CADTH

CADTH COMMON DRUG REVIEW

Clinical Review Report

BELIMUMAB (BENLYSTA)

(GlaxoSmithKline Inc.)

Indication: Indicated in addition to standard therapy for reducing disease activity in adult patients with active, autoantibody-positive systemic lupus erythematosus.

Service Line:	CADTH Common Drug Review
Version:	Final
Publication Date:	June 2020
Report Length:	126 Pages

CADTH

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

CADTH

Table of Contents

Tables

Figures

Abbreviations

AE	adverse event
ACR	American College of Rheumatology
ANA	antinuclear antibody
AZA	azathioprine
anti-Sm	anti-Smith
BILAG	British Isles Lupus Assessment Group
CAPA	Canadian Arthritis Patient Alliance
CI	confidence interval
CNS	central nervous system
CrI	credible interval
CYC	cyclophosphamide
DB	double blind
dsDNA	double-stranded DNA
ESR	erythrocyte sedimentation rate
FACIT-Fatigue	Functional Assessment of Chronic Illness Therapy–Fatigue Scale
GC	glucocorticoid
Hb	hemoglobin
HDA	high disease activity
HR	hazard ratio
ICC	intraclass correlation coefficient
ITC	indirect treatment comparison
ITT	intention to treat
LOCF	last observation carried forward
MCID	minimal clinically important difference
MCS	mental component summary
MID	minimal important difference
MMF	mycophenolate mofetil
NMA	network meta-analysis
NSAID	nonsteroidal anti-inflammatory drug

OR	odds ratio
PCS	physical component summary
PGA	Physician Global Assessment
RCT	randomized controlled trial
RR	rate ratio
SAE	serious adverse event
SC	subcutaneous
SD	standard deviation
SDI	Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index
SELENA	Safety of Estrogen in Lupus Erythematosus National Assessment
SEM	standard error of the mean
SF-36	Short Form (36) Health Survey
SFI	Systemic Lupus Erythematosus Flare Index
SLE	systemic lupus erythematosus
SLEDAI	Systemic Lupus Erythematosus Disease Activity Index
SLICC	Systemic Lupus International Collaborating Clinics
SRI	Systemic Lupus Erythematosus Responder Index
TAC	tacrolimus
TAS	The Arthritis Society
TLC	Toronto Lupus Cohort
TNF	tumour necrosis factor
VAS	visual analogue scale
WDAE	withdrawal due to adverse event

Drug	belimumab (Benlysta)
Indication	Indicated in addition to standard therapy for reducing disease activity in adult patients with active, autoantibody-positive systemic lupus erythematosus (SLE).
Reimbursement request	For the treatment of patients with SLE who meet the following eligibility criteria: • adult patients age 18 years or older; and • patients with active, antibody-positive SLE; and • currently receiving standard therapy; and • has a disease activity SELENA-SLEDAI score ≥ 8. If no improvements are observed in a patient's SLE disease activity and/or symptoms after 6 months, use should be discontinued.
Dosage form(s) and route of administration) and strength(s)	Solution for subcutaneous injection, 200 mg/mL, in pre-filled syringe or autoinjector
NOC date	December 7, 2017
Sponsor	GlaxoSmithKline Inc.

Executive Summary

Introduction

Lupus is an autoimmune disease that affects approximately one in a thousand Canadians, and the most serious form of lupus is systemic lupus erythematosus (SLE). The precise etiology and pathophysiology are unknown; however, women are more commonly afflicted than men (women:men = 9:1). Women are particularly likely to experience onset of SLE between the ages of 15 and 45, although the disease can present at any age. The symptoms of lupus can vary greatly. Patients can experience fatigue and joint pain, which can be disabling, as well as neurological, renal, and cardiovascular sequelae, rash, and a variety of other symptoms. The disease has a variable course, and patients can cycle among a chronic state, flares (acute worsening of their condition), and remission.

SLE is treated with medications, both those taken acutely on an as-needed basis, as well as those taken chronically. Among the chronically administered agents, the first-line drugs are antimalarial drugs, which interfere with intracellular Toll-like receptor signalling. Given that SLE is an autoimmune disorder, immunosuppressants also play an important role, and there are a variety that are used (azathioprine, cyclophosphamide, methotrexate, mycophenolate, and cyclosporine). Immunosuppressants have multiple harms associated with them, including risk of serious infection and malignancy, and present significant tolerability issues for patients. Corticosteroids are well known for toxic effects such as osteoporosis, psychiatric issues, cataracts, diabetes, hypertension, weight gain, hirsutism, glaucoma, and many others, particularly when used chronically. Thus, although corticosteroids are relied upon for flares, chronic use is avoided as much as possible.

Belimumab inhibits the B-lymphocyte simulator, and thus inhibits B-cell functions, which are believed to play an important role in the pathophysiology of SLE.[1] The drug is administered by subcutaneous (SC) injection, 200 mg/mL once weekly and is indicated in addition to standard therapy to reduce disease activity in adult patients with active, autoantibody-positive SLE.

The objective of this report is to perform a systematic review of the beneficial and harmful effects of belimumab 200 mg/mL SC injection in addition to standard therapy for reducing disease activity in adult patients with active, autoantibody-positive SLE.

Stakeholder Engagement

Patient Input

Joint input was provided by the Canadian Arthritis Patient Alliance and The Arthritis Society and was collected via a survey distributed through email and social media. There were 14 respondents to the survey. Patients described being afflicted with a wide range of symptoms, including swelling, pain, rash, fatigue, and cognitive impairment, as well as more serious sequelae such as respiratory and cardiac disorders. The impact on health-related quality of life is significant, as the various symptoms, most notably fatigue and cognitive impairment, limit their ability to carry out daily activities and work productively. The key outcomes patients would like to see addressed by a new therapy are pain and fatigue, organ involvement, disease complications, and lupus nephritis. Patients would like to see enhanced mobility, productivity, and ability to work and carry out activities of daily living and social roles. Overall, it is clear that SLE significantly impairs health-related quality of life and function, as well as eliciting a number of serious sequelae.

Clinician Input

SLE is currently treated chronically with immune modulators such as high-dose corticosteroids, antimalarial drugs, azathioprine, methotrexate, mycophenolate mofetil, cyclophosphamide, and cyclosporine/tacrolimus. Unmet needs include therapeutic failure, (primary or secondary), recurrent flares that cause progressive organ damage, toxic effects of corticosteroids, and lack of safety in pregnancy in a disease that is relatively common in women of child-bearing age. The clinical experts noted that approximately 60% to 70% of patients do not have a good long-term response to therapy without the intermittent or continuous use of corticosteroids. The current place in therapy for belimumab is likely going to be in patients with milder disease (i.e., not first-line). However, this may change as more experience is gained with the drug. In patients with more severe disease, induction with an immunosuppressive drug could be followed by use of belimumab for maintenance therapy if remission is achieved. The clinical experts involved in this review did not identify any biomarker that could be used to select patients for treatment or to predict response to belimumab, although those most in need of belimumab would be 1) those using it as adjunctive therapy for milder forms of the disease to avoid toxic effects of intermittent or long-term corticosteroid use, 2) those using it to maintain a clinical response previously induced with more aggressive therapies that should not be continued long-term, and 3) those using it as a (potential) safe alternative for use in pregnancy. Clinically meaningful response would be signified by meaningful reduction in signs and symptoms of inflammation mediated by the SLE-aberrant immune response, and these include clinical and laboratory variables of organ dysfunction and laboratory biomarkers such autoantibodies. Treatment response should be assessed every six to 12 weeks for those with active disease, and the rapidity of response depends on the treatment (corticosteroids are most rapid) and organ system (renal and central nervous system [CNS] are slowest).

Clinical Evidence

Pivotal Studies and Protocol-Selected Studies

Description of Studies

One sponsor-funded multinational (177 centres in 30 countries, including Canada) double-blind (DB) randomized controlled trial (RCT) was included in this review. BLISS SC randomized 836 patients with active SLE (Safety of Estrogen in Lupus Erythematosus National Assessment–SLE Disease Activity Index [SELENA-SLEDAI] score of ≥ 8), 2:1, to either belimumab (N = 556) or placebo (N = 280) once weekly by SC injection, over a 52-week period. SELENA-SLEDAI is a measure of disease activity at time of visit or in the preceding 10 days. It consists of 24 weighted clinical and laboratory variables, with total possible score ranged from 0 to 105. The primary outcome was the percentage of patients with an SLE Responder Index (SRI) response at 52 weeks, a composite that was defined by 1) a reduction from baseline in the SELENA-SLEDAI score of four points or more, 2) no worsening (increase of < 0.30 points from baseline) in Physician Global Assessment (PGA), and 3) no new British Isles Lupus Assessment Group (BILAG) A organ domain score or two new BILAG B organ domain scores compared with baseline. The two secondary outcomes were time to first severe flare over 52 weeks and percentage of patients with an average prednisone reduction of 25% or more from baseline to less than 7.5 mg per day during weeks 40 through 52.

Efficacy Results

There were three patients in the belimumab group (0.5%) who died (of sepsis, urosepsis, and tuberculosis) and two patients in the placebo group (0.7%) (of thrombocytopenia and cardiac arrest) over the course of the 52-week DB treatment phase. Irreversible organ damage was assessed using the Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index (SDI). However, this was an exploratory outcome. There was no statistically significant difference between belimumab and placebo groups after 52 weeks with respect to change from baseline in the SDI. Remission was not specifically assessed in the included study.

More patients in the belimumab than the placebo group (61% versus 48%) achieved a response on the SRI (odds ratio [OR] = 1.68 [95% CI, 1.25 to 2.25], P = 0.0006), the primary outcome of this study. The clinical experts consulted by CADTH Common Drug Review (CDR) on this review suggested that this was a clinically relevant difference with an adjunctive therapy in a condition with limited therapeutic options that are tolerated by patients. The individual components of the SRI were also improved in a greater percentage of patients receiving belimumab than patients receiving placebo: four-point reduction in SELENA-SLEDAI score by week 52 (62% versus 49%, OR = 1.69 [95% CI, 1.26 to 2.27]); no worsening in PGA by week 52 (81% versus 73%, OR = 1.61 [95% CI, 1.15 to 2.27]), and no new 1A/2B BILAG domain scores by week 52 (81% versus 74%, OR = 1.46 [95% CI, 1.04 to 2.07]). Subgroup analyses were performed on the primary outcome of SRI response based on, for example, SELENA-SLEDAI score (≤ 9 or ≥ 10), baseline prednisone use (no use or use), or baseline medications (steroid, antimalarial drug, and immunosuppressant or other). However, no adjustments were made for multiple comparisons, thus limiting the conclusions that can be drawn from these data. In the subgroup of patients who were not on prednisone at baseline, there was no difference in SRI response between belimumab and placebo.

There were fewer patients treated with belimumab than with placebo (11% versus 18%) who had a severe flare during the 52-week study, and this difference was statistically significant (hazard ratio [HR] = 0.51 [95% CI, 0.35 to 0.74], P = 0.0004). This was a secondary outcome of BLISS SC. The clinical experts consulted by CDR on this review noted the importance of these severe flares to patients, and patients also alluded to the significant impact of these flares in their input to CDR. There was no statistically significant difference between belimumab (18% of patients) and placebo (12%) in the percentage of patients who could reduce their prednisone dosage by at least 25%, to 7.5 mg daily or less (OR = 1.65 [95% CI, 0.95 to 2.84], P = 0.0732), although a higher percentage reduction was observed in patients with belimumab. Because patients had to be on a dose of more than 7.5 mg of prednisone daily at baseline to be included in this analysis, this outcome included only about 60% of the intention-to-treat (ITT) population. The clinical experts consulted by CDR on this review noted the importance of reducing patients' reliance on corticosteroids, given the severe adverse effects associated with this class of drugs with heavy or prolonged use. Health-related quality of life was not studied, and symptoms such as fatigue were assessed only as exploratory outcomes, a significant limitation of this review, given the importance of these outcomes to patients.

Harms Results

Eleven percent of patients receiving belimumab and 16% of patients receiving placebo experienced a serious adverse event (SAE), and 81% of patients receiving belimumab and 84% of patients receiving placebo had an adverse event (AE). The only SAE that occurred in at least 1% of patients was thrombocytopenia, and all three patients (1.1%) experiencing this SAE were in the placebo group. Seven percent of patients receiving belimumab and 9% of patience receiving placebo discontinued the study drug due to an AE. Among notable harms, post-injection reactions occurred in 7% of patients receiving belimumab and 9% of patients receiving placebo; psychiatric AEs occurred in 6% of belimumab and 11% of patients receiving placebo; and infections of special interest occurred in 5% of belimumab and 8% of patients receiving placebo. Tuberculosis occurred in 0.4% of patients in belimumab and 0.7% of patients in placebo groups, and herpes zoster occurred in 3.1% of belimumab and 4.6% of placebo groups.

Table 1: Summary of Key Results From Pivotal and Protocol-Selected Studies

	BLISS SC	
	Belimumab N = 556	**Placebo N = 280**
Mortality		
Deaths by week 52, n (%)	3 (0.5)	2 (0.7)
Organ damage		
SLICC/ACR Damage Index change from baseline at week 52	N = 537	N = 268
Mean (SD) baseline	0.6 (1.00)	0.7 (1.19)
Mean (SD) change to week 52	0.0 (0.19) N = 537	0.1 (0.22) N = 260
Difference between groups [95% CI] [a]	0.0 [–0.1 to 0.0]	
SRI response		
Patients with SRI response, n/N (%)	340/554 (61.4)	135/279 (48.4)
Odds ratio [95% CI][b] versus placebo	1.68 [1.25 to 2.25], P = 0.0006	
Components of response:		
4-point reduction in SELENA-SLEDAI score, n/N (%)	345/554 (62.3)	137/279 (49.1)

	BLISS SC	
Odds ratio [95% CI][b] versus placebo	1.69 [1.26 to 2.27]	
No worsening in PGA, n/N (%)	450/554 (81.2)	203/279 (72.8)
Odds ratio [95% CI][b] versus placebo	1.61 [1.15 to 2.27]	
No new 1A/2B BILAG domain scores, n/N (%)	448/554 (80.9)	207/279 (74.2)
Odds ratio [95% CI][b] versus placebo	1.46 [1.04 to 2.07]	
Flares		
Patients with severe flare, week 52, n/N (%)	59/556 (10.6)	51/280 (18.2)
Hazard ratio [95% CI][c] versus placebo	0.51 [0.35 to 0.74], P = 0.0004	
Patients with any flare, week 52, n/N (%)	241/509 (47.3)	136/248 (54.8)
Hazard ratio [95% CI][c] versus placebo	0.80 [0.65 to 0.99]	
Prednisone dosage		
Prednisone reduction by ≥ 25% from baseline to ≤ 7.5 mg/day during week 40 through week 52, n/N (%)	61/335 (18.2)	20/168 (11.9)
Odds ratio [95% CI][b] versus placebo	1.65 [0.95 to 2.84], P = 0.0732	
PGA		
PGA percent change from baseline, mean (SD) baseline	1.58 (0.429) N = 554	1.54 (0.446) N = 279
Mean (SD) % change from baseline to week 52	–51.19 (37.501)	–37.72 (44.336)
LSM (SE) % change	–47.87 (2.441)	–35.10 (2.914)
Treatment difference [95% CI] versus placebo[d]	–12.77 [–18.46 to –7.09]	
Symptoms		
Patients with improvement (≥ 4) in FACIT week 52, n/N (%)	246/554 (44.4)	101/280 (36.1)
Odds ratio [95% CI] versus placebo[e]	1.42 [1.05 to 1.94]	
Harms		
Patients with > 0 serious AEs, N (%)	60 (11)	44 (16)
AE resulted in study agent discontinuation, N (%)	40 (7)	25 (9)
Post-injection systemic reactions	38 (7)	25 (9)
Psychiatric AEs	35 (6)	32 (11)
All infections of special interest	30 (5)	21 (8)

ACR = American College of Rheumatology; AE = adverse event; CI = confidence interval; FACIT = Functional Assessment of Chronic Illness Therapy; LSM = least squares mean; PGA = Physician Global Assessment; SD = standard deviation; SE = standard error; SELENA = Safety of Estrogen in Lupus Erythematosus National Assessment; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SLICC = Systemic Lupus International Collaborating Clinics; SRI = Systemic Lupus Erythematosus Responder Index.

[a] Analysis of covariance (ANCOVA) model comparing belimumab and placebo with covariates for treatment group, baseline SLICC/ACR Damage Index score, baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

[b] Odds ratio (95% confidence interval) and P value are from a logistic regression model for the comparison between belimumab and placebo with covariates treatment group, baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

[c] From Cox proportional hazards model for the comparison between belimumab and placebo adjusting for baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

[d] Analysis of covariance (ANCOVA) model comparing belimumab and placebo with covariates for treatment group, baseline PGA score, baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

[e] Odds ratio (95% confidence interval) and P value are from a logistic regression model for the comparison between belimumab and placebo with covariates treatment group, baseline FACIT-Fatigue score, baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

Source: Clinical Study Report for BLISS SC.[2]

Critical Appraisal

Key critical appraisal issues included the lack of an active comparator and the fact that health-related quality of life, a key outcome from a patient perspective, was not assessed in the DB phase of the included study. There is also some uncertainty as to the clinical meaningfulness of the thresholds used for the components of the composite primary outcome (SRI). There was a relatively large number of withdrawals across both groups, and numerically fewer withdrawals in the belimumab than in the placebo group (17% versus 24% of patients, respectively). The population enrolled into BLISS SC may have been a selected population, as those with more severe forms of SLE (i.e., those with severe renal or CNS involvement) were excluded from the study.

Indirect Comparisons

Description of Studies

Three indirect treatment comparisons (ITCs) were reviewed: one provided by the sponsor and two that were published, by Lee and Song[3] and by Tian et al.[4] The sponsor-provided and Lee and Song[3] analyses were similar in that they compared the efficacy of the two different formulations of belimumab (IV versus SC) to each other, and both included the same studies. Three of these studies are also reviewed in the Other Studies section of this report (BLISS SC, BLISS 52, BLISS 76). The ITC by Tian et al.[4] compared safety of all available agents (immunosuppressive drugs, biologics, and glucocorticoids [GCs]) used in the treatment of patients with SLE.

Efficacy Results

The two ITCs that compared the formulations of belimumab had different findings. The sponsor-submitted ITC did not favour either the IV or SC formulation of belimumab for SRI response, for patients with a reduction of four points or more in SELENA-SLEDAI score at week 52, and for severe flares over 52 weeks, while the ITC by Lee and Song[3] found that the IV formulation was favoured over the SC formulation for SRI response at 52 weeks. The ITC by Lee and Song[3] may have reported placebo responses from BLISS 52 incorrectly, and this appears to have been the source of the different responses found between the two formulations. Overall findings from the sponsor-submitted ITC suggest there is no difference in efficacy between the SC and IV formulations of belimumab in a population of patients with SLE who have high disease activity (HDA).

Harms Results

The published ITC that focused on safety, Tian et al.,[4] found that no treatment was favoured among belimumab SC and other drugs used for SLE (immunosuppressants, biologics [including the IV formulation of belimumab], and corticosteroids) for mortality, SAEs, AEs, serious infection, or WDAEs. Lee and Song[3] reported that no treatment was favoured among the two formulations of belimumab and placebo with respect to SAEs. Overall findings from ITCs suggest that there are no differences in harms between belimumab SC and other drugs used for SLE.

Critical Appraisal

The sponsor-submitted ITC and the published Lee and Song[3] ITC both had limited networks, as they assessed only the two formulations of belimumab. Therefore, with respect to efficacy, there are no comparisons to the other drugs used to manage SLE.

