0.27 BEL IV = 0.21 BEL SC = 0.52	ANI = 62% BEL IV = 60% BEL SC = 77%

Exploratory 1	STC	BLISS-52, BLISS-76, TULIP-1, TULIP-2	1125	5	4	1.06 (0.65 to 1.72)	NA	NA
Exploratory 2	MAIC	BLISS-52, BLISS-76, TULIP-1, TULIP-2	1125 (350.7)	5	4	1.11 (0.66 to 1.86)	NA	NA

[a]The SRI-4 results from the BEL trials incorporated a modified version of SLEDAI-2K, unless otherwise specified; [b]CrIs in ML-NMR and NMA, confidence intervals in STC and MAIC.

ANI, anifrolumab; BEL, belimumab; CI, confidence interval; CrI, credible interval; ESS, effective sample size; IPD, individual patient data; IV, intraveous; MAIC, matching-adjusted indirect comparison; ML-NMR, multi-level network meta-regression; NA, not applicable; NMA, network meta-analysis; OR, odds ratio; PBO, placebo; SC, subcutaneous; SELENA-SLEDAI, Safety of Estrogens in Lupus Erythematosus National Assessment-SLE Disease Activity Index; SLE, systemic lupus erythematosus; SRI-4, SLE Responder Index-4; STC, simulated treatment comparison; SUCRA, surface under the cumulative ranking curve.

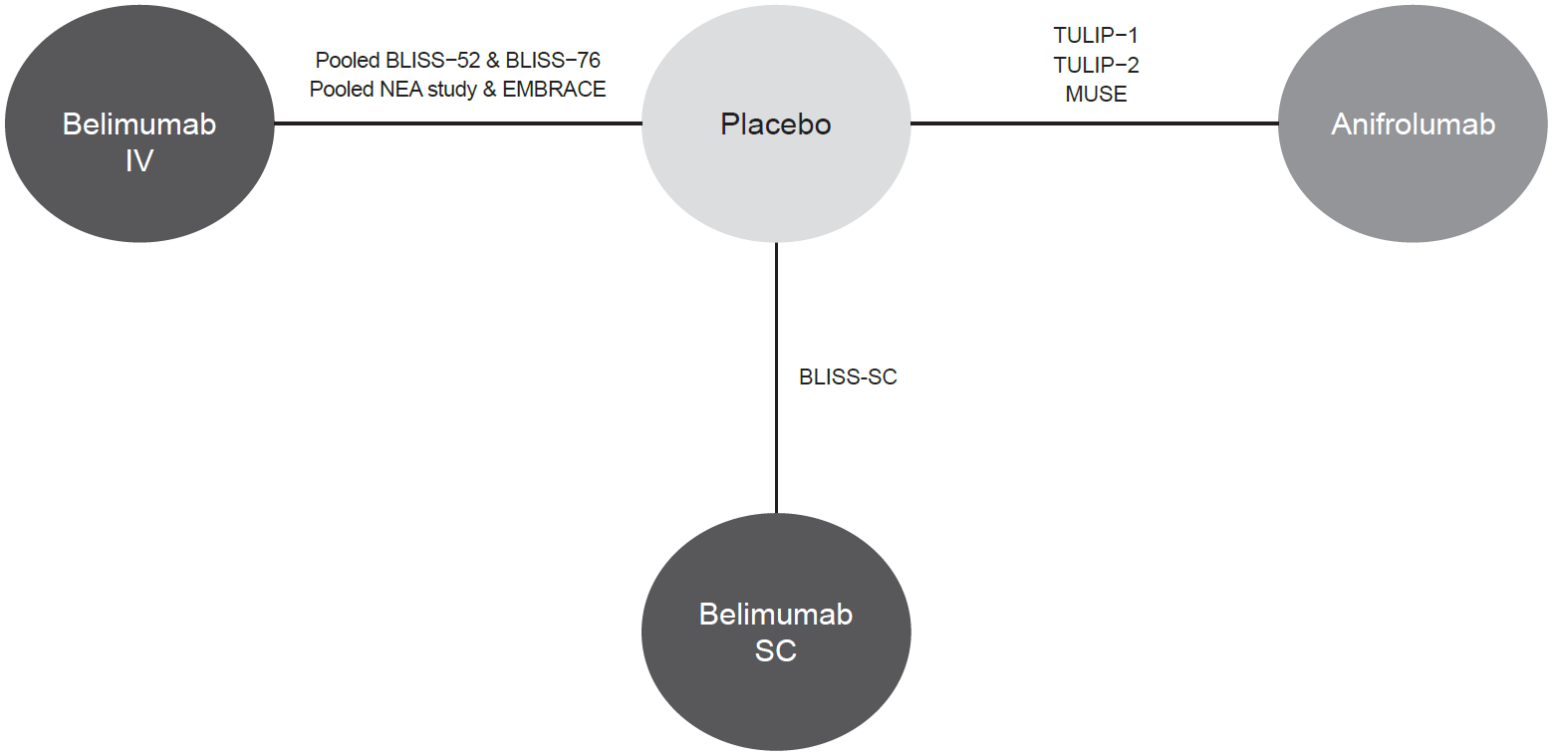
Network structure key:

