Cureus
Part of SPRINGER NATURE

Open Access Review Article

The Role of Biologic Therapies (Belimumab and Anifrolumab) in Systemic Lupus Erythematosus (SLE) Disease Activity and Renal Outcomes in Predominantly Young Female Trial Populations: A Systematic Review and Meta-Analysis

Priyanka Hayre [1]

1. Internal Medicine, University of Medicine and Social Sciences, Kharkiv, UKR

Corresponding author: Priyanka Hayre, hayre.priyanka2@gmail.com

Review began 11/01/2026
Review ended 11/07/2026
Published 11/12/2026

© **Copyright** 2026
Hayre. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.119544

Abstract

Systemic lupus erythematosus (SLE) mainly affects women during their reproductive years, and flares of disease activity, lupus nephritis, and cumulative glucocorticoid exposure shape both long-term organ damage and family planning. Belimumab, which neutralizes soluble B-lymphocyte stimulator, and anifrolumab, which blocks the type I interferon receptor, are the two biologic agents approved for SLE. A systematic review and meta-analysis of randomized, placebo-controlled trials of intravenous belimumab or anifrolumab added to standard therapy, following PRISMA 2020 and searching five sources from inception to 2026, was conducted. Eight trials (BLISS-52, BLISS-76, BLISS-LN, EMBRACE, MUSE, TULIP-1, TULIP-2, and TULIP-LN; 3,854 randomized participants) were included. Pooling the global composite responder endpoint at week 52 across six trials (SRI-4, SRI-SLEDAI-2K, or BICLA; non-interchangeable definitions) gave an odds ratio (OR) of 1.54 (95% confidence interval (CI) 1.18 to 2.02; I^2 = 62%) favoring the biologic, with numerically similar point estimates for belimumab (1.55; 95% CI 1.27 to 1.91) and anifrolumab (1.59; 95% CI 0.82 to 3.09) and no statistically detectable difference between agents (P = 0.95); this indirect, underpowered comparison does not establish equivalent efficacy. Heterogeneity within the anifrolumab trials was substantial (I^2 = 83%) and was driven largely by the neutral TULIP-1 result. Complete renal response favored the biologic in lupus nephritis (OR 1.73, 95% CI 1.17 to 2.55; two trials; Hartung-Knapp interval 0.14 to 21.44, so the estimate is interpreted descriptively). Herpes zoster was more frequent with anifrolumab (OR 4.09). The trials enrolled 83% to 97% women with mean ages of 33 to 42 years. Still, no primary publication reported outcomes for women aged 18 to 49 years, so the findings describe predominantly young female trial populations and apply to this age group only indirectly; Hartung-Knapp intervals were wide, and few trials contributed to each comparison. In these predominantly young female trial populations, both agents added measurable benefit to standard therapy, although the trials were not designed to compare them; evidence specific to young women, including pregnancy, remains the principal gap.

Categories: Rheumatology, Internal Medicine, Nephrology
Keywords: belimumab, lupus nephritis, meta-analysis, reproductive-age women, systemic lupus erythematosus

Introduction And Background

Systemic lupus erythematosus (SLE) is a chronic, multisystem autoimmune disease with a marked female predominance, and onset is most frequent between puberty and the fifth decade [1,2]. Reported incidence ranges from 0.3 to 23.2 per 100,000 person-years, and prevalence ranges from 6.5 to 178.0 per 100,000 population across regions, with higher rates in women and in people of non-European ancestry [3]. For young women, the consequences extend beyond the disease itself. Cumulative glucocorticoid exposure contributes to organ damage; contraception and fertility treatment must be chosen around disease activity and antiphospholipid antibodies, and active disease at conception or during gestation raises the risk of maternal and fetal complications [4-6]. Throughout this review, "reproductive age" refers to women aged 18 to 49 years.

Lupus nephritis, inflammation of the kidney caused by SLE, develops in roughly a quarter to well over half of patients with SLE, depending on the population studied, and it remains the manifestation most closely tied to kidney failure and premature death [7-9]. A renal flare in a woman of childbearing age carries particular weight: induction regimens that contain cyclophosphamide or mycophenolate require reliable contraception because of teratogenic risk, and nephritis activity at the time of conception predicts adverse pregnancy outcomes [4,6,10]. Reducing flare frequency and the glucocorticoid dose needed to control them is therefore a clinical priority in this age group.

Two biologic agents are now used routinely, and they act on distinct immunologic pathways. Belimumab is a human IgG1λ monoclonal antibody (immunoglobulin G1 with a lambda light chain) that binds soluble B-lymphocyte stimulator (BLyS; also called B-cell activating factor, BAFF), a cytokine of the tumor necrosis

factor family that is often elevated in SLE. By preventing BLyS from engaging its receptors on B cells, belimumab reduces the survival signals available to B cells, including autoreactive clones. It is thought to limit their maturation into autoantibody-producing plasma cells [11]. Anifrolumab is a human IgG1κ monoclonal antibody directed against subunit 1 of the type I interferon receptor (IFNAR1), and so blocks signaling by all type I interferons, which share this single receptor [12]. Type I interferon signaling promotes dendritic-cell activation, B-cell differentiation, and autoantibody production, and an interferon-inducible gene signature is present in a large proportion of adults with SLE [12]. About three-quarters of patients in the phase 2 anifrolumab program had a high type I interferon gene signature at baseline [13].

Belimumab was tested in two phase 3 trials that used the SLE Responder Index (SRI-4) at week 52 as the primary endpoint. This composite requires at least a 4-point reduction in the SLE Disease Activity Index (SLEDAI) without worsening in other organ systems or in the physician's global assessment [14-16], then in a trial restricted to patients who self-identified as Black [17] and in a two-year trial in active lupus nephritis [18]. Anifrolumab progressed from a phase 2 study [13] to two phase 3 trials, TULIP-1 and TULIP-2, which used different primary endpoints and reached different conclusions [19,20], and to a phase 2 trial in proliferative lupus nephritis [21]. The trials differ in responder definition (SRI-4, BILAG-based Composite Lupus Assessment (BICLA), which is derived from the British Isles Lupus Assessment Group (BILAG) index and credits partial improvement across organ systems, renal response), population, and follow-up, and no trial reported outcomes separately for women of reproductive age.

Despite the concentration of SLE in women of reproductive age, the evidence for these agents has a specific gap. The pivotal trials enrolled adults aged 18 years or older across a broad age range, and none published efficacy or safety results stratified by age band (for example, 18 to 49 years versus 50 years or older) or by sex. Pregnancy was an exclusion criterion, and contraception was required. Clinicians counseling women aged 18 to 49 years must therefore extrapolate from whole-trial results, in which participants were predominantly female and relatively young but also included older women and some men. This age group has not been tested directly for differences in benefit magnitude, glucocorticoid-sparing effect, or safety profile.

Current European recommendations support both agents as add-on therapy for patients whose disease remains active despite hydroxychloroquine and glucocorticoids [22]. Clinicians counseling young women need to know how large and how consistent the benefit is, how it compares across the two mechanisms, and how much of the evidence applies to this group. Therefore, a systematic review and meta-analysis of randomized placebo-controlled trials of belimumab and anifrolumab was conducted to estimate their effects on global disease activity, renal response, glucocorticoid reduction and safety, and to appraise the applicability of the evidence to women of reproductive age. Because age-stratified data are unavailable, it was not claimed to estimate effects within an isolated 18-to-49-year subgroup; instead, how closely the trial populations approximate this group was described and the pooled estimates were interpreted accordingly.