There was variation between placebo responses between trials involving the IV and SC formulations of belimumab, and this may reflect different background therapies between these trials, or other factors that may have modified the effect. The published ITC by Lee and Song,[3] as noted, may have reported incorrect data, and this appears to have contributed to the differing results found between this published ITC and the sponsor-submitted ITC. For the ITC that focused on safety, Tian et al., there appears to have heterogeneity with respect to baseline duration of disease and SLEDAI score, as well as differences in duration of follow-up. Furthermore, many of the included trials had small sample sizes, all of which contribute to a high degree of uncertainty with respect to the results from this analysis.

Other Relevant Evidence

Description of Studies

Other relevant evidence included the pivotal DB RCTs of the IV formulation of belimumab, which was approved by Health Canada in 2011, and the open-label extensions to both the IV and SC formulations of belimumab. BLISS 52 and BLISS 76 were multinational, manufacturer-sponsored phase III studies that compared belimumab 1 mg/kg and 10 mg/kg to placebo in patients with active SLE (SELENA-SLEDAI score ≥ 6 at screening) and a positive ANA or anti–double-stranded DNA (anti-dsDNA), who were on a stable background regimen for SLE (prednisone [0 mg to 40 mg per day], NSAIDs, antimalarial drugs, or immunosuppressants for at least 30 days). BLISS 52 had a 52-week DB treatment period, and BLISS 76 was 76 weeks. The primary outcome in both studies was patients with an SRI response at 52 weeks. Major secondary outcomes in BLISS 52 included the percentage of patients with at least a four-point reduction from baseline in SELENA-SLEDAI score at week 52, mean change in PGA score at week 24, mean change in the Short Form (36) Health Survey (SF-36) physical component summary (PCS) at week 24, and proportion of patients with average reduction in prednisone dosage of at least 25% from baseline to 7.5 mg per day or less during weeks 40 to 52. For BLISS 76, major secondary outcomes were SRI response rate at week 76, percentage of patients with at least a four-point reduction from baseline in SELENA-SLEDAI score at week 52, change in PGA score at week 24, change in SF-36 PCS at week 24, and percentage of patients with mean prednisone dosage decrease of 25% or more from baseline to 7.5 mg per day or less during weeks 40 to 52.

In addition to the extension to BLISS SC, there was an extension (LSBL02) of a phase II dose-ranging study that followed patients for 13 years, extensions of the BLISS 52 and BLISS 76 studies, and pooled analyses of US and non-US patients enrolled in BLISS 52 and BLISS 76. There was also a long-term study that used propensity-score–matched data to assess organ damage with belimumab plus standard of care versus standard of care alone, with a follow-up of at least five years.

Efficacy Results

There were more patients receiving belimumab than patients receiving placebo achieving an SRI response at 52 weeks in both BLISS 52 (58% versus 44%; OR = 1.83 [95% CI, 1.30 to 2.59], P = 0.0006) and BLISS 76 (43% versus 34%; OR = 1.54 [95% CI, 1.08 to 2.19], P = 0.017), and these differences were statistically significant in both studies. When SRI responses were assessed at 76 weeks, a secondary outcome of BLISS 76, there was no difference between groups. There were no statistically significant differences between belimumab and placebo with respect to the percentage of patients achieving a reduction in

prednisone dosage of 25% or more to 7.5 mg per day or less, or in change from baseline in SF-36 PCS at week 24 in either study. Otherwise, there were more patients receiving belimumab who achieved a four-point reduction in SELENA-SLEDAI score at 52 weeks (but not at 76 weeks in BLISS 76), and a greater reduction in PGA score for belimumab at 24 weeks in BLISS 52 (but not BLISS 76).

In the extension to BLISS SC, 76% of the 435 patients who continued on belimumab achieved an SRI response, while 16% of those who switched from placebo to belimumab achieved an SRI response during the six-month extension. Of the patients taking at least 7.5 mg of prednisone daily, about 20% could reduce their dose below 7.5 mg daily by the end of the six-month extension.

In LSBL02, 13% of patients who were on corticosteroids at baseline could discontinue corticosteroids for the remainder of the study. Patients were assessed at eight weeks and at 24 weeks after discontinuing the study, and 62% and 64% were SRI responders at those time points, respectively. In the pooled extensions to BLISS 52 and BLISS 76, 95% of patients had no change in their SDI score at year 0 to 1, while 83% of patients had no change in their SDI score at six years or more.

In the study that used propensity-matched scoring, results suggest a smaller deterioration in SDI scores with belimumab plus standard of care versus standard of care alone after five years. After at least one year, patients receiving belimumab were less likely to progress to organ damage. While these results provide some preliminary evidence of long-term comparative benefits of belimumab on a single outcome, it was a retrospective study that requires confirmation by additional studies.

Harms Results

There were no clear numerical differences between belimumab and placebo in the percentage of patients experiencing an AE, SAE, or WDAE in either BLISS 52 or BLISS 76. Infections occurred in 74% of belimumab and 69% of patients receiving placebo in BLISS 76 and in 67% of belimumab and 64% of patients receiving placebo in BLISS 52. Infusion reactions occurred in 14% of belimumab versus 10% of patients receiving placebo in BLISS 76 and in 14% of patients in each of the belimumab and placebo groups in BLISS 52.

In the extension to BLISS SC, 49% of patients experienced an AE, 6% experienced an SAE, and 3% withdrew due to an AE over the course of the six-month study.

Critical Appraisal

Data from the extensions are limited by the lack of a control group and by the continual attrition of patients with the progression of the studies. For example, LBSL99 started with 296 patients in year 1, but had 88 patients in the study by 11 years and later.

Conclusions

One multinational manufacturer-sponsored DB RCT of belimumab SC was included in this review. In a population of patients with active SLE (SELENA-SLEDAI score of 8 or more) belimumab reduced disease activity after 52 weeks compared to placebo, as measured by SRI response. Although belimumab reduced the risk of a severe flare over this time period, it did not elicit a reduction in prednisone dosage. Chronic use of corticosteroids contributes to morbidity in patients with SLE due to the severe adverse effects of these drugs. Health-related quality of life was not studied, and this is an important gap in a study of a severe, chronic multi-system disorder such as SLE. The duration of the study was not sufficient to study the effects of belimumab in preventing organ damage. Findings from ITCs suggest there is no difference in efficacy between the SC and IV formulations of belimumab, and no difference between belimumab and other commonly used drugs for SLE (immunosuppressants, corticosteroids, and biologics) with respect to harms, although the latter finding has a high degree of uncertainty. Data from the DB phase as well as the extensions to the SC and IV formulations do not suggest issues with tolerability or safety, although the extensions were limited by a lack of control group.

Introduction

Disease Background

Lupus is an autoimmune disease characterized by inflammatory processes that can occur in various tissues and organs of the body. Approximately one in a thousand Canadians are afflicted with lupus, and the most common form of lupus is SLE. Women are more commonly afflicted than men, and women are particularly likely to experience onset between the ages of 15 and 45. The precise etiology and pathophysiology are unknown.[5] Because lupus affects so many systems, its symptoms can vary greatly from patient to patient. Patients can experience fatigue and joint pain, which can be disabling, as well neurological, renal, and cardiovascular sequelae, rash, and a variety of other symptoms. The disease has a variable course, and patients can cycle among a chronic state, flares (acute worsening of their condition), and remission. The unpredictability of their condition takes a toll on patients, many of whom are unable to maintain a job or schooling because of their disease. According to the clinical experts consulted by CDR on this review, patients with SLE are mainly treated by rheumatologists, although specialists in immunology might also be involved in certain situations.

Standards of Therapy

SLE is treated with medications taken acutely on an as-needed basis, with corticosteroids for rashes (topically) and systemtically for joint pain, with NSAIDs for inflammation, and with medications taken on a chronic basis to keep the disease under control. The first-line chronically administered agents are antimalarial drugs, although their mechanism in managing the disease is unknown. Given that SLE is an autoimmune disorder, immunosuppressants also play an important role, and a variety of these are used (azathioprine, mycophenolate, and cyclosporine). These drugs are all approved for other uses. When the IV formulation of belimumab was approved in 2011, it was marketed as the first new drug for SLE in 50 years. Immunosuppressants are well known for toxic effects, such as serious infections and certain malignancies, and present significant tolerability issues for patients. Corticosteroids are well known for toxic effects, such as osteoporosis, psychiatric issues, cataracts, glaucoma, diabetes, hypertension, and many others, particularly when used chronically. Thus, although they are relied upon for flares, chronic use is avoided as much as possible.

Drug

Belimumab reduces survivability of B cells by inhibiting the B-lymphocyte simulator. B cells are believed to play an important role in the pathophysiology of SLE, perhaps via abnormal activation and differentiation.[1] Belimumab is indicated in addition to standard therapy for reducing disease activity in adult patients with active, autoantibody-positive SLE.[6] The Health Canada–recommended dose is 200 mg once weekly by SC injection. The Health Canada–approved product monograph further states that discontinuation of treatment with belimumab should be considered if there is no improvement in disease control after six months of treatment.[6] The IV formulation of belimumab was previously reviewed by CDR in 2012, also for SLE.

The sponsor-requested criteria for reimbursement is for the treatment of patients with SLE who meet the following eligibility criteria:

- adult patients aged 18 years or older; and

- who have active, antibody-positive SLE; and
- are currently receiving standard therapy; and
- have a disease activity SELENA-SLEDAI score of 8 or higher.

If no improvements are observed in a patient's SLE disease activity and/or symptoms after six months, use should be discontinued.

Table 2: Key Characteristics of Belimumab, Antimalarial Drugs, Corticosteroids, and Immune Suppressants

	Belimumab	Antimalarial drugs	Corticosteroids	Immune suppressants
Mechanism of action	Inhibits B cells by inhibiting the B-lymphocyte stimulator	Mechanism in treating SLE unknown	Possess both and anti-inflammatory and immune-modulating effects through various mechanisms	By various mechanisms, suppress immune responses Examples: • Azathioprine • Cyclophosphamide • Cyclosporine • Methotrexate • Mycophenolate mofetil • Tacrolimus
Indication[a]	In addition to standard therapy for reducing disease activity in adult patients with active, autoantibody-positive SLE	Indicated for malaria prophylaxis Used off-label for SLE	Many indications Off-label for SLE	Most are indicated for preventing organ transplant rejection Used off-label for SLE
Route of administration	SC or IV injection	Oral	Oral, parenteral	Oral, parenteral
Recommended dose	SC: 200 mg/mL once weekly IV infusion: 10 mg/kg every 2 weeks for the first 3 doses then every 4 weeks thereafter	Various doses depending on drug	Various doses depending on drug	Various doses depending on drug
Serious adverse effects or safety issues	Hypersensitivity reactions	Blood dyscrasias Retinopathy	• Osteoporosis • Infections • Cataracts • Glaucoma • Mood disturbances • Hypertension • Cushing syndrome • Hyperglycemia • Ulcer	• Infections • Gastrointestinal adverse effects • Blood dyscrasias • Neoplasia • Fetal harm in pregnancy

SC = subcutaneous; SLE = systemic lupus erythematosus.

[a] Health Canada–approved indication.

Source: eCPS Product Monographs.[7]

Stakeholder Engagement

Patient Group Input

This section was prepared by CADTH staff based on the input provided by patient groups.

1. Brief Description of Patient Group(s) Supplying Input

The Canadian Arthritis Patient Alliance (CAPA) and The Arthritis Society (TAS) submitted a joint patient input for this review. The CAPA is a national grassroots organization that provides education and advocacy for Canadians with arthritis. It is virtual, with no physical location, and communicates with members through a website, quarterly newsletter, email, and social media. The TAS is a health charity founded in 1948. It is the largest non-government funder of arthritis research in Canada and provides education, support, and programs for Canadians with arthritis.

In the past two years, CAPA reported receiving financial payment from Amgen, Abbvie, Janssen, Pfizer, Purdue, Roche, Sanofi, and UCB. TAS reported receiving financial payment from Amgen, Abbvie, Bayer Healthcare, BMS, Celgene, Eli Lilly, IMC, Janssen, Merck, Novartis, Pfizer, and Sanofi. GlaxoSmithKline (GSK), the manufacturer of Benlysta, connected CAPA to organizations that focus exclusively on supporting patients with SLE, and these organizations helped to circulate the survey used to inform this submission.

2. Condition-Related Information

The CAPA and TAS developed a survey that was distributed to patients with SLE. The design of the survey was informed by the lived experience of CAPA board members, who all have various forms of arthritis. The survey was distributed via email and social media, through Canadian networks and communities, such as Lupus Canada, regional SLE support groups, and rheumatologists who run lupus clinics in Canada. The survey was open from May 22, 2019, to June 12, 2019. Reponses were received from 14 people, of whom an equal number have lived with SLE for shorter duration (less than five years) or for longer duration (more than 20 years).

SLE is a chronic autoimmune disease that affects many areas of the body (i.e., skin, joints, kidneys, heart, lungs, blood vessels, nervous system) and results in a range of symptoms, such as swelling, pain, rash, mouth sores, fatigue, and cognitive impairment. The severity of SLE can range from mild to very severe, and is characterized by periods of flares, which can be incapacitating, and remission. One patient described their experience with the disease: "I have joint and muscle pain, and difficulty walking very far... I have shortness of breath… I had a mild heart attack caused by lupus, as myocarditis. I also have recently been told I have asthma." Another patient mentioned, "Complete exhaustion, facial rash, body rash, sore swollen joints, sore muscles, problems with eyesight because of medications, lung problems, I could go on forever." Patients may feel that they are not in control of the disease because flares can occur unpredictably and last from a few hours to months.

The disease has a significant impact on quality of life. It limits patients' abilities to perform daily household chores, work, participate in leisure activities, and care for children and loved ones. One patient mentioned that the disease "makes it difficult to keep up with daily life and socially." Other patients said, "There are so many things that I have had to give up over the years…," "Been off work numerous times. Impacts entire life when in a flare," "For me I need to control the exhaustion the most as it makes it difficult to do my job", and

"Extreme fatigue, nauseous, joint pain, brain fog and lack of concentration — all of this because I find it extremely hard to complete my work, on some days I find it difficult to write a sentence." Spouses, partners, or children often must take on additional responsibilities, such as household chores and taking patients to medical appointments, to support patients with SLE. If family members are not available to provide support, patients may require help from paid caregivers. Patients with SLE are at higher risk for depression and other mental health issues.

3. Current Therapy-Related Information

There is currently no cure for SLE, and treatments are aimed at controlling inflammation and minimizing disease activity, to prevent long-term organ damage. Patients reported having tried several treatments and requiring changes to the treatment approach.

Nonsteroidal anti-inflammatory drugs (NSAIDs) are used for symptomatic pain management; however, they have several side effects, such as stomach upset and kidney function decline. These effects may mimic or further complicate the problems associated with SLE. Antimalarial medications, such as hydroxychloroquine and chloroquine, may be useful for rashes, fatigue, arthritis, and other milder symptoms of SLE. The most common side effect of these medications is stomach upset. In high doses over a long period of time, these medications may accumulate in the retina and cause loss of vision and, rarely, blindness. Corticosteroids are often used for SLE; however, they have many short-term and long-term side effects of concern. Immunomodulation drugs also have many side effects. Anticoagulants may be used in some patients with SLE who have antiphospholipid antibodies but may cause bruising or bleeding. One patient stated, "I used prednisone to begin with...side effects...mood swings, moon face, unable to sleep, agitated at times. Plaquenil...upset stomach sometimes. Methotrexate injection weekly...nausea, weight loss and had to take other medications to control the nausea. Extremely difficult to tolerate...was unable to go out the day I took it."

Patients reported using nonpharmacological treatments, such as physiotherapy, occupational therapy, massage therapy, counselling, and acupuncture, which may be helpful for pain and fatigue. However, these treatments may not be easily accessible to all patients, due to restrictions in reimbursement, lack of availability, or lengthy wait lists.

Patient group input suggests there is an unmet need in this therapeutic area, given that the currently available treatments for SLE have many side effects that are often difficult for patients to tolerate. They also have variable responses on the disease, with some patients responding and others not. Some patients may derive no benefit at all from the available treatments. Also, treatments may manage the disease for a short time, but may need to be switched to another agent due to adaptation of the patient's immune system. Patients therefore need access to many different treatment options to manage this lifelong condition.

4. Expectations About the Drug Being Reviewed

Patients would like treatments that reduce pain and fatigue, organ involvement, disease complications, and lupus nephritis, and that increase mobility, productivity, and ability to work, to carry out activities of daily living and social roles, and to accomplish caregiving and parenting tasks.

Six patients had experience with belimumab. These patients experienced an overall decrease in disease symptoms, such as pain and fatigue, and increased participation in activities of daily living. Patients also mentioned that belimumab allowed them to reduce doses of mycophenolate mofetil and prednisone. Side effects included headache and

tiredness but were minimal. Of the five patients who responded to the question, all indicated that they would still take belimumab, despite the Health Canada warning that belimumab may cause depression, suicidal behaviour, and self-injury.

The administration of belimumab by a three-hour infusion was mentioned as a concern because workplaces do not always allow for time off, which may make it difficult for patients to work full-time while on the medication. Also, the self-injector requires refrigeration, which may restrict travel for some patients.

Clinician Input

All CADTH review teams include at least one clinical specialist with expertise in the diagnosis and management of the condition for which the drug is indicated. Clinical experts are a critical part of the review team and are involved in all phases of the review process (e.g., providing guidance on the development of the review protocol, assisting in the critical appraisal of clinical evidence, interpreting the clinical relevance of the results, and providing guidance on the potential place in therapy). The following input was provided by two clinical specialist(s) with expertise in the diagnosis and management of SLE.

Description of the Current Treatment Paradigm for the Disease

The clinical experts emphasized that the management of SLE is highly variable, a reflection of both the range in clinical features and the unpredictable course that characterizes the disease. The clinical manifestations of SLE can range from relatively benign (e.g., mild rash, oral ulcers, and arthralgia) to life-threatening illness (e.g., severe renal and nervous system disease) requiring aggressive immunosuppression and high doses of corticosteroids. When deciding on a treatment plan, clinical and laboratory manifestations, comorbidities, and the course of the disease, which can transition among acute exacerbations, termed “flares”, and remission or low disease activity, are taken into account.

Nonpharmacological measures are universally recommended, such as avoidance of or protection from direct sunlight, smoking cessation, a balanced diet, exercise, stress reduction, and work/life balance. Pharmacologic interventions can be divided into those that are used for symptom control (NSAIDs, analgesics, low-dose corticosteroids), immune modulators (high-dose corticosteroids, antimalarial drugs, azathioprine, methotrexate, mycophenolate mofetil, cyclophosphamide, and cyclosporine/tacrolimus), and drugs that address comorbidities (hypertension, hyperlipidemia, diabetes, anxiety, and depression). Belimumab is the only biologic with an approved indication for use in SLE, although rituximab is frequently used off-label for this condition.

Treatment Goals

Goals of therapy include controlling the underlying autoimmune and inflammatory response; preventing flares; removing the need for chronic or intermittent corticosteroids or chronic immune suppression; preventing long-term organ damage and toxic effects of SLE and its therapies; restoring health-related quality of life, including functional improvement; and avoiding or reducing risk of medication adverse effects. Accessibility of therapies, both in terms of affordability and ease of administration, and availability of medications that can be used safely during pregnancy are also goals.

Unmet Needs

With respect to unmet needs, the clinical experts identified therapeutic failure (primary or secondary); recurrent flares that cause progressive organ damage (examples include renal failure and cognitive decline); unacceptable toxic effects, including infection and damage due to corticosteroids; and lack of safety in pregnancy in a disease that is relatively common in women of child-bearing age. Approximately 60% to 70% of patients do not have a good long-term response to therapy without the intermittent or continuous use of corticosteroids. This is a significant limitation, given the high burden of adverse effects associated with this class of drugs. Patients also are frequently reluctant to increase corticosteroid doses during flares, because of their awareness of these adverse effects. Nonadherence and lack of persistence with therapy are significant issues, as flares can have serious consequences, such as renal failure. No therapies provide a long-term cure or a long-term medication-free survival in a majority of patients, and there are no therapies that specifically address underlying disease mechanisms in all patients, although targeting B cells is a logical strategy and appears to be effective in some patients.