1. Three treatment nodes in network: belimumab, anifrolumab, placebo. All eight trials; each incorporated separately

2. Three treatment nodes in network: belimumab, anifrolumab, placebo. All eight trials; each anifrolumab trial incorporated separately, while five belimumab trials treated as three trials

3. Three treatment nodes in network: belimumab, anifrolumab, placebo. All eight trials; three anifrolumab trials treated as a single trial and five belimumab trials treated as a single trial

4. Four treatment nodes in network: belimumab IV, belimumab SC, anifrolumab, placebo. All eight trials; each anifrolumab trial incorporated separately while five belimumab trials treated as three trials



IV, intravenous; SC, subcutaneous.

5. Three treatment nodes in network: belimumab, anifrolumab, placebo. Four trials, two anifrolumab trials treated as a single trials and two belimumab trials treated as a single trial

Adjustment approach key:

1. Four imbalanced EMs including baseline (1) SLEDAI-2K; (2) low C3; (3) low C4; (4) anti-dsDNA positive
2. Six EMs including baseline (1) SLEDAI-2K; (2) low C3; (3) low C4; (4) anti-dsDNA positive; (5) OCS use; (6) Black African ancestry
3. Six EMs including (1) SLEDAI-2K; (2) low C3; (3) low C4; (4) anti-dsDNA positive; (5) OCS use; (6) Black African ancestry and two prognostic factors including (1) Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index (SDI) and (2) immunosupressant use

4. Twelve EMs including (1) female sex; (2) White race; (3) age; (4) SLEDAI-2K or Safety of Estrogens in Lupus National Assessment (SELENA)-SLEDAI; (5) BILAG 1A or 2B; (6) low C3; (7) low C4; (8) anti-dsDNA positive; (9) azathioprine use; (10) methotrexate use; (11) mycophenolate use; (12) OCS use of ≥7.5 mg/day

Table S5. Regression estimates for ML-NMR and STC population-adjustments for base-case and sensitivity analyses

Regression estimates Mean (lower CrI to upper CrI) (in linear scale)	Base-case	Sensitivity 1	Sensitivity 3	Sensitivity 4[a]	Sensitivity with SLEDAI	Sensitivity with IV and SC separation
Anifrolumab (reference = placebo)	0.54 (0.25 to 0.82)	0.52 (0.23 to 0.82)	0.54 (0.25 to 0.85)	NA	0.54 (0.25 to 0.84)	0.55 (0.25 to 0.85)
Belimumab (reference = placebo)	0.57 (0.41 to 0.73)	0.58 (0.42 to 0.73)	0.56 (0.41 to 0.72)	0.514 (0.31 to 0.72)	0.51 (0.35 to 0.67)	IV = 0.54 (0.36 to 0.72) SC = 0.63 (0.32 to 0.94)
SLEDAI-2K (continuous)	1 (0.64 to 1.36)	0.99 (0.63 to 1.35)	1.08 (0.72 to 1.45)	1.069 (0.71 to 1.43)	0.99 (0.61 to 1.36)	1.09 (0.72 to 1.46)
Low C3 (reference = “no”)	-0.44 (-0.71 to -0.16)	-0.44 (-0.73 to -0.15)	-0.46 (-0.74 to -0.18)	-0.45 (-0.73 to -0.16)	-0.41 (-0.7 to -0.12)	-0.46 (-0.75 to -0.17)
Low C4 (reference = “no”)	-0.55 (-0.85 to -0.26)	-0.54 (-0.82 to -0.25)	-0.56 (-0.85 to -0.27)	-0.57 (-0.84 to -0.29)	-0.59 (-0.89 to -0.31)	-0.56 (-0.85 to -0.27)
Black African ancestry (reference = “no”)	NA	-0.16 (-0.48 to 0.15)	-0.09 (-0.42 to 0.23)	-0.126 (-0.44 to 0.19)	-0.09 (-0.41 to 0.23)	-0.1 (-0.43 to 0.23)
Anti-dsDNA >30 IU/mL (reference = “no”)	-0.27 (-0.55 to 0.01)	-0.24 (-0.53 to 0.05)	-0.27 (-0.56 to 0.02)	-0.289 (-0.57 to 0)	-0.24 (-0.53 to 0.04)	-0.27 (-0.56 to 0.03)
Any OCS use (reference = “no”)	NA	0.04 (-0.33 to 0.42)	-0.05 (-0.43 to 0.34)	-0.05 (-0.43 to 0.33)	0.01 (-0.36 to 0.4)	-0.05 (-0.43 to 0.33)
Any immunosuppressant use (reference = “no”)	NA	NA	-0.17 (-0.32 to -0.01)	-0.20 (-0.28 to -0.13)	-0.22 (-0.37 to -0.07)	-0.17 (-0.32 to -0.01)