Review

Methods

Protocol and Reporting

The review followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 statement [23]. The review question and outcomes were defined before data extraction. The protocol was not registered prospectively. The analysis used published trial-level data only; ethical approval was not required.

Search Strategy

Five sources, PubMed/MEDLINE, Embase (Ovid), the Cochrane Central Register of Controlled Trials (CENTRAL), Scopus, and ClinicalTrials.gov, from database inception to 10 October 2026 were searched. Strategies were built from the review question and tested to retrieve known trials of each agent. No language restriction was applied. Each strategy combined three concept blocks (disease, intervention, and trial design) with AND; within each block, controlled vocabulary and free-text terms, including brand names and development codes, were combined with OR. Strings were adapted to the field tags, truncation symbols, and controlled vocabulary of each database. No age or sex filter was applied so that trials that did not report these characteristics in their titles or abstracts were not missed. Reference lists of included trials were screened manually. The full strings are as follows.

PubMed/MEDLINE: ("Lupus Erythematosus, Systemic" [Mesh] OR "systemic lupus erythematosus" [tiab] OR "lupus nephritis" [Mesh] OR "lupus nephritis" [tiab] OR SLE [tiab]) AND (belimumab [tiab] OR Benlysta [tiab] OR HGS1006 [tiab] OR LymphoStat-B [tiab] OR anifrolumab [tiab] OR Saphnelo [tiab] OR MEDI-546 [tiab]) AND (randomized controlled trial [pt] OR randomized [tiab] OR randomised [tiab] OR placebo [tiab] OR double-blind [tiab])

Embase (Ovid): ('systemic lupus erythematosus'/exp OR 'lupus nephritis'/exp OR 'systemic lupus

erythematosus':ti, ab OR 'lupus nephritis':ti, ab) AND ('belimumab'/exp OR belimumab: ti, ab OR benlysta: ti, ab OR 'anifrolumab'/exp OR anifrolumab: ti, ab OR saphnelo: ti, ab) AND ('randomized controlled trial'/exp OR random*:ti, ab OR placebo*:ti, ab OR 'double blind*':ti, ab)

CENTRAL (Cochrane Library): #1 MeSH descriptor: [Lupus Erythematosus, Systemic] explode all trees; #2 lupus: ti, ab, kw; #3 MeSH descriptor: [Lupus Nephritis] explode all trees; #4 #1 OR #2 OR #3; #5 (belimumab OR benlysta OR anifrolumab OR saphnelo): ti, ab, kw; #6 #4 AND #5, limited to trials.

Scopus: TITLE-ABS-KEY(("systemic lupus erythematosus" OR "lupus nephritis" OR SLE) AND (belimumab OR benlysta OR anifrolumab OR saphnelo) AND (random* OR placebo OR "double-blind"))

ClinicalTrials.gov: Condition: Lupus Erythematosus, Systemic OR Lupus Nephritis; Intervention: belimumab OR anifrolumab; Study type: Interventional; Status: Completed; results posted or published.

Eligibility Criteria

Population: adults (18 years or older) with SLE, including biopsy-proven lupus nephritis. Intervention: intravenous belimumab or anifrolumab added to standard therapy. Comparator: placebo plus standard therapy. Outcomes: global responder endpoint at 52 weeks (SRI-4 [16], SRI-SLEDAI-2K-modified [16,24], or BICLA [25]); complete renal response (CRR) or partial renal response; severe flare; glucocorticoid reduction; adverse events. Study design: randomized, double-masked, phase 2 to 4 trials. Trials of subcutaneous formulations, open-label extensions, non-randomized studies, and pediatric-only trials were excluded to keep the exposure homogeneous. No upper age limit or sex restriction was applied at the eligibility stage, because the aim was to identify all eligible trials and then appraise their age and sex composition.

Selection, Extraction, and Risk of Bias

The author screened titles, abstracts, and full texts and extracted trial characteristics, event counts, and denominators for the intention-to-treat population into a pre-specified form and then rechecked every extracted value against the primary publication and registry record in a second pass. Selection followed four sequential steps: (1) records from all sources were combined and duplicates removed; (2) titles and abstracts were screened against the eligibility criteria; (3) full texts of potentially eligible reports were assessed, and reasons for exclusion were recorded; and (4) for each included trial, the primary publication was designated the source report and cross-checked against its ClinicalTrials.gov record and, where needed, secondary publications for baseline data. Screening, selection, extraction, and risk-of-bias assessment were performed by a single reviewer (the author) with a second-pass verification; independent duplicate screening was not undertaken, and this is addressed as a limitation. Where only percentages were published, counts were derived from the percentage and the randomized denominator. Risk of bias was assessed using the Cochrane RoB 2 tool across its five domains [26] for the outcome used in the meta-analysis. The domains were (D1) the randomization process, (D2) deviations from intended interventions, (D3) missing outcome data, (D4) measurement of the outcome, and (D5) selection of the reported result. Each domain was judged as low risk, some concerns, or high risk using the RoB 2 signaling-question algorithms, with the effect of assignment to the intervention (intention-to-treat) as the effect of interest. Overall risk was judged as low when all domains were low; some concerns when at least one domain raised some concerns; and none was high; and high risk when any domain was high or several domains raised some concerns. Because the outcome entering the meta-analysis was not always the trial's prespecified primary endpoint, D5 was judged against the registered primary endpoint and analysis plan. Risk of bias was assessed at the trial level and was not used to exclude trials.

Applicability to Women of Reproductive Age

Reproductive age was defined as 18 to 49 years. Because the primary reports do not tabulate efficacy by age band, the analysis cannot be restricted to this group. To judge how far the pooled estimates describe it, three steps were taken. First, the mean (or median) age and the percentage of women in each trial were extracted from the primary articles, registry results, or pooled baseline tables. Second, the share of participants younger than 50 years was estimated for trials reporting a standard deviation, using a normal approximation; this is an approximation, not an observed proportion. Third, an exploratory random-effects meta-regression relating the log odds ratio (OR) to mean trial age across the efficacy trials was ran, with the Hartung-Knapp variance correction and a t distribution with k minus 2 degrees of freedom. The percentage of women was too narrow in range among efficacy trials to support a meta-regression. These steps describe the composition of the study populations; they do not convert pooled whole-trial estimates into age-specific effects.