Place in Therapy

Current evidence supports use of belimumab in milder forms of SLE. The clinical experts noted that belimumab is unlikely to be used first-line; however, as more experience is gained with the drug, this may change. Initial use of an antimalarial drug would be warranted, with or without a concomitant short course of prednisone, to achieve disease control promptly. If the patient does not respond to this approach, or has a secondary failure, belimumab could be added as an adjunct. In patients with more severe disease (such as nephritis), induction with an immunosuppressive drug would be indicated, with belimumab considered for maintenance therapy if remission is achieved.

Patient Population

The clinical experts stated that there is no known biomarker to predict which patients will respond to treatment with belimumab. They did note that the clinical trials were restricted to patients with autoantibodies, specifically antinuclear antibody (ANA) or anti-DNA.

Diagnosis of SLE should not be challenging for trained or experienced specialists, and all such specialists would have access to the laboratory assays (autoantibodies, hematology, and chemistry) and diagnostic tests (conventional and specialized imaging) needed to aid in diagnosis. Pre-clinical screening of patients for SLE is under investigation but is not currently part of the paradigm for managing SLE.

Patients most in need of an intervention with belimumab would be those with 1) active nonrenal and non-CNS involvement despite use of prednisone and an antimalarial drugs (unless intolerant or contraindicated); 2) nonrenal and non-CNS involvement who experience flares during prednisone tapering despite using an antimalarial drug (unless intolerant or contraindicated); 3) inactive nonrenal and non-CNS involvement despite using prednisone and an antimalarial drug (unless intolerant or contra-indicated) to allow for prednisone tapering; 4) inactive nonrenal and non-CNS involvement but with active nonrenal and non-CNS disease who experience flares during prednisone tapering despite using an antimalarial drug (unless intolerant or contra-indicated); 5) as an adjunct to maintain remission in those with more severe involvement (CNS and renal) to maintain remission without use of prednisone.

Some potential roles for belimumab would be 1) as adjunctive therapy for milder forms of the disease to avoid intermittent or long-term corticosteroid use, 2) to maintain a clinical response previously induced with more aggressive therapies that should not be continued long-term, and 3) as a potentially safe alternative for use in pregnancy.

Assessing Response to Treatment

A clinically meaningful response to belimumab would be signified by meaningful reduction in signs and symptoms of inflammation mediated by the SLE-aberrant immune response. Examples of these signs and symptoms are active mucocutaneous disease, serositis, inflammatory arthritis, and constitutional symptoms. Clinically relevant improvements in laboratory indices would include cytopenia, erythrocyte sedimentation rate, anti-dsDNA and C3/C4 levels, and renal indices (serum creatinine, active urine sediment, and proteinuria). A response would also be indicated by a reduction in the number and severity of flares affecting any organ system, reduction in the rate of accumulation of organ damage and reduction in risk of mortality, restored health-related quality of life, and return to full employment.

Treatment response should be assessed every six to 12 weeks for active disease. The rapidity of response depends on the organ system affected and treatment used. Symptoms in some systems (rashes, arthritis, pleurisy/pericarditis, constitutional symptoms, and cytopenia) respond more quickly, in days to weeks, than in other systems (renal, CNS), which respond in four to 12 months. Corticosteroids produce the most rapid response (days to weeks), while antimalarial and immunosuppressive drugs require two to three months for an initial response and six to 12 months for peak response.

Discontinuing Treatment

The decision to discontinue treatment should be based on lack of response, unacceptable toxic effects, patient preference, and access (cost, coverage). For example, if the disease progresses despite adequate treatment and, in particular, it progresses to renal or CNS disease, as described above, this could lead to discontinuation of treatment. Most patients are assessed for an initial response at four to six weeks following the initiation of treatment, but they could be assessed earlier if there is a clinical concern. With antimalarial drugs, response is not expected for two to three months, and further improvement can continue beyond this time. The experts had previously described which parameters would be used to assess a clinical response (remission or significant improvement in clinical features) and laboratory response (hemolytic anemia and/or significant thrombocytopenia, complement levels [C3 and C4], anti-dsDNA levels, and renal parameters [proteinuria, serum creatinine]).

Prescribing Conditions

Rheumatologists typically prescribe belimumab, as they are the specialists most likely to manage patients with SLE. Some clinical immunologists with an interest in SLE could also potentially prescribe this drug.

The drug is administered by SC injection, and, thus, most patients should be able to self-administer belimumab.

Additional Considerations

The clinical experts noted the need for new treatment modalities in SLE, and they noted that belimumab is the only biologic approved for use in this disease.

Clinical Evidence

The clinical evidence included in the review of belimumab is presented in three sections. Section 1, the Systematic Review, includes pivotal studies provided in the sponsor's submission to CDR and Health Canada, as well as studies selected according to an a priori protocol. Section 2 includes indirect evidence from the sponsor and selected from the literature that met the selection criteria specified in the review. Section 3 includes sponsor-submitted long-term extension studies and additional relevant studies that were considered to address important gaps in the evidence included in the systematic review.

Systematic Review (Pivotal and Protocol-Selected Studies)

Objectives

To perform a systematic review of the beneficial and harmful effects of belimumab 200 mg/mL SC injection in addition to standard therapy for reducing disease activity in adult patients with active, autoantibody-positive SLE.

Methods

Studies selected for inclusion in the systematic review included pivotal studies provided in the sponsor's submission to CDR and Health Canada, as well as those meeting the selection criteria presented in Table 3.

Table 3: Inclusion Criteria for the Systematic Review

Patient population	Adults with active, autoantibody-positive SLE Subgroups: • Disease activity • Prior treatment and/or response to prior treatment
Intervention	Belimumab 200 mg by SC injection once weekly in addition to standard treatments
Comparators	Standard treatment, including the following treatments as monotherapy or in combination: • Hydroxychloroquine/chloroquine • Immunosuppressants or immune modulators (azathioprine, methotrexate, mycophenolate mofetil, cyclophosphamide) • Corticosteroids • Rituximab • Belimumab IV • No treatment (placebo)
Outcomes	**Efficacy outcomes:** • Mortality • Morbidity (e.g., organ damage)[a] • Disease activity (e.g., SELENA-SLEDAI scores, SLE response index) • Health-related quality of life[a] • Reduction in symptoms (e.g., rash,[a] pain,[a] fatigue,[a] cognitive impairment,[a] depression,[a] based on validated scales) • Achievement of remission,[a] or low disease activity • Disease flare[a] • Physician Global Assessment (based on a validated scale)

	• Reduction in background corticosteroid use Activities of daily living, including ability to work or attend school (based on a validated scale) Caregiver burden (based on a validated scale) **Harms outcomes:** AEs, SAEs, WDAEs, notable harms (hypersensitivity reactions, serious infections, psychiatric issues [e.g., serious depression, suicidal ideation/behaviour, self-injury])
Study design	Published and unpublished phase III and IV RCTs

AE = adverse event; HRQoL = quality of life; RCT = randomized controlled trial; SAE = serious adverse event; SC = subcutaneous; SELENA-SLEDAI = Safety of Estrogens in Lupus Erythematosus National Assessment Trial–Systemic Lupus Erythematosus Disease Activity Index; SLE = systemic lupus erythematosus; WDAE = withdrawal due to adverse event.

[a] These outcomes were identified as being of particular importance to patients in the input received by CADTH from patient groups.

The literature search for clinical studies was performed by an information specialist using a peer-reviewed search strategy according to the PRESS (Peer Review of Electronic Search Strategies) checklist (https://www.cadth.ca/resources/finding-evidence/press).[8]

Published literature was identified by searching the following bibliographic databases: MEDLINE All (1946–) via Ovid, Embase (1974–) via Ovid, and PubMed. The search strategy was comprised of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concepts were Benlysta (belimumab) and systemic lupus erythematosus (SLE). Clinical trial registries were searched: the US National Institutes of Health's clinicaltrials.gov and the World Health Organization's International Clinical Trials Registry Search Portal (ICTRP).

Search filters were applied to limit retrieval to RCTs or controlled clinical trials. Retrieval was not limited by publication date or by language. Conference abstracts were excluded from the search results. See Appendix 2 for the detailed search strategies.

The initial search was completed on June 26, 2019. Regular alerts updated the search until the meeting of the CADTH Canadian Drug Expert Committee (CDEC) on November 20, 2019.

Grey literature (literature that is not commercially published) was identified by searching relevant websites from the following sections of the *Grey Matters: A Practical Tool For Searching Health-Related Grey Literature* checklist (https://www.cadth.ca/grey-matters):[9] Health Technology Assessment (HTA) Agencies, Health Economics, Clinical Practice Guidelines, Drug and Device Regulatory Approvals, Advisories and Warnings, Drug Class Reviews, Clinical Trials Registries, and Databases (Free). Google was used to search for additional internet-based materials. In addition, the manufacturer of the drug was contacted for information regarding unpublished studies. See Appendix 2 for more information on the grey literature search strategy.

Two CDR clinical reviewers independently selected studies for inclusion in the review based on titles and abstracts, according to the predetermined protocol. Full-text articles of all citations considered potentially relevant by at least one reviewer were acquired. Reviewers independently made the final selection of studies to be included in the review, and differences were resolved through discussion.

Findings From the Literature

A total of one study was identified from the literature for inclusion in the systematic review (Figure 1). The included studies are summarized in Table 4. A list of excluded studies is presented in Appendix 2.

Figure 1: Flow Diagram for Inclusion and Exclusion of Studies

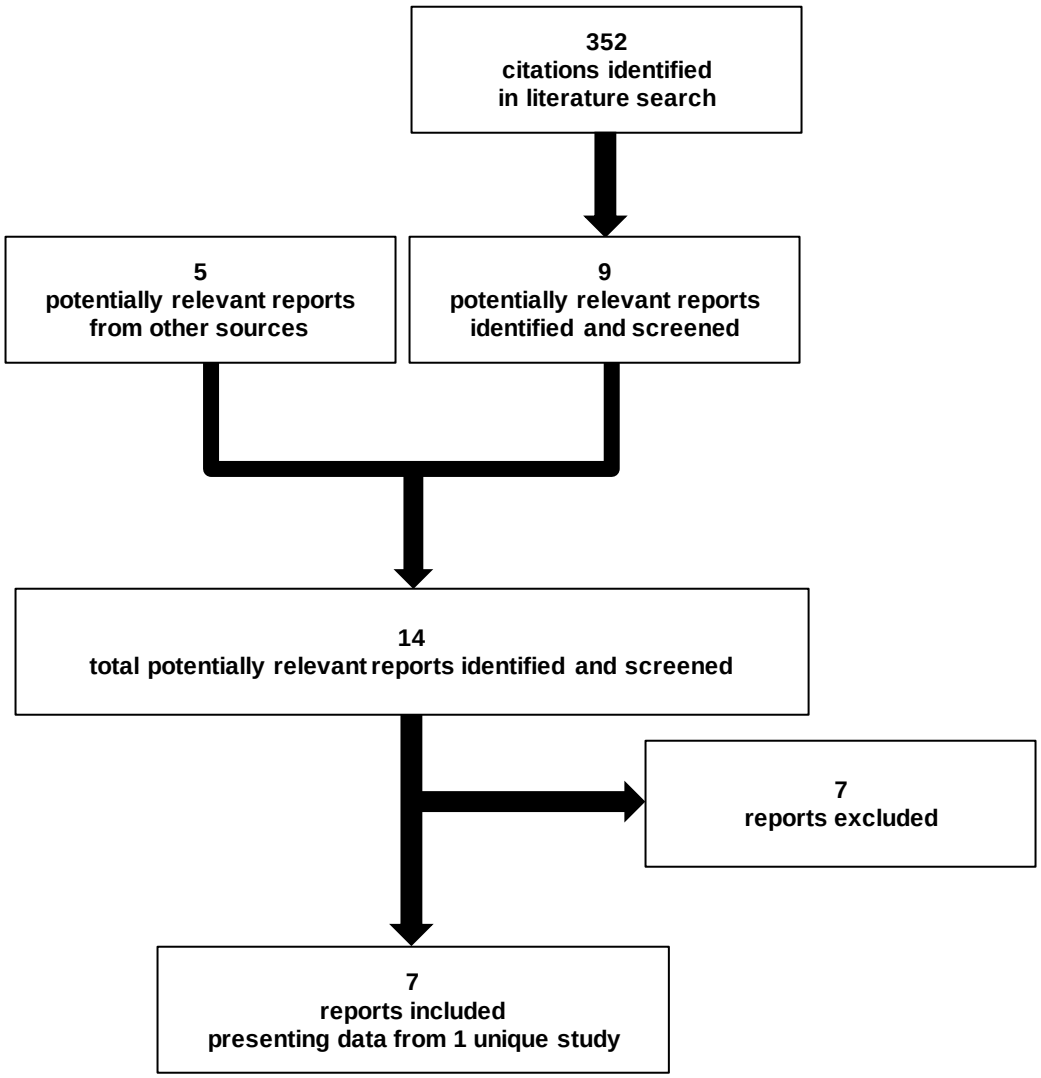


Table 4: Details of Included Studies

		BLISS SC
Designs and Populations	**Study design**	DB RCT
	Locations	177 sites: Canada, US, Europe, Americas, Asia, Australia, Israel
	Randomized (N)	N = 836
	Inclusion criteria	• SLE by ACR criteria (1997) and clinically active disease, defined as SELENA-SLEDAI disease-activity score of ≥ 8 • Unequivocally positive ANA test and/or positive anti-dsDNA serum antibody result from two independent time points within the study screening period or one positive historical test and one positive test result during screening • ANA results obtained in screening were only considered positive if ANA titre was ≥ 1:80 and/or anti-dsDNA serum antibody was ≥ 30 IU/mL • On stable SLE regimen for ≥ 30 days before day 0 • Stable regimens consisted of any of the following, used alone or in combination: • Prednisone or equivalent (0 to 40 mg/day when used in combination with other SLE treatment or 7.5 to 40 mg/day as monotherapy) • Antimalarial drugs • NSAIDs • Any immunosuppressive therapy (methotrexate, azathioprine, leflunomide, mycophenolate, calcineurin inhibitors, sirolimus, oral cyclophosphamide, 6-mercaptopurine, mizoribine, or thalidomide)
	Exclusion criteria	• Severe lupus kidney disease (proteinuria > 6 g in 24 hours or equivalent using spot urine protein-to-creatinine ratio, or serum creatinine > 2.5 mg/dL) or severe active nephritis requiring acute therapy not permitted by protocol or having required hemodialysis or high-dose prednisone or equivalent • Severe active CNS lupus • Pregnancy • Receipt of investigational agent within 60 days of day 0 for non-biologic drugs and within one year for biologics or abatacept • Anti-TNF therapy, IV cyclophosphamide, anakinra, IV immunoglobulin, prednisone > 100 mg/day or plasmapheresis within 3 months • Live vaccine within 1 month • B-cell therapy within 1 year of screening (except US, UK, Spain, Portugal, Sweden, Denmark) • Serious suicide risk or any history of suicidal behaviour and/or any suicidal ideation of type 4 or 5 on the Columbia–Suicide Severity Rating Scale (C-SSRS) in the last 2 months or those who, in the investigator's opinion, posed a significant suicide risk
Drugs	**Intervention**	Belimumab 200 mg by SC injection once weekly
	Comparator(s)	Matching placebo once weekly
Duration	**Phase**	
	Screening	5 weeks
	DB	52 weeks
	Follow-up	8 weeks after last dose of study drug (and optional 6-month extension)
Outcomes	**Primary end point**	SRI response at week 52 (4-point reduction from baseline in SELENA-SLEDAI score and no worsening [increase of < 0.30 points from baseline] in PGA, and no new BILAG A organ domain score or 2 new BILAG B organ domain scores compared to baseline at week 52)
	Other end points	Major secondary end points: • Time to first severe flare (by modified SLE Flare Index) • Percentage of patients whose mean prednisone dosage had been reduced by ≥ 25% from baseline to ≤ 7.5 mg/day during weeks 40 to 52 in patients receiving greater than 7.5 mg/day at baseline

		BLISS SC
		Other efficacy end points: *Supporting the primary efficacy end point:* • SRI by visit • Percentage of patients with no new BILAG A organ domain score or 2 new BILAG B organ domain scores versus baseline at week 52 and by visit • Percentage of patients without PGA worsening (increase of < 0.30 points from baseline) at week 52 and by visit • Percentage of patients with durable SRI week 44 to 52 • Time to first SRI maintained through week 52 • Duration of longest response among patients with at least 1 SRI response • SRI 5 to 8 at week 52 and by visit • SRI 5 to 8 was defined identically to the SRI except for using higher thresholds of improvement (SELENA-SLEDAI score reduction) for a patient to be declared a responder (e.g., SELENA-SLEDAI ≥ 5-point reduction for SR15) *Disease activity:* • Mean percent change and mean change in PGA by visit • Mean percent change and mean change in SELENA-SLEDAI score by visit *Organ-specific:* • Percentage of patients with organ improvement by SELENA-SLEDAI score by visit • Percentage of patients with organ worsening by SELENA-SLEDAI score by visit • Percentage of patients with organ improvement by BILAG by visit • Percentage of patients with organ worsening by BILAG by visit • Mean/median percent (in patients with proteinuria at baseline) and mean/median absolute (in all patients) change in proteinuria by visit • Time to renal flare over 52 weeks • Percentage of patients developing at least 1 renal flare over 52 weeks • Percentage of patients with doubling of serum creatinine (proportion of patients whose serum creatinine attains a level double that of the baseline value and is confirmed with a second measurement at least 3 weeks later) *SFI flare:* • Time to flare over 52 weeks • Time to flare after week 24 • Time to severe flare after week 24 • Rate of flare per 100 patient-years • Rate of severe flare per 100 patient-years *Steroids (based on average steroid dose between visits):* • Percentage of patients with daily prednisone reduced to ≤ 7.5 mg/day from > 7.5 mg/day at baseline by visit • Percentage of patients with daily prednisone dosage increased to > 7.5 mg/day from ≤ 7.5 mg/day at baseline by visit • Percentage of patients with increase in steroid use by visit • Percentage of patients with 50% decrease in steroid dose by visit • Percentage of patients with 50% increase in steroid dose by minimum of ≥ 5 mg/day by visit • Mean/median changes in steroid dose by visit *SLICC/ACR Damage Index (SDI):* • Mean change from baseline in SDI at week 52 • Percentage of patients with any SDI worsening (change > 0) at week 52 compared to baseline

		BLISS SC
Notes	**Publications**	Stohl et al. (2017)[10] Doria et al. (2018)[11]

ACR = American College of Rheumatology; ANA = antinuclear antibody; BILAG = British Isles Lupus Assessment Group; CNS = central nervous system; DB = double-blind; dsDNA = double-stranded DNA; NSAID = nonsteroidal anti-inflammatory drug; PGA = Physician Global Assessment; RCT = randomized controlled trial; SC = subcutaneous; SELENA = Safety of Estrogen in Lupus Erythematosus National Assessment; SLE = systemic lupus erythematosus; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SLICC = Systemic Lupus International Collaborating Clinics; SRI = Systemic Lupus Erythematosus Responder Index; TNF = tumour necrosis factor.

Note: Five additional reports were included (sponsor's submission,[12] Health Canada Reviewers Report,[13] FDA Clinical and Statistical reviews,[1,14] Clinical Study Report for BLISS SC[2]).