SDI (continuous)	NA	NA	-0.21 (-0.29 to -0.14)	-0.18 (-0.33 to -0.03)	-0.21 (-0.28 to -0.13)	-0.21 (-0.29 to -0.14)
Interaction effect: treatment and SLEDAI-2K	0.28 (-0.18 to 0.74)	0.32 (-0.14 to 0.78)	0.3 (-0.18 to 0.78)	0.342 (-0.13 to 0.81)	0.32 (-0.15 to 0.79)	0.29 (-0.19 to 0.78)
Interaction effect: treatment and low C3	0.35 (-0.01 to 0.7)	0.31 (-0.05 to 0.67)	0.29 (-0.07 to 0.66)	0.252 (-0.11 to 0.61)	0.24 (-0.12 to 0.6)	0.3 (-0.07 to 0.66)
Interaction effect: Treatment and low C4	0.37 (0.01 to 0.72)	0.34 (-0.02 to 0.7)	0.35 (-0.01 to 0.71)	0.333 (-0.02 to 0.69)	0.46 (0.1 to 0.81)	0.36 (0 to 0.73)
Interaction effect: treatment and Black African ancestry	NA	-0.29 (-0.68 to 0.1)	-0.3 (-0.69 to 0.1)	-0.314 (-0.7 to 0.08)	-0.21 (-0.6 to 0.18)	-0.28 (-0.67 to 0.11)
Interaction effect: treatment and anti-dsDNA	0.05 (-0.31 to 0.42)	0 (-0.38 to 0.37)	0.04 (-0.33 to 0.4)	0.075 (-0.29 to 0.44)	-0.01 (-0.38 to 0.36)	0.03 (-0.34 to 0.4)
Interaction effect: treatment and OCS use	NA	0.36 (-0.12 to 0.83)	0.44 (-0.05 to 0.92)	0.436 (-0.04 to 0.92)	0.39 (-0.11 to 0.88)	0.44 (-0.04 to 0.93)

[a]The intervals for the STC in Sensitivity 4 are confidence intervals.

Anti-dsDNA, anti–double-stranded DNA antibody; C3/4, complement 3/4; CrI, credible interval; IV, intraveous; ML-NMR, multi-level network meta-regression; NA, not applicable; OCS, oral corticosteroid; SC, subcutaneous; SDI, Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index; SLEDAI-2K, SLE Disease Activity Index-2000; SLE, systemic lupus erythematosus; STC, simulated treatment comparison.

Table S6. Distribution of re-scaled weights of 3080 patients from BEL trials used in MAIC sensitivity analysis 2

Percentile	**Weight**
0%	0.1
1%	0.15
10%	0.28
25%	0.39
50%	0.62
75%	1.28
90%	2.25
99%	4.35
100%	20.46

BEL, belimumab; MAIC, matching-adjusted indirect comparison.

REFERENCES

1 Dias S, Sutton AJ, Ades AE, et al. 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. doi: 10.1177/0272989X12458724.