Statistical Analysis

Dichotomous outcomes were expressed as ORs, with risk ratios (RRs) and risk differences (RDs) as secondary measures. Effects were pooled using the DerSimonian-Laird random-effects model [27], chosen because the trials differed in population, dose, endpoint, and follow-up, so a single common effect could not be

assumed, and because the method is widely used and transparent, allowing comparison with earlier syntheses. It estimates the between-trial variance (τ^2) by the method of moments and weights each trial by the inverse of its within-trial variance plus τ^2; the Mantel-Haenszel fixed-effect estimate [28] was computed as a sensitivity analysis. Because the number of trials was small, the Hartung-Knapp adjustment with the ad hoc variance correction was applied to the confidence intervals [29]. The DerSimonian-Laird method treats the estimated τ^2 as known and relies on a normal approximation, which tends to produce intervals that are too narrow when few trials are pooled. The Hartung-Knapp adjustment instead estimates the variance of the pooled effect from the observed spread of the trial estimates and uses a t distribution with $k - 1$ degrees of freedom (k = number of trials); the ad hoc correction prevents the Hartung-Knapp interval from being narrower than the DerSimonian-Laird interval. With very few trials (for example, $k = 2$, giving 1 degree of freedom), Hartung-Knapp intervals can be extremely wide. Both intervals were reported and, where they disagree, greater weight to the Hartung-Knapp interval was given when interpreting statistical significance. Heterogeneity was quantified with Cochran Q, I^2, and τ^2 [30]. Cochran's Q tests whether between-trial variation exceeds what chance would expect but has low power when few trials are available. I^2 is the percentage of the total variability in effect estimates that is attributable to between-trial heterogeneity rather than sampling error; values of about 25%, 50%, and 75% are conventionally described as low, moderate, and high. τ^2 is the between-trial variance on the log OR scale. Because I^2 is itself imprecise when k is small and does not capture clinical heterogeneity arising from different endpoint definitions, it was interpreted together with τ^2, the prediction interval, and clinical judgment. Prespecified subgroups of belimumab versus anifrolumab using a Q-between test were compared. Because trials were allocated to subgroups by agent, this comparison is indirect (between trials rather than within trials) and is not a substitute for a head-to-head comparison. Leave-one-out analyses assessed influence, and a prediction interval was computed for the overall estimate. Small-study effects were examined with a funnel plot and the Egger regression test [31], recognizing low power with fewer than ten trials. The responder endpoint was SRI-4 where available (BLISS-52, BLISS-76, and TULIP-1), SRI-SLEDAI-2K in EMBRACE (its primary endpoint), and BICLA in TULIP-2 (its primary endpoint); MUSE contributed its week-52 SRI-4. These composite endpoints share a SLEDAI-based core but differ in how they score partial improvement and organ-specific worsening. Pooling them therefore assumes only that each captures a clinically meaningful global response in the same direction; the pooled OR is an average effect on global responder status, not an effect on any single instrument. This limitation is considered in the discussion. Python (NumPy, SciPy) was used for analyses.

Results

Study Selection

The database searches identified 4,222 records. After removing 1,680 duplicates, 2,542 records were screened, 32 full-text reports were assessed, and eight randomized trials were included, each represented by its primary publication (Figure *1*). Six trials contributed to the global efficacy analysis (BLISS-52, BLISS-76, EMBRACE, MUSE, TULIP-1 and TULIP-2) and two to the renal analysis (BLISS-LN and TULIP-LN).

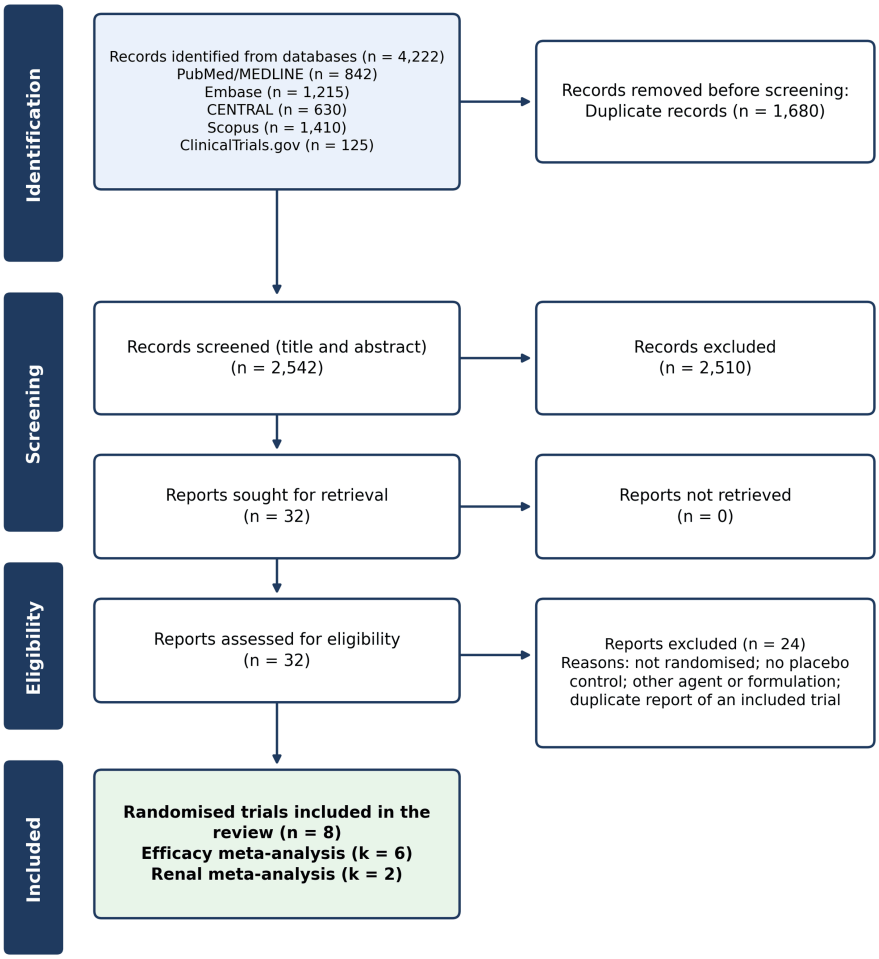

FIGURE 1: PRISMA 2020 flow diagram of study identification, screening, eligibility and inclusion

PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analyses

Characteristics of the Included Trials

The eight trials randomized 3,854 participants (Table *1*) [13-15,17-21]. Five were phase 3, one was phase 3/4 (EMBRACE), and two were phase 2 (MUSE and TULIP-LN). Active-treatment follow-up was 52 weeks in six trials, 76 weeks in BLISS-76, and 104 weeks in BLISS-LN. Belimumab was given at 10 mg/kg intravenously in all analysed arms. Anifrolumab was given at 300 mg every four weeks in MUSE, TULIP-1, and TULIP-2; in TULIP-LN, the intensified regimen (900 mg for three doses, then 300 mg) was compared with placebo for the primary renal analysis. All trials were placebo-controlled and added the study drug to standard therapy. The trial populations were predominantly female (range 82.8% to 96.9%), and the mean or median age ranged from 33.4 to 42.1 years (Table *1*). None reported efficacy by age band or sex, and no trial reported the share of participants aged 18 to 49 years.