Source: Clinical Study Report for BLISS SC.[2]

Description of Studies

One manufacturer-sponsored, multinational (177 centres in 30 countries, including Canada) DB RCT conducted between 2011 and 2015 was included in this review. BLISS SC randomized 836 patients with active SLE (SELENA-SLEDAI ≥ 8), 2:1, to either belimumab 200 mg (N = 556) or placebo (N = 280) once weekly by SC injection, over a 52-week period. SELENA-SLEDAI is a measure of disease activity at time of visit or in the preceding 10 days. It consists of 24 weighted clinical and laboratory variables, with total possible score ranging from 0 to 105. The primary outcome was the percentage of patients with an SRI response at 52 weeks, a composite that was defined by a reduction of four points or more from baseline in SELENA-SLEDAI score, no worsening (increase of < 0.30 points from baseline in PGA), and no new BILAG A organ domain score or two new BILAG B organ domain scores compared with baseline. The two secondary outcomes were time to first severe flare over 52 weeks and percentage of patients with an average prednisone reduction of 25% or more from baseline to 7.5 mg per day or less during weeks 40 through 52.

Randomization was conducted by interactive web response system and was stratified by race (black versus other), screening values for SELENA-SLEDAI score (≤ 9 versus ≥ 10) and complement level (low C3 and/or C4 versus no low C3 or C4).

There was a five-week screening period during which patients were assessed for eligibility to enter the study. Approximately 41% of patients failed screening, either failing to meet inclusion criteria or meeting exclusion criteria.

Populations

Inclusion and Exclusion Criteria

Patients were to have been diagnosed with SLE according to 1997 ACR criteria and had clinically active disease (a SELENA-SLEDAI disease-activity score of ≥ 8) at screening. The 1997 ACR criteria take into account a variety of signs and symptoms associated with SLE: malar rash, discoid rash, photosensitivity, oral ulcers, arthritis, serositis (pleuritis, pericarditis), renal disorders (persistent proteinuria or cellular casts), neurologic disorders (seizures, psychosis), hematologic disorders (hemolytic anemia or leukopenia, lymphopenia, thrombocytopenia), and immunologic disorders (anti-DNA, anti-Smith [anti-Sm], or positive antiphospholipid antibodies, and abnormal ANA). In identifying patients for a clinical trial, SLE is considered any four of these 11 criteria. Patients had to be on stable SLE regimen for 30 days before day 0, consisting of prednisone or equivalent (0 mg to 40 mg per day when used in combination with other SLE treatment, or 7.5 mg to 40 mg per

day when used alone), NSAIDs, antimalarial drugs, or any immunosuppressive therapy. Patients with severe lupus kidney disease (proteinuria > 6 g in 24 hours or equivalent according to a spot urinary protein-to-creatinine ratio or a serum creatinine level > 2.5 mg/dL), or severe active lupus nephritis that required acute therapy, were not allowed in the protocol. Patients with CNS lupus or who were pregnant were excluded. Patients at serious risk of suicide or with a history of suicidal behaviour in the past six months and/or suicidal ideation of type 4 or 5 on the Columbia–Suicide Severity Rating Scale in the past two months or who, in the investigator's opinion, posed a significant suicide risk were also excluded.

Baseline Characteristics

Included patients had a mean age of approximately 39 years and were predominantly women (95%). The majority (60%) were white, and 22% were Asian. Patients had had SLE for 6.6 years, on average, and the majority (72%) had BILAG organ involvement score of at least 1A or 2B. Almost all patients had SELENA-SLEDAI scores of at least nine, indicating active disease, and the mean score was 10.4. Approximately 18% had had a flare, and 1.4% had had a severe flare, during the five-week screening period.

The majority of patients were receiving corticosteroids at baseline (87% versus 86%, belimumab versus placebo groups, respectively), and antimalarial drugs were used by 70% versus 68% of patients, belimumab versus placebo groups, respectively, while 36% versus 33% were receiving steroids in combination with antimalarial drugs alone. Nearly half of the patients were receiving immunosuppressants (44% versus 49%, belimumab versus placebo groups, respectively), with azathioprine being the one most commonly received.

There were fewer patients receiving belimumab with BILAG A/B scores compared to patients receiving placebo (70% versus 75% of patients), and fewer patients receiving belimumab with proteinuria of 2 mg in 24 hours or higher (3% versus 7%). Otherwise, there were no clear differences in reported baseline characteristics between groups in the study.

Table 5: Summary of Baseline Characteristics

Characteristic (ITT population)	Belimumab N = 556	Placebo N = 280
Age, years, mean (SD)	38.1 (12.10)	39.6 (12.61)
Female, N (%)	521 (93.7)	268 (95.7)
Race, n (%)		
White	336 (60.4)	166 (59.3)
Asian	119 (21.4)	63 (22.5)
African-American/African	56 (10.1)	30 (10.7)
American Indian or Alaska Native	43 (7.7)	21 (7.5)
Native Hawaiian or other Pacific islander	2 (0.4)	0
Multiracial	6 (1.1)	3 (1.1)
BMI, kg/m^2, mean (SD)	25.96 (6.293)	26.48 (7.169)
SLE disease duration, years, mean (SD)	6.4 (6.60)	6.8 (6.83)
BILAG organ domain involvement score, n (%)		
At least 1A or 2B	388 (69.8)	210 (75.0)
At least 1A	87 (15.6)	51 (18.2)

Characteristic (ITT population)	Belimumab N = 556	Placebo N = 280
At least 1B	499 (89.7)	258 (92.1)
No A or B	29 (5.2)	13 (4.6)
SELENA-SLEDAI score (range 0 to 105) category,[a] n (%)		
0 to 3	4 (0.7)	0
≤ 9	200 (36.0)	112 (40.0)
10 or 11	161 (29.0)	74 (26.4)
≥ 12	191 (34.4)	94 (33.6)
SELENA-SLEDAI score, mean (SD)	10.5 (3.19) N = 556	10.3 (3.04) N = 280
SLE Flare Index,[a] n (%)		
≥ 1 flare	92 (16.5)	57 (20.4)
≥ 1 severe flare	8 (1.4)	4 (1.4)
PGA category, n (%)		
0 to 1	40 (7.2)	19 (6.8)
> 1 to 2.5	507 (91.2)	255 (91.1)
> 2.5	7 (1.3)	5 (1.8)
Missing	2 (0.4)	1 (0.4)
PGA, mean (SD) score	1.6 (0.43) N = 554	1.5 (0.45) N = 279
SLICC/ACR Damage Index score, mean (SD)	0.6 (0.99) N = 556	0.7 (1.17) N = 280
Proteinuria category (g in 24 hours), n (%)		
≥ 2	19 (3.4)	20 (7.1)
Proteinuria level (g in 24 hours), n, mean (SD)	0.4 (0.71) N = 556	0.4 (0.84) N = 280
Anti-dsDNA, positive (≥ 30 IU/mL), n (%)	404 (72.7)	193 (68.9)
ANA, positive (≥ 80 titre), n (%)	492 (88.6)	254 (90.7)
Anti-dsDNA and/or ANA, positive, n (%)	531 (95.5)	270 (96.4)
Average daily prednisone dosage, n (%)		
0 mg/day	75 (13.5)	39 (13.9)
> 0 to ≤ 7.5 mg/day	146 (26.3)	73 (26.1)
> 7.5 mg/day	335 (60.3)	168 (60.0)
Daily prednisone dosage, mg, mean (SD)	10.8 (8.21)	11.2 (9.09)
Medications at baseline, n (%)		
Steroids	481 (86.5)	241 (86.1)
Antimalarial drugs	391 (70.3)	189 (67.5)
Immunosuppressants	244 (43.9)	137 (48.9)
Aspirin	94 (16.9)	45 (16.1)
NSAIDs	124 (22.3)	72 (25.7)
Combinations used at baseline, n (%)		
Steroid and antimalarial drug only	201 (36)	93 (33)

Characteristic (ITT population)	Belimumab N = 556	Placebo N = 280
Steroid and immunosuppressant and antimalarial drug	133 (24)	67 (24)
Steroid and immunosuppressant only	88 (16)	50 (18)
Steroid only	59 (11)	31 (11)
Antimalarial drug only	44 (8)	16 (6)
Immunosuppressant and antimalarial drug only	13 (2)	13 (5)
Immunosuppressant only	10 (2)	7 (3)

ACR = American College of Rheumatology; ANA = antinuclear antibody; BILAG = British Isles Lupus Assessment Group; BMI = body mass index; dsDNA = double-stranded DNA; ITT = intention-to-treat; NSAID = nonsteroidal anti-inflammatory drug; PGA = Physician Global Assessment; SD = standard deviation; SELENA = Safety of Estrogen in Lupus Erythematosus National Assessment; SLE = systemic lupus erythematosus; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SLICC = Systemic Lupus International Collaborating Clinics.

[a] Assessed during five-week screening period.

Source: Clinical Study Report for BLISS SC.[15]

Interventions

Study drugs, belimumab or matching placebo, were administered by SC injection with a pre-filled syringe and an autoinjector device beginning on day 0 and then every week for the 51 weeks of treatment. The first two doses were administered under supervision, and patients were observed for three hours post-administration. Subsequent doses could then be self-administered by the patient, although this was at the discretion of the investigator. The injection site was rotated weekly between the abdomen and the thigh. Patients used a logbook to record the date, injection site, and dose administered.

To be enrolled in the study, patients had to be on a stable SLE regimen for at least 30 days before day 0. The regimen could include any of the following, alone or in combination: prednisone or equivalent, antimalarial drugs, NSAIDs, or any immunosuppressant. Dose adjustments were allowed, as required clinically, throughout the study, although if patients had adjustments to certain medications, they were deemed a treatment failure and withdrawn from the study. For example, adding a new antimalarial drug after week 16 or increasing the dose of an existing antimalarial drug above what it was at week 0 or week 16 would result in the patient being declared a treatment failure. Switching one antimalarial drug for another was allowed at any time. Corticosteroid doses could be increased as needed during the first 24 weeks. However, they needed to be returned to within 25% (or 5 mg) of the baseline dose; otherwise, the patient was considered a treatment failure. After week 24, a patient with an increase in corticosteroid dose of more than 25% or 5 mg more than the baseline dose for SLE activity was deemed a treatment failure. Any intra-articular injections of corticosteroids within eight weeks of the week 52 visit would result in a patient being deemed a treatment failure. Patients were not allowed to begin any new immunosuppressants after baseline, and doses of currently received immunosuppressants could not be increased after week 16. These situations would lead to declaring the patient a treatment failure. Prohibited medications included other investigational drugs, anti–tumour necrosis factor (anti-TNF) therapy, other biologics, IV immunoglobulin, IV cyclophosphamide, and plasmapheresis. Patients who started a prohibited medication during the study were considered treatment failures.

Outcomes

The percentage of patients achieving SRI response at 52 weeks was the primary outcome of BLISS SC. It is a composite end point that is considered achieved when all three of the

following components are met: 1) reduction from baseline in SELENA-SLEDAI score of four points or more and 2) no worsening (increase of < 0.3 points from baseline) in PGA, and 3) no new BILAG A organ domain score or two new BILAG B organ domain scores compared with baseline.

The SELENA-SLEDAI is a measure of disease activity consisting of 24 items across nine organ systems, which are scored based on the time of visit or the preceding 10 days (see Appendix 4 for detailed summary). The items are answered Yes or No (presence/absence), and answers are weighted to arrive at a total score (range 0 to 105). The items consist of the following: seizure, psychosis, organic brain syndrome, visual disturbance, cranial nerve disorder, lupus headache, cerebrovascular accident, vasculitis, arthritis, myositis, urinary casts, hematuria, proteinuria, pyuria, rash, alopecia, mucosal ulcers, pleurisy, pericarditis, low complement, increased DNA binding, fever, thrombocytopenia, and leukopenia. In BLISS SC, a spot urine protein-to-creatinine ratio was used to determine proteinuria (to be assigned four points for proteinuria, the spot urine assessment had to show proteinuria > 1 g in 24 hours equivalent, or new-onset or recurrent proteinuria > 0.5 g in 24 hours equivalent, or > 0.5 g in 24 hours equivalent increase from previously documented value that was obtained within 26 weeks of the screening value). These values for proteinuria are important, as they are varied when using variations on the SELENA-SLEDAI, such as the SLEDAI 2K score. A reduction from baseline in SELENA-SLEDAI score of at least four points has been identified by the sponsor as clinically meaningful; however, CDR was not able to find evidence of how this threshold was determined.

BILAG measures disease activity across eight organ systems known to be affected in SLE: general, mucocutaneous, neurological, musculoskeletal, cardiorespiratory, vasculitis, renal, and hematologic (see Appendix 4 for detailed summary). It is an ordinal scale from A (most active) to E (never present). Total scores can be calculated by assigning scores of A = 9, B = 3, C = 1, D/E = 0; however, this is typically not done. BILAG is used to identify flares, as a severe flare is defined as a new organ domain score of A, and moderate flare as a new organ domain score of B.

The PGA is a visual analogue scale, scored between 0 and 3, in which physicians are asked the following question: How do you assess your patient's current disease activity?, with possible answers: 0 = none, 1 = mild, 2 = moderate, and 3 = severe. In the SLE Flare Index (SFI), a mild or moderate flare can occur, with an increase in PGA of 1 or more, and a severe flare is indicated by an increase of > 2.5 (see Appendix 4 for detailed summary). No studies were found that described how these cut points were arrived at. The threshold for "no worsening" on the PGA is identified by the sponsor as a change of < 0.3 points; however, this was derived in patients with rheumatoid arthritis. CDR did not find any evidence in the literature of a similar threshold for PGA in patients with SLE.

Time to severe flare was a secondary outcome of BLISS SC. A severe flare was defined as a change in SELENA-SLEDAI score > 12 points; or new or worse CNS SLE, vasculitis, nephritis, myositis, platelet count < 60,000, hemolytic anemia (hemoglobin [Hb] < 70 g/L or decrease in Hb > 30 g/L) that requires double of prednisone dose or prednisone dosage increase to greater than 0.5 mg/kg per day, or hospitalization; or an increase in prednisone dosage to 0.5 mg/kg per day or greater; or a new cyclophosphamide, azathioprine, methotrexate, or mycophenolate therapy for SLE; or a hospitalization for SLE; or an increase in PGA score to greater than 2.5.

The percentage of patients who could reduce their prednisone dosage by at least 25% to 7.5 mg daily or lower was a secondary outcome of BLISS SC. Patients included in this

analysis had to be on a dosage of prednisone greater than 7.5 mg per day at baseline, and this represented about 60% of the ITT population.

Exploratory outcomes included the SDI and the Functional Assessment of Chronic Illness Therapy (FACIT) Fatigue Scale (FACIT-Fatigue). The SDI scores irreversible damage, defined as an irreversible change in an organ system that has occurred since the onset of SLE and has been present for at least six months. The tool is completed by a physician and consists of 42 items in 12 domains, with a maximum score of 46 (higher scores denote more damage).[16,17] The items are rated as present or absent and, in the case of recurring events, there is a possibility of providing a rating of two or three points to an item.[16] At diagnosis of SLE, the SDI score is 0 by definition, and an SDI score of 1 or more indicates damage.[17] In BLISS SC, scores were derived from the following assessments: ocular, neuropsychiatric, renal, pulmonary, cardiovascular, peripheral vascular, gastrointestinal, musculoskeletal, skin, premature gonadal failure, diabetes, and malignancy. These were each scored one point. However, certain items were scored up to two points (cerebrovascular accident, myocardial infarction, significant tissue loss from peripheral vascular disease, infarction/resection of various gastrointestinal organs, avascular necrosis, and malignancy), and signs of end-stage renal disease could receive up to three points. No minimal clinically important difference (MCID) was identified for the SDI. The FACIT-Fatigue is 13 items assessed on a four-point Likert scale (0 = not at all, 1 = a little bit, 2 = somewhat, 3 = quite a bit, and 4 = very much), with a possible range of 0 to 52. Higher scores indicate less fatigue. The anchor-based MCID is a range of 2.5 to 8.4 points, while distribution-based MCIDs for one-third standard deviation are 3.8 to 4.6, for one-half standard deviation are 5.8 to 6.8, and for standard error of the mean are 2.7 to 2.9.

Statistical Analysis

Primary Outcome(s) of the Studies

Power Calculation

The power calculation was based on an assumption of an absolute 12% improvement in SRI response versus placebo (belimumab 56%, placebo 44%), and this was based on observed responses in the phase III studies of the IV formulation (BLISS-52 and BLISS-76). The calculation also assumed a standard deviation of 50%. This yielded a target sample of 544 patients in the belimumab group and 272 in the placebo group, which provided 90% power with a 5% threshold for statistical significance. There does not appear to have been any accounting for loss to follow-up in the calculation of sample size.

Statistical Test or Model

The primary outcome was analyzed using a logistic regression model, with independent variables treatment group, baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other). The primary analysis was performed on the ITT population. Sensitivity analyses were performed, using last observation carried forward (LOCF) imputation, on the completer population, and another on the per-protocol population. An additional sensitivity analysis was performed using the SLEDAI 2K scoring rule, in which proteinuria is scored as four points any time the value is > 0.5 g in 24 hours.

Assessments of disease activity (SELENA-SLEDAI score, SFI, BILAG score) were performed at baseline and every 28 days until the end of the study. The SDI was performed

at baseline and week 52. Specific details about how each scale is calculated are reviewed in the Outcomes section.

Data Imputation Methods

For the primary outcome and its components, patients who began prohibited drugs or who withdrew from the study were considered treatment failures. For patients who did not have a visit on week 52, if they had a visit within 28 days of day 364, the data from the visit closest to that date was used as the week 52 value. If a patient had two visits within equal distance of day 364 (± 28 days), data from the visit before day 364 was used for the week 52 analysis. If the patient had only partial data for the primary outcome (i.e., they were missing a component of the composite outcome), LOCF was used to impute that missing component. A post hoc tipping-point analysis was also performed as an additional sensitivity analysis, at the request of the FDA. This analysis varies the assumptions about what happened to those patients who withdrew from the study, and these could vary independently between groups. For example, a dropout in the belimumab group could have worse outcomes than a dropout in the placebo group.

Subgroup Analyses

Pre-specified subgroup analyses relevant to the protocol for this CDR review were performed based on baseline SELENA-SLEDAI score (≤ 9 or ≥ 10), C3 or C4 levels (low or not low), baseline C3/C4 and anti-dsDNA (not ≥ 1 C3/C4 low and anti-dsDNA ≥ 30 IU/mL or ≥ 1 C3/C4 low and anti-dsDNA ≥ 30 IU/mL), baseline anti-dsDNA (< 30 IU/mL or ≥ 30 IU/mL), baseline prednisone use (no use or use), baseline medications (steroid, antimalarial drug, and immunosuppressant or other), and baseline mycophenolate mofetil (yes or no). Other pre-specified subgroup analyses were performed based on age (< 65 years or ≥ 65 years old), gender, baseline body weight quartile, body mass index, race, country region, baseline average prednisone dosage, and immunogenicity status. The subgroup analyses were not adjusted for any covariates, and interaction P values were reported and tested for significance at the alpha = 0.010 level. No hypotheses were reported as to whether greater efficacy was anticipated in one subgroup versus another. Multiplicity does not appear to have been accounted for in the analyses.

Sensitivity Analyses

Sensitivity analyses were performed on the various analysis sets: per-protocol, completers, and others. A post hoc tipping-point sensitivity analysis was also performed on the primary and key secondary outcomes at the request of a regulatory agency. Other sensitivity analyses were performed by adjusting the parameters for the composite itself. For example, the SLEDAI 2K scoring rule was used, in which proteinuria is scored as four points when it is > 0.5 g in 24 hours. Higher thresholds were also used for SELENA-SLEDAI in order for patients to be considered a responder (SELENA-SLEDAI ≥ 5 [SRI5], ≥ 6 [SRI6], ≥ 7 [SRI7], and ≥ 8 [SRI8] point reductions).

Secondary Outcomes of the Studies

The first secondary outcome, time to first severe flare according to SFI response over 52 weeks, was analyzed using a Cox proportional hazards model adjusting for baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus low C3 or C4), and race (black versus other). If patients withdrew early or completed the study to week 52, time to first severe flare was censored at the time of last observation in the time period being analyzed. The next secondary outcome, the percentage of patients achieving a reduction in prednisone dosage of 25% or greater from baseline to 7.5 mg per

day or less, was analyzed using the same statistical method as the primary outcome. Those who withdrew early or those receiving prohibited drugs were considered nonresponders, as was the case in the primary analysis.