2 Salanti G, Higgins JP, Ades AE, et al. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17:279-301. doi: 10.1177/0962280207080643.

3 Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health* 2014;17:157-73. doi: 10.1016/j.jval.2014.01.004.

4 Song F, Altman DG, Glenny AM, et al. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003;326:472. doi: 10.1136/bmj.326.7387.472.

5 Dias S, Ades AE. Absolute or relative effects? Arm-based synthesis of trial data. *Res Synth Methods* 2016;7:23-8. doi: 10.1002/jrsm.1184.

6 Cochrane. Cochrane Handbook for Systematic Reviews of Interventions. Available from: https://training.cochrane.org/handbook/current Accessed: October 2, 2022.

7 Phillippo DM, Ades AE, Dias S, et al. Methods for Population-Adjusted Indirect Comparisons in Health Technology Appraisal. *Med Decis Making* 2018;38:200-11. doi: 10.1177/0272989X17725740.

8 Phillippo DM, Dias S, Ades AE, et al. Multilevel network meta-regression for population-adjusted treatment comparisons. *J R Stat Soc Ser A Stat Soc* 2020;183:1189-210. doi: 10.1111/rssa.12579.

9 Phillippo DM, Dias S, Ades AE, et al. Assessing the performance of population adjustment methods for anchored indirect comparisons: A simulation study. *Stat Med* 2020;39:4885-911. doi: 10.1002/sim.8759.

10 Phillippo DM, Ades T, Dias S, et al. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Available from: https://research-information.bris.ac.uk/ws/portalfiles/portal/94868463/Population_adjustment_TSD_FINAL.pdf Accessed: October, 2022.

11 Phillippo DM. multinma: Bayesian Network Meta-Analysis of Individual and Aggregate Data. R package version 0.5.0. Available from: https://zenodo.org/record/7033094#.Y8bhAHb7Q2w Accessed: October 2, 2022.

12 Signorovitch JE, Wu EQ, Yu AP, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. *Pharmacoeconomics* 2010;28:935-45. doi: 10.2165/11538370-000000000-00000.

13 Bucher HC, Guyatt GH, Griffith LE, et al. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology* 1997;50:683-91. doi: 10.1016/s0895-4356(97)00049-8.

14 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* 2022;11:765-77. doi: 10.2217/cer-2022-0040.

15 Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71.