Author and year	Trial	Design	Randomized population (N)	Demographics	Biologic agent and regimen	Standard-of-care baseline	Primary outcome measure	Follow-up
Navarra 2011 [14]	BLISS-52	Phase 3, double-blind RCT; Asia, South America, Eastern Europe	N = 865	Mean age 35.5 y (SD 11.1); female 94.9% (821/865); share aged 18-49 y not reported	Belimumab 1 or 10 mg/kg IV (loading doses, then every 28 days) vs placebo	Standard SLE therapy (glucocorticoids, antimalarials, immunosuppressants)	SRI-4 at week 52	52 weeks
Furie 2011 [15]	BLISS-76	Phase 3, double-blind RCT; North America, Europe	N = 819	Mean age 40.2 y (SD 11.5); female 93.3% (764/819); share aged 18-49 y not reported	Belimumab 1 or 10 mg/kg IV (loading doses, then every 28 days) vs placebo	Standard SLE therapy	SRI-4 at week 52 (week 76 secondary)	76 weeks
Furie 2020 [18]	BLISS-LN	Phase 3, double-blind RCT; 107 sites, 21 countries	N = 448 with biopsy-proven lupus nephritis	Mean age 33.4 y (33.1 placebo, 33.7 belimumab); female 88% (393/446); share aged 18-49 y not reported	Belimumab 10 mg/kg IV on days 1, 15, 29, then every 28 days vs placebo	Mycophenolate or cyclophosphamide-azathioprine induction plus glucocorticoids	Primary efficacy renal response at week 104 (CRR secondary)	104 weeks
Ginzler 2022 [17]	EMBRACE	Phase 3/4, double-blind RCT	N = 448	Self-identified Black race; mean age 38.8 y (SD 11.4); female 96.9%; share aged 18-49 y not reported	Belimumab 10 mg/kg IV monthly vs placebo	Standard SLE therapy	SRI-SLEDAI-2K response at week 52	52 weeks
Furie 2017 [13]	MUSE	Phase 2b, double-masked RCT	N = 305 (placebo 102; 300 mg 99; 1000 mg 104)	Mean age 39.8 y (SD 12.2); female 93.5% (reported for the trial population of 307); share aged 18-49 y not reported	Anifrolumab 300 or 1000 mg IV every four weeks for 48 weeks vs placebo	Oral glucocorticoids, azathioprine, antimalarial, mycophenolate or methotrexate	SRI-4 at day 169 with sustained glucocorticoid reduction (week-52 SRI-4 used here)	52 weeks
Furie 2019 [19]	TULIP-1	Phase 3, double-masked RCT	N = 457 (300 mg 180; 150 mg 93; placebo 184)	Mean age 41.3 y (SD 12.1); female 92.3% (422/457); share aged 18-49 y not reported	Anifrolumab 150 or 300 mg IV every four weeks for 48 weeks vs placebo (300 mg analyzed)	Standard SLE therapy	SRI-4 at week 52	52 weeks
Morand 2020 [20]	TULIP-2	Phase 3, double-masked RCT	N = 365 randomized, 362 treated	Mean age 42.1 y (SD 11.7); female 93.4% (338/362); aged 65 y or older 1.7% (6/362); share aged 18-49 y not reported	Anifrolumab 300 mg IV every four weeks for 48 weeks vs placebo	Standard SLE therapy	BICLA response at week 52	52 weeks
Jayne 2022 [21]	TULIP-LN	Phase 2, double-masked RCT	N = 147 with biopsy-proven class III/IV lupus nephritis (basic 45; intensified 51; placebo 49)	Median age 34.5 y anifrolumab, 32.0 y placebo; female 82.8% (120/145 treated); share aged 18-49 y not reported	Anifrolumab basic (300 mg) or intensified (900 mg x3, then 300 mg) IV monthly vs placebo	Oral glucocorticoids plus mycophenolate mofetil	Change in 24-hour urinary protein-creatinine ratio at week 52 (CRR secondary)	52 weeks (two-year trial)

TABLE 1: Characteristics of the included randomized controlled trials

RCT, randomized controlled trial; IV, intravenous; SRI-4, Systemic Lupus Erythematosus Responder Index; SLEDAI-2K, SLE Disease Activity Index 2000; BICLA, BILAG-based Composite Lupus Assessment; CRR, complete renal response. Mean (SD) age and female percentage are for the whole randomized or treated population as reported in the primary article, ClinicalTrials.gov results record or pooled baseline tables (BLISS-52, BLISS-76 and BLISS-LN from pooled or secondary baseline tables; TULIP-1 and TULIP-2 from registry results; TULIP-LN reports median age). No trial reported the share of participants aged 18-49 years.

Risk of Bias

Randomization, deviation from intended interventions, missing outcome data, and outcome measurement were rated at low risk for all trials (Table *2*). Selection of the reported result raised some concerns in three trials, because the outcome used here was not the prespecified primary endpoint: week-52 SRI-4 in MUSE (the primary endpoint was a day-169 composite with glucocorticoid tapering), CRR in BLISS-LN (primary efficacy renal response was the primary endpoint), and CRR in TULIP-LN (change in the urinary protein-creatinine ratio was primary). The overall judgement was low risk for five trials and some concerns for three.

Trial	D1 Randomization	D2 Deviations from intended interventions	D3 Missing outcome data	D4 Outcome measurement	D5 Selection of reported result	Overall
BLISS-52 [14]	Low	Low	Low	Low	Low	Low
BLISS-76 [15]	Low	Low	Low	Low	Low	Low
BLISS-LN [18]	Low	Low	Low	Low	Some concerns	Some concerns
EMBRACE [17]	Low	Low	Low	Low	Low	Low
MUSE [13]	Low	Low	Low	Low	Some concerns	Some concerns
TULIP-1 [19]	Low	Low	Low	Low	Low	Low
TULIP-2 [20]	Low	Low	Low	Low	Low	Low
TULIP-LN [21]	Low	Low	Low	Low	Some concerns	Some concerns

TABLE 2: Risk of bias assessment using the Cochrane RoB 2 tool

D1-D5, RoB 2 domains assessed for the outcome used in the meta-analysis.

Applicability to Women of Reproductive Age

Using the normal approximation, an estimated 75% to 94% of participants in the seven trials reporting a standard deviation were younger than 50 years (BLISS-LN, the youngest, 94%; TULIP-2, the oldest, 75%), and TULIP-LN had a median age of 34.5 years in the anifrolumab arms. With 83% to 97% female enrolment, most pooled participants were women within or near the reproductive age range, so the pooled estimates largely reflect this group but are not estimates for it. In the exploratory meta-regression across the six efficacy trials, the OR changed by a factor of 0.76 (95% CI 0.10 to 5.63; P = 0.72) per 10 years of mean trial age. With six trials and a narrow age range, this analysis cannot exclude age-related differences in effect.

Global Disease Activity

Six trials (n = 2,500) contributed to the global efficacy analysis (Figure *2*). The pooled random-effects OR for a responder endpoint at week 52 was 1.54 (95% CI 1.18 to 2.02), with moderate heterogeneity (I^2 = 62%, τ^2 = 0.07, Q P = 0.02). The pooled RR was 1.27 (95% CI 1.10 to 1.47) and the RD was 10.6 percentage points (95% CI 4.0 to 17.1). The Hartung-Knapp interval was 1.06 to 2.25, and the 95% prediction interval was 0.68 to 3.51. The Mantel-Haenszel fixed-effect OR was 1.53 (95% CI 1.30 to 1.80). Leave-one-out estimates ranged from 1.44 to 1.70. These estimates pool different composite endpoints (SRI-4, SRI-SLEDAI-2K and BICLA) and are best read as an average effect on global responder status, not as an effect on a single validated instrument.