A hierarchical testing procedure was used to account for multiplicity among outcomes, in which outcomes were tested in the following order: 1) SRI response at week 52, 2) time to first SLE flare, and 3) percentage of patients with average prednisone dosage reduced by 25% from baseline to 7.5 mg daily during weeks 40 to 52. Failure to achieve statistical significance (two-sided alpha of 0.05) resulted in an acknowledgement that subsequent outcomes should not be deemed statistically significant, although P values were still reported for descriptive purposes. Other efficacy end points were not adjusted for multiple comparisons, and it was not mentioned whether the composite was adjusted for multiple comparisons.

Analysis Populations

The ITT population consisted of randomized patients who received at least one dose of study medication, and patients were analyzed according to the group to which they were assigned rather than the actual treatment they received. The as-treated group were all patients who received at least one dose of study drug, and patients were analyzed according to the actual treatment they received > 50% of the time rather than the group to which they were assigned. The per-protocol group were all patients randomized and treated with at least one dose of study medication who had no major protocol deviations. Completers were those who completed all 52 weeks of DB treatment.

Results

Patient Disposition

There were 17% of patients receiving belimumab and 24% of patients receiving placebo who discontinued the study. The most common reason for discontinuation was AE (7% of patients receiving belimumab and 9% of those receiving placebo).

Table 6: Patient Disposition

	BLISS SC	
	Belimumab N = 556	**Placebo N = 280**
Screened, N	1,427	
Randomized, N (%)	559	280
Discontinued study, N (%)	93 (16.7)	66 (23.6)
Adverse event	40 (7.2)	25 (8.9)
Patient request	12 (2.2)	15 (5.4)
Disease progression/lack of efficacy	15 (2.7)	10 (3.6)
Other	14 (2.5)	4 (1.4)
Lost to follow-up	6 (1.1)	2 (0.7)
Protocol violation	4 (0.7)	3 (1.1)
Investigator decision	1 (0.2)	5 (1.8)
Lack of compliance	1 (0.2)	2 (0.7)
ITT, N	556	280
Per-protocol, N	521	268

	BLISS SC	
	Belimumab N = 556	Placebo N = 280
As treated	556	280
Completers	463	214

ITT = intention-to-treat.

Source: Clinical Study Report for BLISS SC.[15]

Exposure to Study Treatments

The mean (SD) duration of exposure was 330 (84.6) days in the belimumab group and 319 (95.4) days in the placebo group. Compliance, defined as the percentage of prescribed doses administered, was 96% in each group. Most patients were on some form of combination therapy for SLE as background, in addition to the study drugs (belimumab or placebo). The most common combination was a steroid and an antimalarial drug (about 35% of patients); a steroid, antimalarial drug, and an immunosuppressant (24%); or a steroid and an immunosuppressant (17%). Only 11% of patients were on a steroid only, and only 7% were on an antimalarial drug only. No information was provided about the extent of exposure to these concomitant therapies.

Efficacy

Only those efficacy outcomes and analyses of subgroups identified in the review protocol are reported here. See Appendix 3 for detailed efficacy data.

Mortality

Three patients in the belimumab group died and two patients in the placebo group over the course of 52 weeks. Deaths in the belimumab group were all infection-related (sepsis, urosepsis, tuberculosis), and deaths in the placebo group were due to thrombocytopenia and cardiac arrest.

Morbidity

Organ damage was assessed using the SDI. The mean difference between belimumab and placebo groups after 52 weeks was 0.0 (95% CI, –0.1 to 0.0), P = 0.1174.

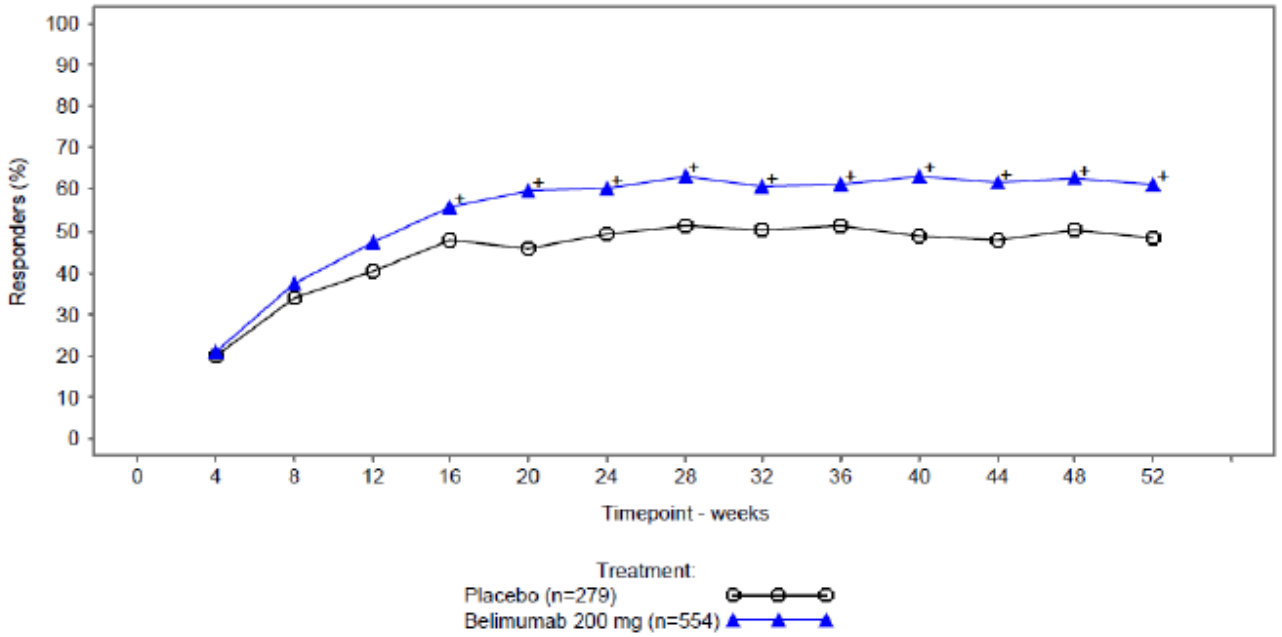
Other indicators of organ damage included proteinuria, and the difference between belimumab and placebo groups in change from baseline to week 52 was –0.13 (95% CI, –0.21 to –0.04).

Disease Activity

More patients in the belimumab group than in the placebo group (61% versus 48%) achieved an SRI response (OR = 1.68 [95% CI, 1.25 to 2.25], P = 0.0006); thus, belimumab met its primary outcome. Sensitivity analyses were performed using LOCF (65% versus 55% with SRI response, belimumab versus placebo, respectively), as well as the completer (73% versus 63%) and per-protocol (62% versus 48%) populations, as well as a post hoc tipping-point analysis requested by the FDA, and results for all were consistent with those of the primary analysis. The components of the SRI were also improved in a greater percentage of patients receiving belimumab than those receiving placebo: a four-point reduction in SELENA-SLEDAI score by week 52 (62% versus 49%, OR = 1.69 [95% CI, 1.26 to 2.27]); no worsening in PGA by week 52 (81% versus 73%, OR = 1.61 [95% CI, 1.15 to 2.27]), and no new 1A/2B BILAG domain scores by week 52 (81% versus 74%,

OR = 1.46 [95% CI, 1.04 to 2.07]). Improvement was observed in the belimumab group as compared with the placebo group as early as week 16, and the difference was sustained up to week 52.

Figure 2: SRI Response by Visit (DB Phase)



DB = double blind; SRI = Systemic Lupus Erythematosus Responder Index.

Source: Clinical Study Report, page 83.[15]

Other exploratory outcomes related to disease activity with respect to durability of response were all in favour of belimumab, including durable SRI response from weeks 40 to 52 (OR = 1.71 [95% CI, 1.27 to 2.30]), time to first SRI response maintained through week 52 (HR = 1.48 [95% CI, 1.21 to 1.81]), and duration of longest SRI response in patients with at least one SRI response (treatment difference of 24.65 days [95% CI, 7.46 to 41.84]).

Subgroup

Subgroup analyses were reported for baseline SELENA-SLEDAI score (≤ 9 or ≥ 10), C3 or C4 levels (low or not low), baseline C3/C4 and anti-dsDNA (not ≥ 1 C3/C4 low and anti-dsDNA ≥ 30 IU/mL or ≥ 1 C3/C4 low and anti-dsDNA ≥ 30 IU/mL), baseline anti-dsDNA (< 30 IU/mL or ≥ 30 IU/mL), baseline prednisone use (no use or use), baseline medications (steroid, antimalarial drug, and immunosuppressant or other), and baseline mycophenolate mofetil (yes or no). Results for these subgroup analyses are summarized in Table 39. Of note, patients who were not on prednisone at baseline exhibited little difference between belimumab (49% of patients with SRI response) and placebo (46% with SRI response). For comparison, those who were using prednisone at baseline had responses of 63% with belimumab and 49% with placebo. Patients treated with mycophenolate mofetil also exhibited little difference in SRI response between belimumab (49%) and placebo (44%), compared with responses of 63% with belimumab and 49% with placebo in those not taking mycophenolate mofetil at baseline.

Health-Related Quality of Life

Health-related quality of life was not investigated in BLISS SC.

Symptoms

The percentage of patients with an improvement in FACIT-Fatigue at week 52 was an exploratory outcome. There were 44.4% of patients receiving belimumab and 36.1% of patients receiving placebo who had at least a four-point improvement in FACIT-Fatigue by week 52 (OR =1.42 [95% CI, 1.05 to 1.94]). FACIT-Fatigue scores increased (improved) from baseline in both studies, and the difference between belimumab and placebo groups at 52 weeks was 1.6 points (95% CI, 0.3 to 2.9).

Other symptoms of importance to patients, according to their input to CDR, such as pain, mouth sores, cognitive impairment, and rash, were not specifically studied in BLISS SC.

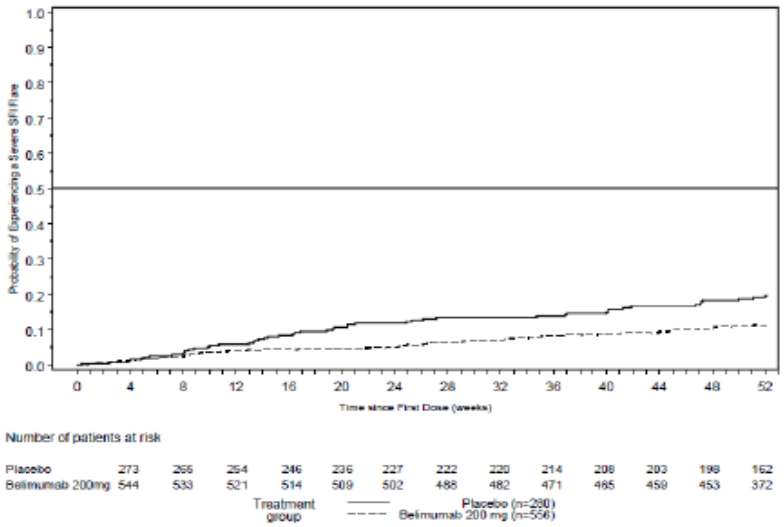
Remission

Remission was not specifically assessed in the included study.

Disease Flare

The risk of experiencing a severe flare was assessed according to the SFI. Fewer patients treated with belimumab (n = 59, 11% of patients) had a severe flare over 52 weeks when compared with patients in the placebo group (n = 51, 18%), and this difference was statistically significant (HR = 0.51 [95% CI, 0.35 to 0.74], P = 0.0004). The adjusted rate ratio (RR) per patient-year was also reduced versus placebo (RR = 0.54 [95% CI, 0.33 to 0.88]), indicating a 46% reduction in the risk of experiencing a severe flare over the course of one year.

Figure 3: Time to First Severe SFI Flare Over 52 Weeks



DB = double blind; SFI = System Lupus Erythematosus Flare Index.
Source: Clinical Study Report, page 96.[15]

The HR with any flare was also reduced in patients on belimumab versus those on placebo (HR = 0.80 [95% CI, 0.65 to 0.99]).

Physician Global Assessment

The percentage of patients for whom PGA had not worsened at week 52 was reported previously as part of the composite SRI response.

Reduction in Corticosteroid Use

A secondary outcome was the reduction of prednisone dosage by 25% or more from baseline to 7.5 mg per day or less during week 40 to 52, and there was no statistically significant difference between belimumab (18%) and placebo (12%) groups for this outcome (OR = 1.65 [95% CI, 0.95 to 2.84], P = 0.0732), although, numerically, a higher percentage of patients with on belimumab had a reduced dosage.

Activities of Daily Living and Caregiver Burden

This outcome was not investigated in the included study.

Table 7: Key Efficacy Outcomes

	BLISS SC	
	Belimumab N = 556	**Placebo N = 280**
Mortality		
Deaths by week 52, n (%)	3 (0.5)	2 (0.7)
Organ damage		
SLICC/ACR Damage Index (SDI) change from baseline at week 52	N = 537	N = 268
Mean (SD) baseline	0.6 (1.00)	0.7 (1.19)
Mean (SD) change to week 52	0.0 (0.19) N = 537	A. (0.22) N = 260
Difference between groups [95% CI] [a]	0.0 [–0.1 to 0.0]	
Patients with doubling of serum creatinine, n (%)	1/462 (0.2)	1/213 (0.5)
Baseline proteinuria, mean (SD)	0.39 (0.707) N = 556	0.44 (0.838) N = 280
Mean (SD) change in proteinuria from baseline to week 52	–0.09 (0.488) N = 456	0.01 (0.851) N = 211
LSM (SE) change in proteinuria from baseline to week 52	–0.07 (0.04)	0.06 (0.05)
Difference between groups [95% CI] [b]	–0.13 [–0.21 to –0.04]	
Disease activity (SRI response; primary outcome)		
Patients with SRI response by week 52, n/N (%)	340/554 (61.4)	135/279 (48.4)
Odds ratio [95% CI] [c] versus placebo	1.68 [1.25 to 2.25] P = 0.0006	
Components of response:		
4-point reduction in SELENA-SLEDAI score, n/N (%)	345/554 (62.3)	137/279 (49.1)
Odds ratio [95% CI] [c] versus placebo	1.69 [1.26 to 2.27]	
No worsening in PGA	450/554 (81.2)	203/279 (72.8)
Odds ratio [95% CI] [c] versus placebo	1.61 [1.15 to 2.27]	
No new 1A/2B BILAG domain scores n/N (%)	448/554 (80.9)	207/279 (74.2)

	BLISS SC	
	Belimumab N = 556	**Placebo N = 280**
Odds ratio [95% CI] [c] versus placebo	1.46 [1.04 to 2.07]	
Reasons for non-response, n (%)		
Dropout without medication failure	55 (9.9)	42 (15.1)
Medication failure	42 (7.6)	27 (9.7)
SELENA-SLEDAI score, < 4-point reduction	112 (20.2)	73 (26.2)
SELENA-SLEDAI score, ≥ 4-point reduction	5 (0.9)	2 (0.7)
PGA worsening only	1 (0.2)	0
BILAG new 1A/2B only	4 (0.7)	1 (0.4)
Both PGA worsening and BILAG new 1A/2B	0	1 (0.4)
Response using SLEDAI 2K n/N (%)	342/554 (61.7)	130/279 (46.6)
Odds ratio [95% CI][c] versus placebo	1.83 [1.36 to 2.46]	
Disease activity (exploratory outcomes)		
Durable SRI response from week 44 to week 52, n (%)	301 (54.3)	114 (40.9)
Odds ratio [95% CI] [c] versus placebo	1.71 [1.27 to 2.30]	
Duration of longest SRI response among patients with ≥ 1 SRI response, LSM (SE), days	193.9 (7.22)	169.3 (8.82)
Treatment differences [95% CI] versus placebo[e]	24.65 [7.46 to 41.84]	
Flares		
Patients with severe flare, week 52 N (%)	59 (10.6)	51 (18.2)
Hazard ratio [95% CI][d] versus placebo	0.51 [0.35 to 0.74] P = 0.0004	
Unadjusted rate per patient-year	0.2	0.3
Adjusted rate per patient-year	0.2	0.4
Adjusted rate ratio [95% CI][e] versus placebo	0.54 [0.33 to 0.88]	
Patients with any flare, week 52, n/N (%)	241/509 (47.3)	136/248 (54.8)
Hazard ratio [95% CI][d] versus placebo	0.80 [0.65 to 0.99]	
Unadjusted rate per patient-year	1.7	2.0
Adjusted rate per patient-year	2.0	2.5
Adjusted rate ratio [95% CI][f] versus placebo	0.81 [0.69 to 0.97]	
Patients with renal flare-week 52, n (%)	26 (4.7)	21 (7.5)
Hazard ratio [95% CI]	0.57 [0.32 to 1.01]	
Prednisone dosage		
Prednisone reduction by ≥ 25% from baseline to ≤ 7.5 mg/day during week 40 through week 52, n/N (%)	61/335 (18.2)	20/168 (11.9)
Odds ratio [95% CI][c] versus placebo	1.65 [0.95 to 2.84] P = 0.0732	
Symptoms		
Mean (SD) baseline FACIT-Fatigue score	31.9 (12.17)	32.1 (11.35)
Mean (SD) change to week 52	5.1 (10.28) N = 554	3.4 (9.22) N = 280
LSM (SE) change to week 52	4.4 (0.55)	2.7 (0.65)
Treatment difference [95% CI] versus placebo	1.6 [0.3 to 2.9]	
Patients with improvement (≥ 4) in FACIT week 52, n/N (%)	246/554 (44.4)	101/280 (36.1)

	BLISS SC	
	Belimumab N = 556	Placebo N = 280
OR [95% CI] versus placebo [g]	1.42 [1.05 to 1.94]	

ACR = American College of Rheumatology; BILAG = British Isles Lupus Assessment Group; CI = confidence interval; FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy–Fatigue Scale; LSM = least squares mean; OR = odds ratio; PGA = Physician Global Assessment; SD = standard deviation; SE = standard error; SELENA = Safety of Estrogen in Lupus Erythematosus National Assessment; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SLICC = Systemic Lupus International Collaborating Clinics; SRI = Systemic Lupus Erythematosus Responder Index.

[a] ANCOVA model comparing belimumab and placebo with covariates for treatment group, baseline SLICC/ACR Damage Index score, baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

[b] All statistics are from an analysis of covariance (ANCOVA) model comparing belimumab and placebo with covariates for treatment group, baseline proteinuria value, baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

[c] Odds ratio (95% confidence interval) and P value are from a logistic regression model for the comparison between belimumab and placebo with covariates treatment group, baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

[d] From Cox proportional hazards model for the comparison between belimumab and placebo, adjusting for baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

[e] Analysis of covariance (ANCOVA) model comparing belimumab and placebo with covariates for treatment group, baseline PGA score, baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

[f] From negative binomial regression with the number of flares/severe flares as the dependent variable and adjusting for baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other). Adjustment is also made for patient's follow-up time by including log follow-up time (years) as an offset variable in the model.

[g] Odds ratio (95% confidence interval) and P value are from a logistic regression model for the comparison between belimumab and placebo with covariates treatment group, baseline FACIT-Fatigue score, baseline SELENA-SLEDAI score (≤ 9 versus ≥ 10), baseline complement levels (low C3 and/or C4 versus no low C3 or C4), and race (black versus other).

Source: Clinical Study Report for BLISS SC.[15]

Harms

Only those harms identified in the review protocol are reported here. See Table 8 for detailed harms data.

Adverse Events

There were 81% of patients receiving belimumab and 84% of patients receiving placebo who had an AE. The most common AE in either group was headache, and there were no clear or consistent differences in the percentage of patients with a specific AE.