16 Furie R, Khamashta M, Merrill JT, et al. Anifrolumab, an Anti-Interferon-alpha Receptor Monoclonal Antibody, in Moderate-to-Severe Systemic Lupus Erythematosus. *Arthritis Rheumatol* 2017;69:376-86. doi: 10.1002/art.39962.
17 Wan X, Wang W, Liu J, et al. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 2014;14:135. doi: 10.1186/1471-2288-14-135.
18 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 3 trial. *Lancet* 2011;377:721-31. doi: 10.1016/S0140-6736(10)61354-2.
19 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* 2011;63:3918-30. doi: 10.1002/art.30613.
20 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* 2017;69:1016-27. doi: 10.1002/art.40049.
21 Zhang F, Bae SC, Bass D, et al. A pivotal phase III, randomised, placebo-controlled study of belimumab in patients with systemic lupus erythematosus located in China, Japan and South Korea. *Ann Rheum Dis* 2018;77:355-63. doi: 10.1136/annrheumdis-2017-211631.
22 Morand EF, Furie R, Tanaka Y, et al. Trial of Anifrolumab in Active Systemic Lupus Erythematosus. *N Engl J Med* 2020;382:211-21. doi: 10.1056/NEJMoa1912196.
23 Furie RA, Morand EF, Bruce IN, et al. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. *Lancet Rheumatol* 2019;1:E208-E19. doi: https://doi.org/10.1016/S2665-9913(19)30076-1.
24 Ginzler E, Guedes Barbosa LS, D'Cruz D, et al. Phase III/IV, Randomized, Fifty-Two-Week Study of the Efficacy and Safety of Belimumab in Patients of Black African Ancestry With Systemic Lupus Erythematosus. *Arthritis Rheumatol* 2022;74:112-23. doi: 10.1002/art.41900.
25 Sheikh SZ, Scheinberg MA, Wei JC-C, et al. Mortality and adverse events of special interest with intravenous belimumab for adults with active, autoantibody-positive systemic lupus erythematosus (BASE): a multicentre, double-blind, randomised, placebo-controlled, phase 4 trial. *Lancet Rheumatol* 2021;3:e122-e30. doi: 10.1016/s2665-9913(20)30355-6.
26 Wallace DJ, Stohl W, Furie RA, et al. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. *Arthritis Rheum* 2009;61:1168-78. doi: 10.1002/art.24699.
27 Furie R, Stohl W, Ginzler EM, et al. Biologic activity and safety of belimumab, a neutralizing anti-B-lymphocyte stimulator (BLyS) monoclonal antibody: a phase I trial in patients with systemic lupus erythematosus. *Arthritis Res Ther* 2008;10:R109. doi: 10.1186/ar2506.
28 Yamada M, Akita M, Nakagawa T, et al. Safety, tolerability, pharmacokinetics and pharmacodynamics of belimumab in Japanese patients with mild-to-moderate systemic lupus erythematosus. *J Drug Assess* 2013;2:40-8. doi: 10.3109/21556660.2013.792823.
29 Ceccarelli F, Cipriano E, Natalucci F, et al. Belimumab is Able to Induce a Significant Improvement of Joint Activity Status in Patients Diagnosed with Systemic Lupus Erythematosus: Results From a 12-Month Longitudinal Study. *Isr Med Assoc J* 2020;22:415-19. doi: https://www.ncbi.nlm.nih.gov/pubmed/33236565
30 Tanaka Y, Takeuchi T, Okada M, et al. Safety and tolerability of anifrolumab, a monoclonal antibody targeting type I interferon receptor, in Japanese patients with systemic lupus erythematosus: A multicenter, phase 2, open-label study. *Mod Rheumatol* 2020;30:101-08. doi: 10.1080/14397595.2019.1583833.
31 Struemper H, Kurtinecz M, Edwards L, et al. Reductions in circulating B cell subsets and immunoglobulin G levels with long-term belimumab treatment in patients with SLE. *Lupus Sci Med* 2022;9:e000499. doi: 10.1136/lupus-2021-000499.

32 Tanaka Y, Bae SC, Bass D, et al. Long-term open-label continuation study of the safety and efficacy of belimumab for up to 7 years in patients with systemic lupus erythematosus from Japan and South Korea. *RMD Open* 2021;7:e001629. doi: 10.1136/rmdopen-2021-001629.

33 Bae S, Dimelow R, Ji B., Kurrasch R, et al. Results of the Open-label, Non-randomized 52-Week Study to Evaluate Treatment Holidays and Rebound Phenomenon After Treatment with Belimumab in Patients with SLE [abstract]. Paper presented at: ACR/ARP Annual Meeting; November 12, 2019, 2019; Atlanta, GA.

34 Bruce I, Nami A, Schwetje E, et al. Pharmacokinetics, pharmacodynamics, and safety of subcutaneous anifrolumab in patients with systemic lupus erythematosus, active skin disease, and high type I interferon gene signature: a multicentre, randomised, double-blind, placebo-controlled, phase 2 study. *Lancet Rheumatol* 2020;3:E101-E10. doi: 10.1016/S2665-9913(20)30342-8.

35 Shipa M, Embleton-Thirsk A, Parvaz M, et al. Effectiveness of Belimumab After Rituximab in Systemic Lupus Erythematosus : A Randomized Controlled Trial. *Ann Intern Med* 2021;174:1647-57. doi: 10.7326/M21-2078.

36 Furie R, Kalunian K, Merrill JT, et al. Lupus Disease Activity After Cessation of Anifrolumab Treatment During the Phase 2b MUSE Trial Follow-up Period [abstract]. ACR Convergence; November 9, 2020, 2020; Virtual.

37 Gladman DD, Ibanez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol* 2002;29:288-91. doi: https://www.ncbi.nlm.nih.gov/pubmed/11838846

38 Furie RA, Morand EF, Bruce IN, et al. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. *Lancet Rheumatol* 2019;1:e208-e19. doi: 10.1016/S2665-9913(19)30076-1.

39 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* 2022. doi: 10.2217/cer-2022-0040.

40 Tummala R, Abreu G, Pineda L, et al. Safety profile of anifrolumab in patients with active SLE: an integrated analysis of phase II and III trials. *Lupus Sci Med* 2021;8:e000464. doi: 10.1136/lupus-2020-000464.