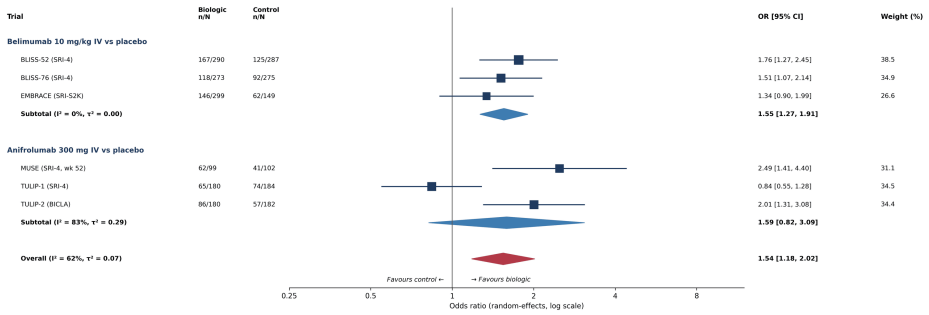

FIGURE 2: Forest plot of the primary efficacy endpoint

Included trials: BLISS-52 [14], BLISS-76 [15] and EMBRACE [17] (belimumab); MUSE [13], TULIP-1 [19] and TULIP-2 [20] (anifrolumab). Odds ratios for a week-52 responder endpoint (SRI-4, SRI-SLEDAI-2K or BICLA) with intravenous belimumab or anifrolumab versus placebo; DerSimonian-Laird random-effects model. Weights are within subgroup. Black squares represent the odds ratio of each trial (point estimate), and the horizontal line through each square is its 95% confidence interval; square size is proportional to the trial weight. Blue diamonds represent the pooled random-effects odds ratio for each subgroup (belimumab and anifrolumab). The red diamond represents the overall pooled random-effects odds ratio across all six trials. For each diamond, the center marks the pooled odds ratio and the left and right tips mark the limits of its 95% confidence interval.

Among the three belimumab trials, the pooled OR was 1.55 (95% CI 1.27 to 1.91) with no heterogeneity (I^2 = 0%); the Hartung-Knapp interval was 0.99 to 2.43 because only three trials contributed. EMBRACE, restricted to Black patients, gave an estimate in the same direction (OR 1.34) with an interval that included 1. Among the three anifrolumab trials, the pooled OR was 1.59 (95% CI 0.82 to 3.09) with substantial heterogeneity (I^2 = 83%, τ^2 = 0.29). TULIP-1, whose primary SRI-4 endpoint was not met (36.2% versus 40.4%), was the outlier; without it, the pooled anifrolumab OR for MUSE and TULIP-2 was 2.17 (95% CI 1.54 to 3.06). This leave-one-out result is post hoc and does not justify discarding TULIP-1, a well-conducted trial rated as low risk of bias. The Hartung-Knapp interval for the three-trial anifrolumab estimate was 0.37 to 6.83, and I^2 is imprecise with only three trials. Heterogeneity may reflect differences in endpoints (SRI-4 in MUSE and TULIP-1; BICLA in TULIP-2) as well as chance. The test for subgroup differences did not suggest that the agents differed (Q = 0.00, df = 1, P = 0.95), although the test has little power with three trials per subgroup. A non-significant test for subgroup differences does not demonstrate equivalence: the intervals for both subgroups (belimumab 1.27 to 1.91; anifrolumab 0.82 to 3.09) are compatible with a wide range of relative effects, and the comparison is indirect. Responder definitions differ between SRI-4, SRI-SLEDAI-2K, and BICLA, adding clinical heterogeneity that I^2 does not capture.

Renal Response and Glucocorticoid Sparing

Two trials reported CRR (Figure 3, Table 3). In BLISS-LN, 30% of patients on belimumab and 20% on placebo achieved CRR at week 104 (OR 1.7, 95% CI 1.1 to 2.7) [18]; in TULIP-LN, 45.5% of patients on the intensified anifrolumab regimen and 31.1% on placebo did so at week 52 [21]. The pooled OR was 1.73 (95% CI 1.17 to 2.55), with no heterogeneity (I^2 = 0%). With only two trials, the Hartung-Knapp interval is very wide (0.14 to 21.44), and the pooled estimate should be read descriptively. The DerSimonian-Laird interval (1.17 to 2.55) is narrow partly because it ignores uncertainty in the between-trial variance estimate; with k = 2, the Hartung-Knapp t distribution has only one degree of freedom, and an I^2 of 0% carries little information for two trials. The trials also differ in phase (3 versus 2), time point (week 104 versus week 52), regimen (belimumab 10 mg/kg versus intensified anifrolumab), and background therapy, and TULIP-LN did not meet its primary endpoint. The pooled estimate is therefore hypothesis-generating and does not support conclusions about the relative efficacy of the two agents in lupus nephritis. In BLISS-LN, the primary efficacy renal response at week 104 was reached by 43% versus 32% [18].

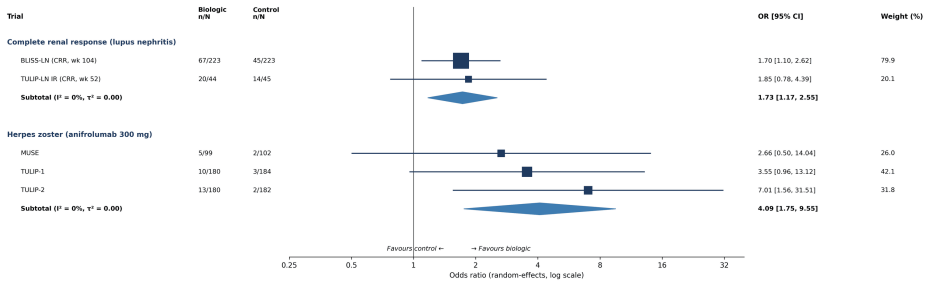

FIGURE 3: Forest plot of renal and safety outcomes

Included trials: BLISS-LN [18] and TULIP-LN [21] (complete renal response); MUSE [13], TULIP-1 [19], and TULIP-2 [20] (herpes zoster). Odds ratios for complete renal response in lupus nephritis (BLISS-LN; TULIP-LN intensified regimen) and for herpes zoster with anifrolumab 300 mg versus placebo: DerSimonian-Laird random-effects model. Black squares represent the odds ratio of each trial (point estimate), and the horizontal line through each square is its 95% confidence interval; square size is proportional to the trial weight. Blue diamonds represent the pooled random-effects odds ratio for each outcome (complete renal response and herpes zoster); the center of each diamond marks the pooled odds ratio, and its left and right tips mark the limits of the 95% confidence interval. No overall pooled estimate (red diamond) is shown because the two outcomes are not combined.