Serious Adverse Events

There were 11% of patients receiving belimumab and 16% of patients receiving placebo who experienced an SAE. The only SAE that occurred in at least 1% of patients was thrombocytopenia, and all three patients (1.1%) experiencing this SAE were in the placebo group.

Withdrawals Due to Adverse Events

There were 7% of patients in the belimumab group and 9% of patients receiving placebo who discontinued study drug due to an AE.

Notable Harms

Post-injection reactions occurred in 7% of patients receiving belimumab and 9% of patients receiving placebo. Psychiatric AEs occurred in 6% of belimumab and 11% of patients receiving placebo, infections of special interest occurred in 5% of belimumab and 8% of patients receiving placebo. Tuberculosis occurred in 0.4% of patients in the belimumab

group and 0.7% of patients in the placebo group, and herpes zoster occurred in 3.1% of patients receiving belimumab and 4.6% of patients receiving placebo.

Table 8: Summary of Harms

	BLISS SC	
	Belimumab N = 556	**Placebo N = 280**
Adverse events		
Patients with > 0 AE, N (%)	449 (81)	236 (84)
Most common AE, 5% in any group		
Headache	57 (10)	26 (9)
Viral upper respiratory tract infection	49 (9)	24 (9)
Nasopharyngitis	38 (7)	22 (8)
Nausea	38 (7)	22 (8)
Urinary tract infection, bacterial	42 (8)	18 (6)
Back pain	28 (5)	16 (6)
Upper respiratory tract infection, bacterial	30 (5)	14 (5)
Arthralgia	32 (6)	11 (4)
Diarrhea	28 (5)	14 (5)
Cough	22 (4)	19 (7)
Hypertension	25 (5)	14 (5)
Insomnia	18 (3)	20 (7)
Serious AEs		
Patients with > 0 SAEs, N (%)	60 (11)	44 (16)
Most common SAEs, 1% in any group		
Thrombocytopenia	0	3 (1.1)
Withdrawals due to AEs		
AE resulted in study agent discontinuation, N (%)	40 (7)	25 (9)
Mortality		
Number of deaths, N (%)	3 (0.5)	2 (0.7)
Most common reasons		
Reason	Sepsis	Thrombocytopenia
Reason	Tuberculosis	Cardiac arrest
Reason	Urosepsis	
Notable harms		
Post-injection systemic reactions	38 (7)	25 (9)
Post-injection systemic reactions per anaphylactic reaction, CMQ narrow search	2 (0.4)	1 (0.4)
Post-injection systemic reactions per anaphylactic reaction, CMQ broad search	38 (7)	24 (9)
Serious	1 (0.2)	0 (0)
Post-injection systemic reactions per anaphylactic reaction, CMQ algorithmic search	2 (0.4)	1 (0.4)
Serious delayed non-acute hypersensitivity reactions per manufacturer adjudication	0	1 (0.4)
Psychiatric		
Psychiatric AEs	35 (6)	32 (11)
Suicidal ideation	3 (0.5)	0
Suicidal ideation or behaviour by C-SSRS	7 (1.3)	2 (0.7)

	BLISS SC	
	Belimumab N = 556	**Placebo N = 280**
Suicidal ideation	7 (1.3)	2 (0.7)
Suicidal behaviour	0	0
Infection		
All infections of special interest	30 (5)	21 (8)
Serious	8 (1.4)	3 (1.1)
Opportunistic per manufacturer adjudication	2 (0.4)	1 (0.4)
Serious	1 (0.2)	0
Opportunistic infections, manufacturer adjudication excluding tuberculosis or herpes zoster	0	1 (0.4)
Serious	0	0
Active TB	2 (0.4)	2 (0.7)
Serious	1 (0.2)	1 (0.4)
Nonopportunistic	1 (0.2)	2 (0.7)
Serious	0	1 (0.4)
Opportunistic	1 (0.2)	0
Serious	1 (0.2)	0
Herpes zoster	18 (3.2)	13 (4.6)
Serious	1 (0.2)	0
Nonopportunistic	17 (3.1)	13 (4.6)
Serious	1 (0.2)	0
Opportunistic	1 (0.2)	0
Disseminated	1 (0.2)	0
Sepsis	6 (1.1)	3 (1.1)
Serious	4 (0.7)	2 (0.7)

AE = adverse event; CMQ = customized Medical Dictionary for Regulatory Activities query; C-SSRS = Columbia–Suicide Severity Rating Scale; SAE = serious adverse event; TB = tuberculosis.

Source: Clinical Study Report for BLISS SC.[2]

Critical Appraisal

Internal Validity

There was both a relatively high percentage of patients who withdrew from the study in each of the belimumab (17%) and placebo (24%) groups, and a higher percentage of patients receiving placebo who withdrew from the study compared to patients receiving belimumab. The high discontinuations, including a large proportion of patients who had prematurely discontinued the treatment, would have led to the ITT analysis at 52 weeks. The direction of the bias is difficult to ascertain. The causes of discontinuation were primarily AE, patient request, and disease progression/lack of efficacy. There was a slightly higher proportion of patients who discontinued due to disease progression/lack of efficacy in the placebo group (3.6%) than in the belimumab (2.7%), and there were also more patients in the placebo group who requested to withdraw (5.4% versus 2.2%) before the end of the study. The sponsor used a nonresponder (e.g., treatment failure) imputation approach for any early withdrawals or for patients who began taking a prohibited drug. In this approach, when more patients withdraw from the placebo group, this may have biased the results in favour of belimumab, as these patients would be considered nonresponders, whether they were responding at the time of withdrawal or not. The impact of this potential bias would be mitigated somewhat if the difference in withdrawals could be largely

accounted for by lack of efficacy; however, this was not the case, as there was only a 1% difference in withdrawals between groups for this reason. The sensitivity analyses performed by the sponsor seem to support the findings of their primary analysis, and these include LOCF approaches as well as completer analyses and per-protocol analyses. LOCF was also used to impute missing data when individual components of the primary composite outcome were missing. The FDA believed these analyses were inadequate and thus requested a post hoc tipping-point analysis, in which potential responses were varied independently between belimumab and placebo groups. The results of this analysis also supported that of the primary outcome.[14]

A hierarchical testing procedure was used to account for multiplicity, and included the primary and two secondary outcomes; however, no adjustments were made for the subgroup analyses that were performed. Tests for interactions between subgroups were performed; however, no hypotheses were provided for the expected results from these tests. Therefore, they can be considered only hypothesis-generating. The sponsor also continued to report P values for outcomes that fell outside of their hierarchy and were deemed exploratory, and these P values were not reported by CDR.

The sponsor otherwise adhered to its hierarchy, testing outcomes in sequence. It was only the final outcome in the hierarchy, reduction in prednisone dosage by 25% to 7.5 mg daily or less, where statistical significance was not reached. For this outcome, only those receiving prednisone at a baseline dosage of more than 7.5 mg daily were included in the analysis, and this was only about 60% of the ITT population.

The primary outcome was a composite, described as an SRI response that included three different assessments related to disease activity: the SELENA-SLEDAI score, BILAG score, and PGA score. CDR reviewed these outcomes (Appendix 4), and the thresholds used to indicate a clinically significant response on the SELENA-SLEDAI (minimum four points) and the PGA (no worsening indicated by change of < 0.3 points) have been identified by the sponsor. According to clinical experts consulted for this review, the SRI is a clinically relevant outcome to assess response in patients with SLE. Responders and nonresponders on the SRI have been shown to differ on several measures of disease activity, biomarkers, and health-related quality of life. However, there is some uncertainty as to the clinical meaningfulness of the thresholds of 0.3 for PGA score and four points for SELENA-SLEDAI score. For example, one study showed the MCID of improvement on SELENA-SLEDAI was seven points (Appendix 4). Of note, a clinically meaningful difference was estimated based on the change in disease-activity score that corresponded to a 0.70 or more probability that experts would rate the patient as improved or worsened. Some issues have been identified with the SLEDAI. In comparison with BILAG, the SLEDAI is less responsive to change; it does not capture improvement or worsening; and it does not assess severity in an organ system. As discussed in the Appendix, however, using a single weighted score to summarize disease activity, on the one hand, makes the judgment of disease activities much easier and standardized, while, on the other, it could mask the underlying importance of organ systems that are contributing to the total score (i.e., the same score could represent mild disease in many organs or severe disease in a single organ; or a score could remain unchanged despite worsening in one organ system if there is also improvement in another system).

External Validity

BLISS SC excluded patients with severe renal or CNS involvement. This may have resulted in a relatively homogeneous population with less severe disease, who may therefore have been more likely to respond to belimumab therapy and to have a more favourable benefit:risk profile. Additionally, 1997 ACR criteria were used to identify patients with SLE in BLISS SC, and these are rigorous criteria that are designed for use in clinical trials, rather than clinical practice. Thus, there is a higher risk of misdiagnosis of SLE in clinical practice, although the clinical experts consulted by CDR on this review note that diagnosis of SLE should be straightforward for clinicians with specialized training.

According to the clinical experts, a 52-week DB treatment phase is unlikely to be of sufficient duration to detect improvements in organ damage or other longer-term outcomes of treatment with belimumab. The composite primary outcome, patients with an SRI response, would not routinely be used to assess patient status in clinical practice. However, the components of the composite would be an important part of the assessment of patients with SLE.

The included study did not assess health-related quality of life in its DB phase. SLE has a significant impact on quality of life, and this is clear from patients' input to CDR. The sponsor did not provide a rationale for not assessing quality of life, but it did include the SF-36 as an outcome in the open-label extension. However, without a comparator, limited conclusions can be drawn from these data. The lack of data on health-related quality of life is a significant limitation of this review. Additionally, fatigue, a key symptom identified by patients in their input to CDR, was assessed only as an exploratory outcome using the FACIT-Fatigue, thus limiting conclusions that can be drawn about fatigue.

Belimumab has not been studied against an active comparator; therefore, the efficacy and harms of this drug compared to the addition of other drugs used in the treatment of SLE is unknown. There are a variety of drugs used chronically to manage SLE, none of which were specifically developed for managing this disease. ITCs are available; however, these suffer from their own limitations, outlined in the Indirect Evidence section of this report.

Indirect Evidence

Objectives and Methods for the Summary of Indirect Evidence

Given the absence of head-to-head studies comparing the SC formulation of belimumab with the IV formulation of belimumab or other active therapies for reducing disease activity in adult patients with active, autoantibody-positive SLE, ITCs may provide information on the effectiveness and safety of belimumab SC compared with existing therapies. The objective of this section is to summarize and critically appraise available indirect evidence comparing belimumab SC with relevant treatment regimens (as specified in the CDR review protocol) for adult patients with active, autoantibody-positive SLE.

The sponsor submitted one ITC,[18,19] which was reviewed, summarized, and critically appraised. CADTH CDR conducted an independent literature search for published ITCs that compared belimumab SC with other relevant comparators for reducing disease activity in adult patients with active, autoantibody-positive SLE. MEDLINE, Embase, and PubMed were searched. The search was limited to documents published between January 1, 2011,

and July 19, 2019. Two relevant publications were identified in the literature and were included in this summary.[3,4]

Description of Indirect Comparison(s)

The sponsor submitted an ITC that compared the efficacy of the IV formulation with the SC formulation of belimumab in adult patients with active, autoantibody-positive SLE and I who are receiving standard therapy.[18,19]

One ITC by Lee and Song[3] identified in the literature compared the efficacy and safety of the IV formulation with the SC formulation of belimumab in combination with standard therapy in patients with active SLE.

Another ITC by Tian et al.[4] identified in the literature compared specific adverse effects of immunosuppressive drugs, biologics, and GCs in patients with SLE.

The population, intervention, comparators, outcomes, and design of studies included in the ITCs are provided in Table 9.

Table 9: Study Selection Criteria and Methods for ITCs

	Ramachandran et al.[18,19]	Lee and Song[3]	Tian et al.[4]
Population	Adult patients with diagnosis of SLE according to ACR criteria, and active SLE defined by SELENA-SLEDAI score at screening of ≥ 6. In addition, patients had to have positive ANA test (titre ≥ 1:80) or anti-dsDNA (≥ 30 IU/mL) from two independent time points, 1 within the screening period, and stable SLE treatment regimen for a period of 30 days or more (depending on agent/study) before day 0 Analysis was conducted on patients with I. The I subgroups were defined in two ways: 1) Criteria I — Patients who meet a and b criteria at baseline: 2) Criteria II — Patients who meet a or c criteria at baseline: a) C3 (< 0.9 g/L) or C4 (< 0.16 g/L): For BLISS 76 and BLISS 52 C3 (< 0.9 g/L) or C4 (< 0.10 g/L): For BLISS SC and BEL113750 b) Anti-dsDNA positive (≥ 30 IU/mL) c) SELENA-SLEDAI score ≥ 10	Patients with active SLE (score ≥ 4 at screening on SELENA-SLEDAI) despite having received standard therapy	Patients who met the 1987 American College of Rheumatology Classification criteria for SLE
Intervention	Belimumab 200 mg SC	Belimumab 200 mg SC	

	Ramachandran et al.[18,19]	Lee and Song[3]	Tian et al.[4]
Comparator	Belimumab IV 10 mg/kg	Belimumab IV 10 mg/kg Belimumab IV 1 mg/kg	Agents used to treat SLE, exposure of interest was treatment with immunosuppressants or biologics (only included rituximab and belimumab)
Outcome	• SRI response (a composite end point of ≥ 4-point reduction in SELENA-SLEDAI score; no worsening in PGA, and no new 1A/2B BILAG domain scores at week 52) • ≥ 4-point reduction in SELENA-SLEDAI at week 52 Occurrence of severe flare occurrence measured by the SLE Flare Index (SFI) over 52 weeks	• SRI response rate at week 52 (defined as a >4-point reduction in SELENA-SLEDAI score, no new BILAG A organ domain scores, and no more than 1 new BILAG B score, and no worsening in PGA score versus baseline) SAEs until week 52	• All-cause mortality • AE-related withdrawals • AEs • SAEs • Cardiovascular events (acute coronary syndrome, chronic ischemic heart disease, coronary revascularization, cardiovascular disease death, cerebrovascular events, or peripheral vascular events) • Serious infections (serious infection, major infection, severe infection, sepsis, cardiovascular infection, or bacterial pneumonia) • Bone toxicity (avascular necrosis or fracture) • Malignant transformation • Serious gastrointestinal events (leading to dose reduction or withdrawal) • Ovarian failure (sustained amenorrhea) • Menstrual disorder • New-onset hypertension • Serious leucopenia (white cell count < 2 × 109 L leading to dose reduction or withdrawal) • Leucopenia • Hyperglycemia (hyperglycemia or new-onset diabetes)
Study design	Pivotal phase III RCTs	RCTs	RCTs
Exclusion criteria	Studies were excluded in they enrolled patients with any of the following criteria: • B-cell targeted therapy, any prior • Investigational biologic (including abatacept) within past year • IV cyclophosphamide within past 180 days • Anti-TNF, anakinra, IVIG, high-dose prednisone, plasmapheresis within past 90 days	• Inclusion of duplicate data, lack of adequate data for inclusion	• Duplicate reports • Studies that did not report on the outcomes of interest or in which all arms had 0 events • Studies that lasted 24 weeks or less • Studies that included children younger than 10 years old or women during pregnancy or lactation

	Ramachandran et al.[18,19]	Lee and Song[3]	Tian et al.[4]
	• New immunosuppressive, antimalarial drug, NSAID within past 60 days • Severe lupus kidney disease/active nephritis within past 90 days • CNS lupus requiring intervention within past 60 days • History of renal transplant • Grade 3 or 4 laboratory abnormality (some lupus-related exceptions permitted) History of anaphylactic reaction to parenterally administered contrast agents/proteins		• Studies that included fewer than 20 patients • Scientific reports that presented pooled trial data for which the individual trials could not be identified to prevent double counting
Databases searched	NR	MEDLINE, Embase, and the Cochrane Central Register of Controlled Trials	PubMed, the Cochrane Central Register of Controlled Trials, and Embase (from their inception to September 2017)
Selection process	NR	NR	Two reviewers independently assessed the full text of the articles. Any disagreement was resolved by consensus discussion.
Data extraction process	NR	NR	Standardized data forms and data extraction training exercises were developed.
Quality assessment	NR	Jadad scores	The Cochrane Collaboration's tool for assessing risk of bias

ACR = American College of Rheumatology; AE = adverse event; ANA = antinuclear antibody; anti-dsDNA = anti–double-stranded DNA; BILAG = British Isles Lupus Assessment Group; CNS = central nervous system; I = high disease activity; ITC = indirect treatment comparison; IVIG = intravenous immunoglobulin; NR = not reported; NSAID = nonsteroidal anti-inflammatory drug; PGA = Physician Global Assessment; RCT = randomized controlled trial; SAE = serious adverse event; SC = subcutaneous; SELENA-SLEDAI = Safety of Estrogens in Lupus Erythematosus National Assessment Trial–Systemic Lupus Erythematosus Disease Activity Index; SLE = systemic lupus erythematosus; SRI = SLE Responder Index; TNF = tumour necrosis factor.

Source : Ramachandran et al.,[18,19] Lee and Song,[3] Tian et al.[4]

Methods of the Sponsor-Submitted ITC

Objectives

The aim of this ITC was to evaluate the comparative clinical efficacy of the IV 10 mg/kg and SC 200 mg per week formulations of belimumab for the treatment of adult patients with active, autoantibody-positive SLE who are receiving standard therapy, via a network meta-analysis (NMA).

Study Selection Methods

No literature search was conducted. It is not reported how studies were selected. Studies included in the ITC were published, pivotal, phase III RCTs of belimumab in patients with active, autoantibody-positive SLE. It is not reported how data extraction was conducted and whether more than one reviewer was involved in data extraction. No quality assessment of included studies was reported.

Studies that met the following inclusion/exclusion criterion were included in the analysis:

- Inclusion criteria: Patients had to be at least 18 years of age at screening visit, with diagnosis of SLE according to ACR criteria, and active SLE defined by SELENA-SLEDAI score at screening of 6 or higher. In addition, patients had to have positive ANA test (titre ≥ 1:80) or anti-dsDNA (≥ 30 IU/mL) from two independent time points, 1 of which was within the screening period, as well as a stable SLE treatment regimen for a period of 30 days or more (depending on agent/study) before day 0.
- Exclusion criteria: Studies were excluded if they enrolled patients with B-cell targeted therapy (any prior), used an investigational biologic (including abatacept) within past year, used IV cyclophosphamide within past 180 days, used anti-TNF, anakinra, IV immunoglobulin, high-dose prednisone, or plasmapheresis within the past 90 days, or used a new immunosuppressant, antimalarial drug, or NSAID within the past 60 days. Studies were also excluded in enrolled patients who had had severe lupus kidney disease or active nephritis within the past 90 days, CNS lupus requiring intervention within the past 60 days, history of renal transplant, Grade 3 or 4 laboratory abnormality (some lupus-related exceptions permitted), or history of anaphylactic reaction to parenterally administered contrast agents/proteins.

The included studies were BLISS 52,[20] BLISS 76,[21] BEL113750,[22] and BLISS SC.[10] Patient-level data were obtained from these studies for patients with I. The I subgroups were defined in two ways: criteria I and criteria II. Criteria I were defined as patients who were anti-dsDNA positive (≥ 30 IU/mL) and whose complement levels were C3 (< 0.9 g/L) or C4 (< 0.16 g/L), for patients in BLISS 52 and BLISS 76 trials, and C3 (< 0.9 g/L) or C4 (< 0.10 g/L), for patients in BLISS SC and BEL113750 trials. Criteria II were defined as patients with SELENA-SLEDAI score ≥ 10, or patients whose complement levels were as defined for criteria I.

The primary end point assessed was SRI response at week 52, defined as a composite end point of no new 1A/2B BILAG domain scores, reduction in SELENA-SLEDAI score of four points or more, and no worsening in PGA. Other end points were the percentage of patients with reduction in SELENA-SLEDAI score of four points or more at week 52 and the occurrence of severe flare occurrences over 52 weeks, as measured by the SFI. Safety end points were not assessed.