Outcome and comparison	Reference	Trials (participants)	OR (95% CI), DL	OR, Hartung-Knapp CI	RR (95% CI)	RD, percentage points (95% CI)	I^2/τ^2
Responder endpoint (SRI-4/BICLA), all biologics	[13-15,17,19,20]	6 (2,500)	1.54 (1.18-2.02)	1.06-2.25	1.27 (1.10-1.47)	10.6 (4.0 to 17.1)	62%/0.07
Belimumab 10 mg/kg	[14,15,17]	3 (1,573)	1.55 (1.27-1.91)	0.99-2.43	1.27 (1.14-1.43)	10.7 (5.8 to 15.6)	0%/0.00
Anifrolumab 300 mg	[13,19,20]	3 (927)	1.59 (0.82-3.09)	0.37-6.83	1.29 (0.90-1.84)	11.2 (-4.7 to 27.2)	83%/0.29
Complete renal response (BLISS-LN, TULIP-LN)	[18,21]	2 (535)	1.73 (1.17-2.55)	0.14-21.44	1.48 (1.12-1.96)	10.5 (3.1 to 17.9)	0%/0.00
Glucocorticoid reduction to ≤7.5 mg/day, anifrolumab (MUSE, TULIP-1; exploratory)	[13,19]	2 (324)	2.18 (0.90-5.30)	0.01-686.80	1.60 (0.96-2.66)	18.4 (-2.5 to 39.3)	71%/0.29
Herpes zoster, anifrolumab 300 mg	[13,19,20]	3 (927)	4.09 (1.75-9.55)	0.63-26.34	3.89 (1.70-8.91)	4.5 (2.1 to 7.0)	0%/0.00
Severe flare reduction	-	Not pooled	-	-	-	-	-

TABLE 3: Summary of pooled findings and subgroup analyses

OR, odds ratio; RR, risk ratio; RD, risk difference; CI, confidence interval; DL, DerSimonian-Laird random effects; Hartung-Knapp CI with ad hoc variance correction; I^2, heterogeneity statistic; τ^2, between-trial variance. Counts for some trials were derived from published percentages and denominators. Severe flare, infusion, and hypersensitivity reactions, and non-zoster infections were not pooled. Test for subgroup difference (belimumab vs. anifrolumab): Q = 0.00, df = 1, P = 0.95. Uniform event counts could not be extracted for severe flare. Reference column: numbers in square brackets refer to the reference list. -, not applicable, or not pooled.

Glucocorticoid reduction was defined differently across trials. In BLISS-52 and BLISS-76, a prednisone reduction of at least 25% to 7.5 mg/day or less during weeks 40 to 52 occurred in 19% versus 12% and 17.5% versus 12.7% of belimumab 10 mg/kg and placebo patients, respectively [14,15]. In TULIP-2, oral glucocorticoid reduction was achieved by 51.5% versus 30.2% [20]. In MUSE and TULIP-1, sustained reduction to 7.5 mg/day or less among patients taking at least 10 mg/day at baseline favored anifrolumab 300

mg (pooled OR 2.18 (95% CI 0.90 to 5.30); two trials, exploratory) [13,19].

Safety

Herpes zoster was more frequent with anifrolumab 300 mg than with placebo in all three trials that reported it (Figure 3); the pooled OR was 4.09 (95% CI 1.75 to 9.55) (I^2 = 0%), and the Hartung-Knapp interval was 0.63 to 26.34 [13,19,20]. Serious adverse events were reported in 14% of anifrolumab 300 mg and 16% of placebo patients in TULIP-1 [19]. In a post hoc mycophenolate subgroup of BLISS-LN, serious adverse events occurred in 24.4% on belimumab and 32.1% on placebo, an analysis that was not randomized for that stratum [18]. Uniform counts for other infections, infusion and hypersensitivity reactions and severe flares could not be extracted from all trials, and these outcomes were not pooled.

Publication Bias

The funnel plot of the six efficacy trials (Figure 4) showed no marked asymmetry, and the Egger test did not indicate small-study effects (intercept 1.63, P = 0.73). With six trials, this test has low power, and the absence of a signal does not exclude selective publication.

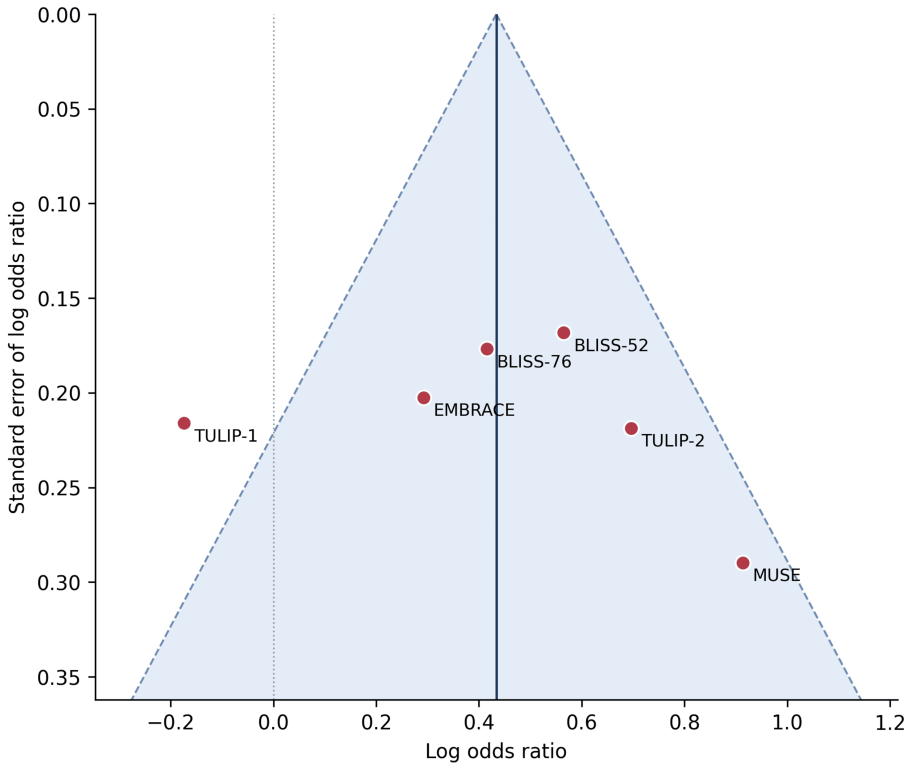

FIGURE 4: Funnel plot for assessment of publication bias

Included trials: BLISS-52 [14], BLISS-76 [15], EMBRACE [17], MUSE [13], TULIP-1 [19], and TULIP-2 [20]. Log odds ratio against its standard error for the six trials in the efficacy analysis; the vertical line marks the pooled random-effects estimate, and the dashed lines the pseudo 95% confidence limits.

Discussion

This meta-analysis of eight randomized placebo-controlled trials (six contributing to the pooled responder analysis and two to the renal analysis) shows that adding belimumab or anifrolumab to standard therapy increases the odds of a week-52 responder endpoint by about half (pooled OR 1.54; the endpoint pools non-identical composite definitions, as discussed below), equivalent to roughly 11 additional responders per 100 treated patients. The estimate was stable on leave-one-out analysis, and the Hartung-Knapp interval (1.06 to 2.25) still excluded 1. The 95% prediction interval (0.68 to 3.51) includes 1, so the average effect is reasonably secure, while the effect in a new trial or a new patient group is less certain.

Belimumab produced consistent effects across BLISS-52, BLISS-76 and EMBRACE, with no heterogeneity. The effect in EMBRACE was smaller, and its interval included 1, matching the trial's conclusion that the primary endpoint was not met while the direction favored treatment [17]. The anifrolumab trials were less consistent. TULIP-1 and TULIP-2 had identical regimens and similar populations but different primary

endpoints, and the SRI-4 in TULIP-1 and BICLA in TULIP-2 are not interchangeable: SRI-4 requires complete resolution of a manifestation before an item score changes, whereas BICLA credits partial improvement across organs. The anifrolumab estimate therefore depends on which endpoint is chosen, and the apparent heterogeneity (I^2 = 83%) is partly a property of the measurement. Differences in population, trial era, background therapy, and chance may also contribute, and I^2 is imprecise with only three trials. Excluding TULIP-1 raises the pooled anifrolumab OR to 2.17 (95% CI 1.54 to 3.06). Still, this post hoc exclusion would be selective: TULIP-1 was methodologically sound, and its neutral result is part of the evidence. Overall, anifrolumab appears to improve global response on average, but the effect is inconsistent across trials. The subgroup comparison found no difference between agents (P = 0.95), but this is an indirect comparison across trials with different endpoints, eras, and populations, with no head-to-head trial; it cannot be used to rank the agents. The similar point estimates (1.55 and 1.59) should not be read as equivalent efficacy: the absence of a statistically detectable difference is not evidence of equivalence, and the anifrolumab interval is wide enough to include no effect.

Another consideration is pooling composite responder endpoints. SRI-4, SRI-SLEDAI-2K and BICLA all aim to capture global improvement in disease activity without worsening, but they weight organ systems differently, treat partial improvement differently and rely on different instruments (SLEDAI-based scores versus a BILAG-based assessment). Pooling them yields a meaningful summary only if each is read as a marker of clinically relevant global response in the same direction; it does not yield an effect on a single validated scale, and the pooled OR cannot be translated into change on any one instrument. Endpoint-specific properties may influence the OR, and clinical heterogeneity of this kind is not captured by I^2 or τ^2. A sensitivity analysis restricted to a single endpoint was not feasible without discarding trials: only four trials (BLISS-52, BLISS-76, MUSE and TULIP-1) contributed SRI-4 as the analyzed endpoint, and SRI-4 data for EMBRACE and TULIP-2 were not extracted. Harmonized endpoints in future individual participant data analyses would address this limitation.

In lupus nephritis, belimumab improved primary efficacy renal response and CRR over two years in BLISS-LN [18], and this evidence led to regulatory approval of belimumab for lupus nephritis in December 2020. Anifrolumab did not meet the primary urinary protein-creatinine endpoint in TULIP-LN, although the intensified regimen showed numerically higher CRR [21]. A phase 3 trial of anifrolumab in proliferative lupus nephritis (IRIS, NCT05138133) is underway. The pooled CRR OR of 1.73 combines a phase 3 and a phase 2 trial with different time points, doses, and background regimens, and I regard it as a summary of two consistent trials, not evidence of a class effect. Interpretation should be cautious: with only two trials, the Hartung-Knapp interval (0.14 to 21.44) is too wide to exclude either no effect or a very large effect, the narrower DerSimonian-Laird interval (1.17 to 2.55) understates uncertainty, an I^2 of 0% carries little information, and TULIP-LN did not meet its primary endpoint. The estimate is hypothesis-generating and provides no basis for comparing the two agents in lupus nephritis. Renal preservation, the outcome that matters most for a young woman with decades of life ahead, has been assessed only through response measures; these trials have not yet reported hard renal outcomes over longer follow-up.

The applicability to the population named in the title deserves an explicit statement. Women made up 83% to 97% of every included trial, and the youngest trials were the two lupus nephritis studies (mean or median age 33 to 35 years), so the evidence base is largely a young female one, and an estimated 75% to 94% of participants were younger than 50 years. That is indirect support, not subgroup evidence: no trial published age-stratified or pregnancy-specific efficacy, the exploratory age meta-regression is underpowered, and women older than 49 years were included. Read the pooled estimates as applying to a predominantly reproductive-age population, with the caveat that effects specific to women aged 18 to 49 years remain unmeasured. Accordingly, the title and conclusions refer to predominantly young female trial populations rather than to a strictly defined 18-to-49-year subgroup. For women of reproductive age, three practical points follow. First, both agents reduce glucocorticoids in trials, and lower cumulative exposure matters for bone health, weight, glucose metabolism, and organ-damage risk [14,15,20]. Second, the disease-control benefit appears over 52 weeks, which fits the time frame of pre-conception planning, since disease quiescence before conception is associated with better pregnancy outcomes [4,5]. Third, the evidence on pregnancy and lactation is thin. These trials excluded pregnancy and required contraception, so the pregnancy-specific safety of both agents rests on registries and case series that this review did not assess. Decisions should follow current reproductive-health guidance [5,6]. The higher frequency of herpes zoster with anifrolumab is relevant to counseling, and assessing varicella and zoster immunity before treatment is reasonable.

Several limitations apply. The principal one is that no trial reported outcomes for women aged 18 to 49 years, so the target population is approximated by predominantly female trial populations with a young mean age, and the estimates may not transfer if age or sex modifies the response; the age meta-regression cannot settle this. Second, only eight trials were available, so Hartung-Knapp intervals are wide, subgroup tests are underpowered, and the funnel plot and Egger test are uninformative about small-study effects. The renal analysis rests on two trials and the anifrolumab subgroup on three, so their estimates and I^2 values are unstable. Third, outcome definitions differed (SRI-4, SRI-SLEDAI-2K, BICLA, renal response at different time points), and MUSE and TULIP-LN were phase 2 studies. Pooling these composite endpoints limits the interpretability of the pooled OR, as discussed above. Fourth, some event counts were derived from published percentages and denominators and should be checked against the source tables. Fifth, severe

flare, infusion and hypersensitivity reactions and non-zoster infections were not pooled because uniform counts were not extracted. Sixth, the pharmaceutical manufacturers sponsored the trials. Finally, a single reviewer screened and extracted data, with a second-pass check but no independent duplicate extraction, which increases the risk of selection or extraction error. Risk of bias was also judged by one reviewer, and the protocol was not registered prospectively, which limits protection against undeclared changes to outcomes or methods. Trials of subcutaneous formulations were excluded by design.

Future work should prioritize individual participant data analyses of the BLISS and TULIP programs that report responses by age band and sex, prospective pregnancy registries with comparator groups, head-to-head or network comparisons using a harmonized endpoint such as BICLA or the Lupus Low Disease Activity State [32], and long-term renal outcomes, including sustained response, flare, kidney function and kidney failure. The results of the anifrolumab lupus nephritis phase 3 program will determine whether the numerical signal seen in TULIP-LN is confirmed.

Conclusions

Across eight randomized trials enrolling predominantly young women, belimumab and anifrolumab each added measurable benefit to standard therapy in active SLE, and no head-to-head trial exists. In lupus nephritis, both trials favored CRR, with the strongest evidence for belimumab, whose phase 3 trial ran for two years. The renal meta-analysis rested on two trials with very wide Hartung-Knapp intervals and should be interpreted cautiously. Anifrolumab was associated with more herpes zoster.