ITC Analysis Methods

A Bayesian NMA was conducted for the outcomes SRI response at 52 weeks, percentage of patients with a reduction in SELENA-SLEDAI score of four points or more at week 52, and the rate of flare occurrences at week 52. Treatment effects for all outcomes were analyzed using a binomial likelihood and logit link function. These analyses were conducted on both subgroups of patients identified, i.e., the subgroup of patients who met criteria I and the subgroup of patients who met criteria II. The treatment effects were presented as odd ratios with 95% credible intervals (CrIs). It was concluded that one treatment was favoured over another if the CrI excluded the null.

Analyzes were conducted using both a fixed-effects model and a random-effects model. The final model selection was based on residual deviance and deviance information criterion. Noninformative prior distributions were used for all model parameters. For the random-effects model, the prior for the heterogeneity parameter used was uniform (0, 2). The fixed-effects model was selected by the sponsor for the analysis of all three end points; the rationale provided was that results from the fixed-effects model were easier to interpret

and more conservative. Also, the small number of studies and the homogeneous nature of the characteristics of the studies included in the ITC made the fixed-effects model more appropriate within this analysis. In addition, the residual deviance within each analysis showed no difference between the two models.

The first 30,000 iterations were discarded as "burn-in," results were based on an additional 50,000 iterations using three chains, and convergence was assessed using the Brooks–Gelman–Rubin convergence diagnostics. Given that there were no closed loops, the consistency assumption was not checked.

Analysis was conducted using the Bucher method as a sensitivity analysis. Another sensitivity analysis was conducted by excluding study BEL113750,[22] which was different from the other three studies on region, race, and SFI.

Table 10 presents a summary of the methods used for the ITC.

Table 10: ITC Analysis Methods

	Ramachandran et al.
ITC methods	Network meta-analysis
Priors	Uniform (0, 2)
Assessment of model fit	The model selected was chosen based on residual deviance and deviance information criterion.
Assessment of consistency	Not conducted, given that there were no closed loops.
Assessment of convergence	Brooks–Gelman–Rubin convergence diagnostics
Follow-up time points	52 weeks
Sensitivity analyses	• Reanalysis of the data using the Bucher method • ITC sensitivity analysis of all studies excluding the northeast Asian study
Subgroup analysis	• Criteria I • Criteria II

ITC = indirect treatment comparison.

Source: Ramachandran et al.[18]

Results of the Sponsor-Submitted ITC

Summary of Included Studies

Four phase III RCTs (BLISS 52,[20] BLISS 76,[21] BEL113750,[22] and BLISS SC[10]) were included in the ITC. Characteristics of these studies are presented in Table 11.

Table 12 presents demographic and baseline disease characteristics for the subgroup of patients with I who met criteria I. In total, 1,375 patients with I who met the criteria I subgroup were included in the ITC (IV, n = 597; SC, n = 248; placebo, n = 530). Most patients were female and the sex ratio was similar across studies, but the mean weight of the study population varied across trials, ranging from 56.2 kg (in BEL113750) to 70.6 kg (in BLISS 76). The BEL113750 study had no distribution of race or region category, other than Asian, compared to the other studies. In the BEL113750 study, the average age of patients included was slightly younger than that in the other studies, and the SFI showed a slightly different profile compared to the other studies. The percentage of patients with at least one flare on SFI was lower in the BEL113750 study than the percentage reported in the other three studies; however, the percentage of patients with at least one severe flare on SFI was

higher in the BEL113750 study than the percentage reported in the other three studies. The SELENA-SLEDAI score and subcategory were homogeneous across studies. The PGA scale for the four studies was also comparable, but the percentage of patients within the 0 to 1 PGA category for BEL113750 (6.1%) and BLISS SC (7.6%) was less than that for BLISS 52 (15.0%), and BLISS 76 (14.3%).

Table 13 presents demographic and baseline disease characteristics for the subgroup of patients with I who met criteria II. In total, 1,892 patients with I who met the criteria II subgroup were included in the ITC (IV, n = 739; SC, n = 421; placebo, n = 732). For patients who met criteria II, the demographic and baseline disease characteristics were similar to those for patients who met criteria I.

Table 11: Study Design of Included Studies in the Sponsor-Submitted ITC

Study name	Geographical location	Design	Intervention	Comparator	Follow-up	Outcomes reported
BLISS SC[10]	US, Canada, South America, Western Europe, Australia, Israel, Eastern Europe, Asia	Phase III, multi-centre, double-blind 836 patients randomized (2:1)	• Belimumab 200 mg + SoC n = 556	Placebo plus SoC n = 280	• 52 weeks • SC in every week for 52 weeks	• SRI response • 4-point reduction in SELENA-SLEDAI • No worsening in PGA • No new 1A 2B BILAG domain scores • Reduction in prednisone
BLISS 76[21]	US, Canada, South America, Western Europe, Australia, Israel, Eastern Europe	Phase III, multi-centre, double-blind 819 patients randomized (1:1:1)	• Belimumab 1 mg/kg + SoC n = 271 • Belimumab 10 mg/kg + SoC n = 273	Placebo plus SoC n = 275	• 76 weeks • IV infusion in 1 h on days 0, 14, and 28, and then every 28 days for 72 weeks	• SRI response • 4-point reduction in SELENA-SLEDAI • No worsening in PGA • Reduction in prednisone • SF-36 PCS score change
BLISS 52[20]	South America, Western Europe, Australia, Israel, Eastern Europe, Asia	Phase III, multi-centre, double-blind 867 patients randomized (1:1:1)	• Belimumab 1 mg/kg + SoC n = 288 • Belimumab 10 mg/kg + SoC n = 290	Placebo plus SoC n = 287	• 52 weeks • IV infusion in 1 h on days 0, 14, and 28, and then every 28 days for 52 weeks.	• SRI response • 4-point reduction in SELENA-SLEDAI • No worsening in PGA • Reduction in prednisone • SF-36 PCS score change
BEL113750[22]	Northeast Asia (China, Japan, and South Korea)	Phase III, multi-centre, double-blind 677 patients randomized (2:1)	• Belimumab 10 mg/kg + SoC n = 451	Placebo plus SoC n = 226	• 52 weeks • The total duration of patient participation in the blinded period,	• SRI response • 4-point reduction in SELENA-SLEDAI • No worsening in PGA

Study name	Geographical location	Design	Intervention	Comparator	Follow-up	Outcomes reported
					including follow-up, was approximately 69 weeks, or 57 weeks if entering the open-label	• Reduction in prednisone • Adverse events

BILAG = British Isles Lupus Assessment Group; ITC = indirect treatment comparison; PCS = physical component summary; PGA = Physician Global Assessment; SC = subcutaneous; SELENA-SLEDAI = Safety of Estrogens in Lupus Erythematosus National Assessment Trial–Systemic Lupus Erythematosus Disease Activity Index; SF-36 = Short Form (36) Health Survey; SoC = standard of care; SRI = Systemic Lupus Erythematosus Responder Index.

Source: Ramachandran et al.[18]

Table 12: Demographic and Baseline Disease Characteristics for Included Studies on High Disease Activity Subgroup (Criteria I)

	BLISS 76		BLISS 52		BEL113750		BLISS SC	
	Belimumab 10 mg/kg	Placebo	Belimumab 10 mg/kg	Placebo	Belimumab 10 mg/kg	Placebo	Belimumab 200 mg	Placebo
Original RCT, N	273	275	290	287	451	226	556	280
High disease activity (criteria I), n	134	131	171	156	292	135	248	108
Female, n (%)	125 (93.3)	120 (91.6)	167 (97.7)	144 (92.3)	270 (92.5)	127 (94.1)	236 (95.2)	106 (98.1)
Race, n (%)								
White	90 (67.2)	84 (64.1)	35 (20.5)	39 (25.0)	0	0	140 (56.5)	58 (53.7)
African-American	17 (12.7)	18 (13.7)	6 (3.5)	2 (1.3)	0	0	26(10.5)	7 (6.5)
Asian	7 (5.2)	8 (6.1)	84 (49.1)	69 (44.2)	292 (100)	135 (100)	64 (25.8)	32 (29.6)
Native race[a]	20 (14.9)	21 (16.0)	46 (26.9)	46 (29.5)	0	0	18 (7.3)	11 (10.2)
Region, n (%)								
North America	48 (35.8)	60 (45.8)	0	0	0	0	55 (22.2)	24 (22.2)
Europe/Australia	48 (35.8)	36 (27.5)	17 (9.9)	15 (9.6)	0	0	82 (33.1)	30 (27.8)
Latin America	18 (13.4)	17 (13.0)	71 (41.5)	73 (46.8)	0	0	52 (21.0)	21 (19.4)
Asian	20 (14.9)	18 (13.7)	83 (48.5)	68 (43.6)	292 (100)	135 (100)	59 (23.8)	33 (30.6)
Mean age, years (SD)	36.6 (10.5)	35.5 (9.9)	33.4 (10.4)	33.6 (11.3)	31.5 (9.5)	30.8 (8.3)	34.6 (11.0)	34.6 (10.4)
Age group (years), n (%)								
≤ 45	104 (77.6)	109 (83.2)	147 (86.0)	131 (84.0)	266 (91.1)	126 (93.3)	207 (83.8)	92 (85.2)
> 45 to < 65	29 (21.6)	21 (16.0)	23 (13.5)	23 (14.7)	25 (8.6)	9 (6.7)	40(16.2)	14 (13.0)

	BLISS 76		BLISS 52		BEL113750		BLISS SC	
	Belimumab 10 mg/kg	Placebo	Belimumab 10 mg/kg	Placebo	Belimumab 10 mg/kg	Placebo	Belimumab 200 mg	Placebo
≥ 65	1 (0.7)	1 (0.8)	1 (0.6)	2 (1.3)				
Mean weight, kg (SD)	70.6 (18.9)	68.8 (14.4)	60.2 (13.3)	61.2 (13.2)	56.9 (10.0)	56.2 (10.3)	64.5 (16.3)	64.8 (17.9)
BILAG organ domain involvement, n (%)								
At least 1A or 1B	124 (92.5)	122 (93.1)	151 (88.3)	136 (87.2)	235 (80.5)	235 (80.5)	221 (89.1)	98 (90.7)
At least 1A	0	0	0	0	23 (7.9)	16 (11.9)	35 (14.1)	25 (23.1)
At least 1A or 2B	75 (56.0)	90 (68.7)	100 (58.5)	76 (48.7)	129 (44.2)	66 (48.9)	164 (66.1)	77 (71.3)
Mean SELENA-SLEDAI score (SD)	10.3 (3.4)	11.4 (4.1)	10.8 (4.0)	10.8 (3.7)	10.4 (3.8)	11.3 (4.0)	11.5 (3.3)	11.7 (3.1)
SELENA-SLEDAI score category								
4 to 9	55 (41.0)	35 (26.7)	59 (34.5)	50 (32.1)	122 (41.8)	46 (34.1)	62 (25.0)	30 (27.8)
10 to 11	37 (27.6)	35 (26.7)	45 (26.3)	41 (26.3)	64 (21.9)	30 (22.2)	57 (23.0)	13 (12.0)
≥ 12	42 (31.3)	61 (46.6)	67 (39.2)	65 (41.7)	106 (36.3)	59 (43.7)	129 (52.0)	65 (60.2)
SFI, n (%)								
At least 1 SFI flare in 35 days before baseline	29 (21.6)	45 (34.4)	27 (15.8)	35 (22.4)	27 (9.2)	14 (10.4)	43 (17.3)	23 (21.3)
At least severe SFI flare in 35 days before baseline	5 (3.7)	3 (2.3)	3 (1.8)	1 (0.6)	15 (5.1)	7 (5.2)	5 (2.0)	1 (0.9)
PGA score category, n (%)								
0 to 1	26 (19.4)	12 (9.2)	19 (11.1)	30 (19.2)	20 (6.9)	6 (4.4)	17 (6.9)	10 (9.3)
> 1 to 2.5	105 (78.4)	117 (89.3)	150 (87.7)	125 (80.1)	262 (90.0)	124 (91.9)	226 (91.9)	96 (88.9)
> 2.5 to 3	3 (2.2)	2 (1.5)	2 (1.2)	1 (0.6)	9 (3.1)	5 (3.7)	3 (1.2)	2 (1.9)
Mean PGA (SD)	1.4 (0.6)	1.5 (0.5)	1.4 (0.5)	1.4 (0.5)	1.6 (0.5)	1.7 (0.5)	1.6 (0.4)	1.6 (0.5)

BILAG = British Isles Lupus Assessment Group; PGA = Physician Global Assessment; RCT = randomized controlled trial; SD = standard deviation; SFI = System Lupus Erythematosus Flare Index;. SELENA-SLEDAI = Safety of Estrogen in Lupus Erythematosus National Assessment–SLE Disease Activity Index.

[a] Native Alaskan, Native Hawaiian, or Native Indian.

Source: Ramachandran et al.[18]

Table 13: Demographic and Baseline Disease Characteristics for Included Studies on High Disease Activity Subgroup (Criteria II)

	BEL110751-IV		BEL110752-IV		BEL113750-IV		BEL112341-SC	
	Belimumab 10 mg/kg	Placebo	Belimumab 10 mg/kg	Placebo	Belimumab 10 mg/kg	Placebo	Belimumab 200 mg	Placebo
Original RCT	273	275	290	287	451	226	556	280
High disease activity (criteria II)	161	162	201	185	377	182	421	203
Female, n (%)	151 (93.8)	150 (92.6)	196 (97.5)	173 (93.5)	348 (92.3)	170 (93.4)	398 (94.5)	197 (97.0)
Race, n (%)								
White	111 (68.9)	108 (66.7)	44 (21.9)	48 (25.9)	0	0	241 (57.2)	116 (57.1)
African-American	19 (11.8)	20 (12.3)	9 (4.5)	2 (1.1)	0	0	46 (10.9)	20 (9.9)
Asian	8 (5.0)	9 (5.6)	92 (45.8)	78 (42.2)	377 (100)	182 (100)	97 (23.0)	51 (25.1)
Native race[a]	23 (14.3)	25 (15.4)	56 (27.9)	57 (30.8)	0	0	37 (8.8)	16 (7.9)
Region, n (%)								
North America	62 (38.5)	76 (46.9)	0	0	0	0	109 (25.9)	58 (28.6)
Europe/Australia	56 (34.8)	42 (25.9)	22 (10.9)	22 (11.9)	0	0	126 (29.9)	56 (27.6)
Latin America	20 (12.4)	23 (14.2)	88 (43.8)	87 (47.0)	0	0	97 (23.0)	39 (19.2)
Asian	23 (14.3)	21 (13.0)	91 (45.3)	76 (41.1)	377(100)	182(100)	89 (21.1)	50 (24.6)
Mean age, years (SD)	37.4 (10.6)	36.5 (10.2)	33.9 (10.4)	34.6 (11.4)	32.0 (9.6)	31.3 (8.9)	36.5 (11.7)	38.0 (11.9)
Age group (years), n (%)								
≤ 45	120 (74.5)	130 (80.2)	171 (85.1)	151 (81.6)	340 (90.2)	168 (92.3)	326 (77.6)	151 (74.4)
> 45 to < 65	40 (24.8)	31 (19.1)	29 (14.4)	31 (16.8)	36 (9.5)	14 (7.7)	90 (21.4)	47 (23.2)
≥ 65	1 (0.6)	1 (0.6)	1 (0.5)	3 (1.6)	NR	0 NR	4 (1.0)	5 (2.5)
Mean weight, kg (SD)	69.5 (18.3)	68.2 (14.3)	60.3 (12.9)	61.2 (12.8)	57.3 (10.4)	57.1 (11.5)	66.9 (17.7)	67.4 (19.1)
BILAG organ domain involvement, n (%)								
At least 1A or 1B	150 (93.2)	153 (94.4)	181 (90.0)	163 (88.1)	307 (81.4)	145 (79.7)	383 (91.0)	188 (92.6)
At least 1A	0	0	0	0	37 (9.8)	21 (11.5)	74 (17.6)	42 (20.7)
At least 1A or 2B	88 (54.7)	110 (67.9)	120 (59.7)	95 (51.4)	180 (47.7)	93 (51.1)	303 (72.0)	162 (79.8)
Mean SELENA-SLEDAI score (SD)	10.0 (3.5)	10.9 (4.1)	10.6 (3.9)	10.4 (3.7)	10.5 (3.7)	11.1 (3.9)	11.4 (3.0)	11.4 (2.8)

	BEL110751-IV		BEL110752-IV		BEL113750-IV		BEL112341-SC	
	Belimumab 10 mg/kg	Placebo	Belimumab 10 mg/kg	Placebo	Belimumab 10 mg/kg	Placebo	Belimumab 200 mg	Placebo
SELENA-SLEDAI score category								
0 to 3			74 (36.8)	1 (0.5)	144 (38.2)	3 (1.6)	0	0
4 to 9	1 (0.6)	55 (34.0)	53 (26.4)	68 (36.8)	93 (24.7)	55 (30.2)	69 (16.4)	35 (17.2)
10 to 11	71 (44.1)	42 (25.9)	74 (36.8)	48 (25.9)	140 (37.1)	46 (25.3)	161 (38.2)	74 (36.5)
≥ 12	41 (25.5)	65 (40.1)	0	NR	0	NR	191 (45.4)	94 (46.3)
SFI, n (%)								
At least 1 SFI flare in 35 days before baseline	36 (22.4)	54 (33.3)	34 (16.9)	37 (20.0)	34 (9.0)	18 (9.9)	74 (17.6)	39 (19.2)
At least severe SFI flare in 35 days before baseline	5 (3.1)	3 (1.9)	3 (1.5)	1 (0.5)	15 (4.0)	10 (5.5)	8 (1.9)	4 (2.0)
PGA score category, n (%)								
0 to 1	29 (18.0)	13 (8.0)	22 (10.9)	34 (18.4)	23 (6.1)	6 (3.3)	23 (5.5)	13 (6.4)
> 1 to 2.5	129 (80.1)	147 (90.7)	177 (88.1)	150 (81.1)	342 (91.0)	168 (92.8)	389 (92.8)	185 (91.1)
> 2.5 to 3	3 (1.9)	2 (1.2)	2 (1.0)	1 (0.5)	11 (2.9)	7 (3.9)	7 (1.7)	5 (2.5)
Mean PGA (SD)	1.4 (0.5)	1.5 (0.5)	1.4 (0.4)	1.4 (0.5)	1.6 (0.5)	1.7 (0.5)	1.6 (0.4)	1.6 (0.5)

BILAG = British Isles Lupus Assessment Group; PGA = Physician Global Assessment; RCT = randomized controlled trial; SC = subcutaneous; SD = standard deviation; SFI = System Lupus Erythematosus Flare Index.; SELENA-SLEDAI = Safety of Estrogen in Lupus Erythematosus National Assessment–SLE Disease Activity Index.