For rheumatologists treating young women, the choice between agents should rest on the dominant organ manifestation, prior treatment, interferon-driven mucocutaneous disease, immunization status, and reproductive plans, not on the pooled estimates, because the trials were not designed to rank the agents. Counseling should state that no trial reported results for women of reproductive age as a group and that pregnancy was excluded. Pre-conception planning, contraception while on treatment, and review of glucocorticoid dose should remain part of every consultation. Trial analyses by age and sex, together with pregnancy registries, are needed before the evidence can be considered complete for this population.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Priyanka Hayre

Acquisition, analysis, or interpretation of data: Priyanka Hayre

Drafting of the manuscript: Priyanka Hayre

Critical review of the manuscript for important intellectual content: Priyanka Hayre

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Tsokos GC: Systemic lupus erythematosus. N Engl J Med. 2011, 365:2110-21. 10.1056/NEJMra1100359
2. Kaul A, Gordon C, Crow MK, et al.: Systemic lupus erythematosus. Nat Rev Dis Primers. 2016, 2:16039. 10.1038/nrdp.2016.39
3. Rees F, Doherty M, Grainge MJ, Lanyon P, Zhang W: The worldwide incidence and prevalence of systemic lupus erythematosus: a systematic review of epidemiological studies. Rheumatology (Oxford). 2017, 56:1945-61. 10.1093/rheumatology/kex260
4. Smyth A, Oliveira GH, Lahr BD, Bailey KR, Norby SM, Garovic VD: A systematic review and meta-analysis of pregnancy outcomes in patients with systemic lupus erythematosus and lupus nephritis. Clin J Am Soc Nephrol. 2010, 5:2060-8. 10.2215/CJN.00240110
5. Andreoli L, Bertsias GK, Agmon-Levin N, et al.: EULAR recommendations for women's health and the management of family planning, assisted reproduction, pregnancy and menopause in patients with systemic lupus erythematosus and/or antiphospholipid syndrome. Ann Rheum Dis. 2017, 76:476-85. 10.1136/annrheumdis-2016-209770
6. Sammaritano LR, Bermas BL, Chakravarty EE, et al.: 2020 American College of Rheumatology guideline for

the management of reproductive health in rheumatic and musculoskeletal diseases. Arthritis Rheumatol. 2020, 72:529-56. 10.1002/art.41191
7. Tektonidou MG, Dasgupta A, Ward MM: Risk of end-stage renal disease in patients with lupus nephritis, 1971-2015: a systematic review and Bayesian meta-analysis. Arthritis Rheumatol. 2016, 68:1432-41. 10.1002/art.39594
8. Hahn BH, McMahon MA, Wilkinson A, et al.: American College of Rheumatology guidelines for screening, treatment, and management of lupus nephritis. Arthritis Care Res (Hoboken). 2012, 64:797-808. 10.1002/acr.21664
9. KDIGO 2024 clinical practice guideline for the management of lupus nephritis . Kidney Int. 2024, 105:S1-S69. 10.1016/j.kint.2023.09.002
10. Fanouriakis A, Kostopoulou M, Cheema K, et al.: 2019 update of the Joint European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association (EULAR/ERA-EDTA) recommendations for the management of lupus nephritis. Ann Rheum Dis. 2020, 79:713-23. 10.1136/annrheumdis-2020-216924
11. Stohl W, Hilbert DM: The discovery and development of belimumab: the anti-BLyS-lupus connection . Nat Biotechnol. 2012, 30:69-77. 10.1038/nbt.2076
12. Banchereau J, Pascual V: Type I interferon in systemic lupus erythematosus and other autoimmune diseases. Immunity. 2006, 25:383-92. 10.1016/j.immuni.2006.08.010
13. 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. 10.1002/art.39962
14. Navarra SV, Guzman RM, Gallacher AE, et al.: Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomized, placebo-controlled, phase 3 trial. Lancet. 2011, 377:721-31. 10.1016/S0140-6736(10)61354-2
15. 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. 10.1002/art.30613
16. Furie RA, Petri MA, Wallace DJ, et al.: Novel evidence-based systemic lupus erythematosus responder index . Arthritis Rheum. 2009, 61:1143-51. 10.1002/art.24698
17. 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. 10.1002/art.41900
18. Furie R, Rovin BH, Houssiau F, et al.: Two-year, randomized, controlled trial of belimumab in lupus nephritis. N Engl J Med. 2020, 383:1117-28. 10.1056/NEJMoa2001180
19. Furie RA, Morand EF, Bruce IN, et al.: Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP- 1): a randomized, controlled, phase 3 trial. Lancet Rheumatol. 2019, 1:208-19. 10.1016/S2665-9913(19)30076-1
20. Morand EF, Furie R, Tanaka Y, et al.: Trial of anifrolumab in active systemic lupus erythematosus . N Engl J Med. 2020, 382:211-21. 10.1056/NEJMoa1912196
21. Jayne D, Rovin B, Mysler EF, et al.: Phase II randomised trial of type I interferon inhibitor anifrolumab in patients with active lupus nephritis. Ann Rheum Dis. 2022, 81:496-506. 10.1136/annrheumdis-2021-221478
22. Fanouriakis A, Kostopoulou M, Andersen J, et al.: EULAR recommendations for the management of systemic lupus erythematosus: 2023 update. Ann Rheum Dis. 2024, 83:15-29. 10.1136/ard-2023-224762
23. Page MJ, McKenzie JE, Bossuyt PM, et al.: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021, 372:n71. 10.1136/bmj.n71
24. Gladman DD, Ibañez D, Urowitz MB: Systemic lupus erythematosus disease activity index 2000 . J Rheumatol. 2002, 29:288-91.
25. Wallace DJ, Kalunian K, Petri MA, et al.: Efficacy and safety of epratuzumab in patients with moderate/severe active systemic lupus erythematosus: results from EMBLEM, a phase IIb, randomised, double-blind, placebo-controlled, multicentre study. Ann Rheum Dis. 2014, 73:183-90. 10.1136/annrheumdis-2012-202760
26. Sterne JA, Savović J, Page MJ, et al.: RoB 2: a revised tool for assessing risk of bias in randomised trials . BMJ. 2019, 366:l4898. 10.1136/bmj.l4898
27. DerSimonian R, Laird N: Meta-analysis in clinical trials. Controlled Clinical Trials. 1986, 7:177-88. 10.1016/0197-2456(86)90046-2
28. MA N, HA W: Statistical aspects of the analysis of data from retrospective studies of disease . J Natl Cancer Inst. 1959, 22:719-48.
29. IntHout J, Ioannidis JP, Borm GF: The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol. 2014, 14:25. 10.1186/1471-2288-14-25
30. Higgins JP, Thompson SG: Quantifying heterogeneity in a meta-analysis . Stat Med. 2002, 21:1539-58. 10.1002/sim.1186
31. Egger M, Davey Smith G, Schneider M, Minder C: Bias in meta-analysis detected by a simple, graphical test . BMJ. 1997, 315:629-34. 10.1136/bmj.315.7109.629
32. Franklyn K, Lau CS, Navarra SV, et al.: Definition and initial validation of a Lupus Low Disease Activity State (LLDAS). Ann Rheum Dis. 2016, 75:1615-21. 10.1136/annrheumdis-2015-207726