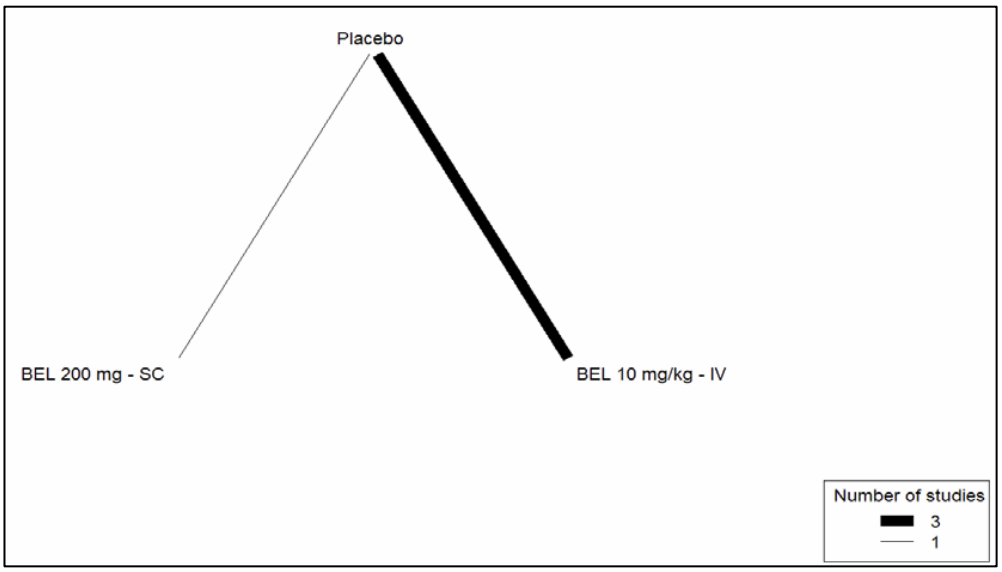
[a] Native Alaskan, Native Hawaiian, or Native Indian.

Source: Ramachandran et al.[18]

Results

Figure 4 presents the evidence network for the included studies in the ITC. The same network was used for all outcomes assessed.

Figure 4: Network of Studies Included in the Sponsor-Submitted ITC



ITC = indirect treatment comparison ; SC = subcutaneous.

Source: Ramachandran et al.[18]

Table 15 present the results for the outcomes SRI response at 52 weeks, for patients with a reduction of four points or more in SELENA-SLEDAI score at week 52, and severe flare occurrences over 52 weeks, for subgroups of patients who met criteria I and criteria II, respectively, for belimumab SC from NMAs. When belimumab SC was compared with belimumab IV, no difference was notable for any of the outcomes assessed. In comparison with placebo, belimumab SC was favoured for all of the outcomes assessed.

Results from the sensitivity analysis by excluding the northeast Asian study from the NMA or by using the Bucher method showed similar results to the base-case Bayesian model analysis. The ORs for the fixed-effects and random-effects models were in the same direction, with a wider CrI reported in the random-effects model.

Table 14: Efficacy Results for Patients in High Disease Activity Subgroup (Criteria I)

	OR (95% CrI) fixed-effects model		OR (95% CrI) random-effects model	
Outcome	**Belimumab 200 mg (SC) versus belimumab 10 mg/kg (IV)**	**Belimumab 200 mg (SC) versus placebo**	**Belimumab 200 mg (SC) versus belimumab 10 mg/kg (IV)**	**Belimumab 200 mg (SC) versus placebo**
SRI response at 52 weeks				
Base case	1.13 (0.67 to 1.92)	2.01 (1.27 to 3.19)	1.13 (0.29 to 4.35)	NR
sensitivity analysis based on data excluding North-east Asian study	1.15 (0.65 to 2.04)	NR	NR	NR
Sensitivity analysis using Bucher method — fixed-effects model, OR, 95% CI	1.13 (0.67 to 1.92)	2.00 (1.26 to 3.16)	NR	NR

Outcome	OR (95% CrI) fixed-effects model		OR (95% CrI) random-effects model	
	Belimumab 200 mg (SC) versus belimumab 10 mg/kg (IV)	Belimumab 200 mg (SC) versus placebo	Belimumab 200 mg (SC) versus belimumab 10 mg/kg (IV)	Belimumab 200 mg (SC) versus placebo
Patients with ≥ 4-point reduction in SELENA-SLEDAI score at week 52				
Base case	1.01 (0.60 to 1.72)	2.16 (1.37 to 3.45)	1.01 (0.27 to 3.85)	NR
Sensitivity analysis based on data excluding northeast Asian study	1.00 (0.56 to 1.79)	NR	NR	NR
Sensitivity analysis using Bucher method — fixed-effects model, OR, 95% CI	1.01 (0.60 to 1.72)	2.17 (1.35 to 3.45)	NR	NR
Severe flare occurrences over 52 weeks				
Base case	1.45 (0.78 to 2.73)	0.36 (0.21 to 0.62)	1.47 (0.29 to 7.69)	NR
sensitivity analysis based on data excluding northeast Asian study	1.56 (0.80 to 3.03)	NR	NR	NR
Sensitivity analysis using Bucher method — fixed-effects model, OR, 95% CI	1.47 (0.79 to 2.70)	0.36 (0.21 to 0.61)	NR	NR

CI = confidence interval; CrI = credible interval; OR = odds ratio; NR = not reported; SC = subcutaneous; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SRI = Systemic Lupus Erythematosus Responder Index.

Source: Ramachandran et al.[18]

Table 15: Efficacy Results for Patients in High Disease Activity Subgroup (Criteria II)

Outcome	OR (95% CrI) fixed-effects model		OR (95% CrI) random-effects model	
	Belimumab 200 mg (SC) versus belimumab 10 mg/kg (IV)	Belimumab 200 mg (SC) versus placebo	Belimumab 200 mg (SC) versus belimumab 10 mg/kg (IV)	Belimumab 200 mg (SC) versus placebo
SRI response at 52 weeks				
Base case	1.14 (0.75 to 1.72)	1.89 (1.35 to 2.66)	1.01 (0.39 to 3.13)	NR
Sensitivity analysis based on data excluding northeast Asian study	1.15 (0.72 to 1.82)	NR	NR	NR
Sensitivity analysis using Bucher method — fixed-effects model, OR, 95% CI	1.14 (0.75 to 1.72)	1.88 (1.34 to 2.65)	NR	NR
Patients with ≥ 4-point reduction in SELENA-SLEDAI score at week 52				
Base case	1.06 (0.70 to 1.59)	1.97 (1.40 to 2.78)	1.05 (0.24 to 4.35)	NR
Sensitivity analysis based on data excluding northeast Asian study	1.04 (0.66 to 1.64)	NR	NR	NR
Sensitivity analysis using Bucher method — fixed-effects model, OR, 95% CI	1.06 (0.70 to 1.59)	1.96 (1.39 to 2.78)	NR	NR
Severe flare occurrences over 52 weeks				
Base case	1.12 (0.67 to 1.89)	0.47 (0.30 to 0.73)	1.12 (0.23 to 5.56)	NR

Outcome	OR (95% CrI) fixed-effects model		OR (95% CrI) random-effects model	
	Belimumab 200 mg (SC) versus belimumab 10 mg/kg (IV)	Belimumab 200 mg (SC) versus placebo	Belimumab 200 mg (SC) versus belimumab 10 mg/kg (IV)	Belimumab 200 mg (SC) versus placebo
Sensitivity analysis based on data excluding northeast Asian study	1.25 (0.71 to 2.17)	NR	NR	NR
Sensitivity analysis using Bucher method — fixed-effects model, OR, 95% CI	1.12 (0.67 to 1.89)	0.47 (0.30 to 0.72)	NR	NR

CI = confidence interval; CrI = credible interval; NR = not reported; OR = odds ratio; SC = subcutaneous; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SRI = Systemic Lupus Erythematosus Responder Index.

Source: Ramachandran et al.[18]

Critical Appraisal of the Sponsor-Submitted ITC

No literature search was conducted to identify the appropriate trials to be included in this ITC. It is not clear how the studies were identified and selected to be included in this ITC.

There were substantial differences in the placebo effect for SRI response, which ranged from 0.282 to 0.341 in the belimumab IV trials (BLISS 52, BLISS 76, and BEL113750) and was 0.472 in the belimumab SC trial (BLISS SC) for patients in criteria I subgroup; and it ranged from 0.309 to 0.376 in the belimumab IV trials (BLISS 52, BLISS 76, and BEL113750) and was 0.488 in the belimumab SC trial (BLISS SC) for patients in criteria II subgroup. This variation in the placebo response among the studies signals the potential difference in background therapies in the patient population among the trials or other potential effect modifiers. Therefore, the transitivity or homogeneity assumption may not hold.

In all included studies in the ITC, patients were not stratified by criteria I or criteria II at randomization. Hence, the subgroup comparison breaks randomization, and patient characteristics (e.g., SELENA-SLEDAI category and SFI, in particular) appear to be imbalanced between treatment arms.

The strength of the network was low, with only four studies for two treatment options, and with only one of the four studies considering the clinical efficacy of belimumab SC formulation. In addition, the networks were centred on placebo, and all comparisons were indirect.

No quality assessment of included studies was reported.

The clinical experts consulted on this review indicated that patients who met criteria I or criteria II were patients with I.

Results between the Bayesian NMA and the Bucher method were consistent for all end points, supporting the validity of the findings from the NMA.

The ITC did not include any data on health-related quality of life, or on key safety outcomes, SAEs, or WDAEs.

The indirect comparison between the different administration routes of SC and IV resulted in wide CrIs, including null on all the outcomes analyzed. Of note, however, including the

null value in the 95% CrIs of the difference between treatments does not necessarily imply that the treatments are equivalent, comparable, or noninferior. There might still be differences between the two administration routes that the network is unable to identify.

Methods of ITC by Lee and Song[3]

Objectives

This ITC aimed to compare the efficacy and safety of belimumab 200 mg SC injection, and belimumab 1 and 10 mg/kg IV administration with those of placebo in patients with active SLE despite having received standard therapy, via a Bayesian NMA. The authors indicated that they did not receive financial support for the research, authorship, and/or publication of this article. Results for belimumab 1 mg/kg were not presented in this summary, as it is not a Health Canada–recommended dose.

Study Selection Methods

Multiple electronic databases, such as MEDLINE, Embase, and the Cochrane Central Register of Controlled Trials, were searched. The search covered articles from the databases' inception until January 2017 and had no language or race restrictions. Article references were also reviewed to identify additional studies that were not included in the electronic databases.

Studies were included if they were RCTs that compared belimumab with placebo in the treatment of patients with active SLE (defined as score ≥ 4 at screening on SELENA-SLEDAI) despite having received standard therapy. Studies provided end points for the clinical efficacy and safety of belimumab, and they included patients diagnosed with SLE based on the ACR criteria. Reports were excluded if they included duplicate data, or if they lacked adequate data for inclusion. It was not reported whether the study selection was performed by two independent reviewers.

Information on the first author, year of publication, the country in which the study was conducted, belimumab dose, number of patients treated with belimumab and placebo, clinical features, and safety and efficacy outcomes at 52 weeks were extracted. It was not reported whether data extraction was undertaken by more than one reviewer and whether quality was checked by a second reviewer.

The quality of the RCTs selected for inclusion in the NMA were assessed using Jadad scores. Quality was classified as high if the Jadad score was 3 to 5 or low if the Jadad score was 0 to 2. The Jadad scores of all of the included studies ranged from 3 to 5.

The efficacy outcome assessed was the SRI response rate at week 52 (defined as a ≥ 4-point reduction in SELENA-SLEDAI score, no new BILAG A organ domain scores and no more than one new BILAG B score, and no worsening in PGA score versus baseline). The safety outcome assessed was the number of patients who experienced SAEs to week 52.

ITC Analysis Methods

A Bayesian NMA was conducted using vague priors. A fixed-effects model was used; the rationale provided was that the patient population was homogeneous. The first 10,000 iterations were discarded as "burn-in," and results were based on an additional 10,000 iterations. The OR and 95% CrI were reported.

A sensitivity analysis was conducted by eliminating one study (BLISS 52[20]) from the analysis due to the low SRI response rate in the placebo group (SRI response rate in the placebo group was 11% in BLISS 52[20] versus 33%, 40%, and 48% in BLISS 76,[21] BEL113750,[22] and BLISS SC,[10] respectively).

The posterior mean deviance of the individual data points in the inconsistency model was plotted against their posterior mean deviance in the consistency model to assess network inconsistency between direct and indirect estimates in each loop.

Results of ITC Lee and Song[3]

Summary of Included Studies

Five RCTs (BLISS SC [Stohl et al.[10]], BEL113750 [Zhang et al.[22]], BLISS 52 [Navarra et al.[20]], BLISS 76 [Furie et al.[21]], and Wallace et al.[23]) met the inclusion criteria. Four of the RCTs (BLISS SC[10], BEL113750[22], BLISS 52[20], BLISS 76[21]) were included in the efficacy analysis of SRI at 52 weeks, and all five RCTs provided the data on SAEs at 52 weeks.

The studies BLISS SC, BEL113750, and Wallace et al. received Jadad score of 3 because they did not provide information on the appropriate method on randomization and blinding. The other two studies BLISS 52 and BLISS 76, were assigned a Jadad score of 5.

All patients included in the studies received standard therapy. The study design of BLISS SC, BEL113750, BLISS 52, and BLISS 76 are presented in Table 11 (in results of the sponsor-submitted ITC). Wallace et al. was conducted in the US and Canada, enrolled 338 patients, and compared belimumab 1 mg/kg and 10 mg/kg IV administration with placebo over 52 weeks. SELENA-SLEDAI score at entry had to be at least 4, average age of enrolled patients was 42.7 years, and average disease duration was nine years.

Demographics and disease characteristics for individual studies were not reported in the article by Lee and Song;[3] hence, it is not possible to comment on whether there were potential sources of heterogeneity among the patients in the included studies. However, the authors indicated in their discussion of results that there was heterogeneity in the design and patient characteristics of the included trials. Hence, these differences across studies may have affected the results of the ITC.

Results

Four RCTs (BLISS SC, BEL113750, BLISS 52, and BLISS 76 were included in the efficacy analysis of SRI at 52 weeks, and all five RCTs provided the data on SAE at 52 weeks.

No evidence network was presented in the article by Lee and Song.[3] The pairwise comparisons consisted of 11 direct comparisons (one study of belimumab 200 mg SC versus placebo, three studies of belimumab 1 mg/kg IV versus placebo, four studies of belimumab 10 mg/kg IV versus placebo, and three studies of belimumab 10 mg/kg IV versus belimumab 1 mg/kg IV).

In the base-case analysis, when belimumab SC was compared with belimumab IV for the outcome SRI response at 52 weeks, the belimumab IV group had a statistically higher percentage of responders (OR = 0.65 [95% CI, 0.45 to 0.93]). However, when a sensitivity analysis was conducted by omitting one outlier study (BLISS 52 [Navarra et al.[20]]) that reported low SRI response rate in the placebo group compared with the other three studies, no treatment was favoured (OR = 1.04 [95% CI, 0.71 to 1.51]). In comparison with placebo, belimumab SC was favoured in the base-case analysis and the sensitivity analysis.

Results for the SAE indicated that no treatment was favoured when belimumab SC was compared with belimumab IV (OR = 0.75 [95% CI, 0.46 to 1.22]). Belimumab SC was significantly better (reported less SAE) when compared with placebo (OR = 0.65 [95% CI, 0.43 to 0.99]).

Inconsistency plots assessing network inconsistencies showed a low possibility that inconsistencies significantly affected the ITC results.

Critical Appraisal of ITC Lee and Song[3]

The main limitation in the ITC by Lee and Song,[3] was potentially inaccurate data used in the NMA analysis. For example, Lee and Song[3] reported that the SRI response rate in BLISS 52[20] was 11% in the placebo group; however, Navarra et al.[20] reported that the placebo SRI response rate was 44% at week 52. If the reported SRI response rate of 44% was used instead of 11%, then, when belimumab SC is compared with belimumab IV, most likely no treatment would be favoured. This raises questions about the accuracy of all of the results reported by Lee and Song[3], and the results are questionable. It was not possible to check the accuracy of the data used in the ITC because such data used in the analysis (other than placebo response for SRI) were not reported.

The search covered literature from the databases' inception until January 2017, more than two years ago. Since then, there may have been new trials published, and these would have been excluded from the analysis, potentially impacting the conclusions of the NMA.

It was not reported whether literature selection was performed by two independent reviewers. It was not reported whether data extraction was undertaken by more than one reviewer and whether quality was checked by a second reviewer.

Demographics and disease characteristics for individual studies were not reported in the article; hence, it is not possible to comment on whether there were potential sources of heterogeneity among the patients in the included studies. However, the authors indicated in their discussion of results that there was heterogeneity in the design and patient characteristics of the included trials. Hence, these differences across studies may have affected the results of the ITC. In addition, that the transitivity assumption may not have been met.

A fixed-effects model was used; the rationale provided was that the patient population was homogeneous. However, the authors indicated in their discussion of results that there was heterogeneity in the design and patient characteristics of the included trials, which contradicts their rationale for using a fixed-effects model.

Another deficiency in the Lee and Song[3] ITC is the lack of reporting on several key items that would have allowed the reader to better assess the validity of the reported results. We cannot determine, for example, whether the model used in the analysis converged properly; we are unable to determine any level of potential statistical heterogeneity; and we are unable to determine whether the fixed-effects model was appropriate. Without these pieces of information, an assessment of the assumptions behind the use of indirect comparisons cannot be made. As a result, there is high uncertainty involved in the results presented by Lee and Song.[3]

The ITC did not include any data on health-related quality of life or WDAEs.

Methods of ITC by Tian et al.[4]

Objectives

The objective of this ITC was to assess the comparative safety of all available agents (immunosuppressive drugs, biologics, and GCs) used in the treatment of patients with SLE using an NMA. The authors indicated that their work was supported by the National Natural Science Foundation of China, the National Key Research and Development Program of China, and the National Key Clinical Speciality Construction Project of National Health and Family Planning Commission of the People's Republic of China. The authors also indicated that there were no competing interests.

Study Selection Methods

Multiple electronic databases such as PubMed, the Cochrane Central Register of Controlled Trials (CENTRAL), and Embase were searched. The search covered articles from the databases' inception until September 2017. Article references, review articles, and reports were manually scanned to identify additional studies that were not included in the electronic databases.

Studies were included if they were RCTs that reported the outcomes of interest. Studies were included if they included one of the following treatments of interest: immunosuppressants, biologics (rituximab and belimumab), or GCs. Studies had to enrol patients who met the 1987 ACR classification criteria for SLE in order to be included. Studies were excluded if they did not report on the outcomes of interest; if all arms of the trial had no events; or if trials lasted 24 weeks or less, enrolled children younger than 10 years old, enrolled women during pregnancy or lactation, or enrolled fewer than 20 patients. Scientific reports that presented duplicate data, or presented pooled trial data for which the individual trials could not be identified, were also excluded.

Two independent reviewers assessed the full text of the articles. In the case of disagreement in the final decision (i.e., inclusion or exclusion), a consensus was reached through discussion. In order to achieve a high level of agreement between reviewers, standardized data forms and data extraction training exercises were developed. However, it was not reported whether data extraction was done in duplicate, or by a single reviewer with checking by another reviewer, or if there was no checking at all.

The quality of the RCTs selected for inclusion in the NMA was assessed by the authors using the Cochrane Collaboration's tool for assessing risk of bias.

Outcomes assessed were all-cause mortality, withdrawals related to AEs, AEs, SAEs, cardiovascular events, serious infections (major infection, severe infection, sepsis, cardiovascular infection, or bacterial pneumonia), bone toxicity, malignant transformation, serious gastrointestinal events, ovarian failure, menstrual disorder, new-onset hypertension, serious leucopenia, leucopenia. and hyperglycemia. However, results for comparison of belimumab 200 mg SC versus other treatments was only available for all-cause mortality, withdrawals related to AEs, AEs, SAEs, and serious infection.

ITC Analysis Methods

An NMA was conducted to obtain estimates for all outcomes. The estimates were presented as ORS with 95% CIs. A random-effects model was used; however, no rationale was provided. In order to evaluate the assumptions of transitivity in the studies, network meta-regression within the frequentist framework was implemented. In addition, graphical representation of the results was also provided. The extent of heterogeneity in every network was assessed by comparing the empirical distribution of heterogeneity variances specific to the types of outcomes and treatments being compared with the magnitude of tau (τ) for the network. Values higher than 1.0 represented high heterogeneity, outcomes from 0.1 to 1.0 were considered moderate, and outcomes lower than 0.1 were considered low.

Inconsistency was assessed using a loop-specific approach by investigating the consistency within every closed triangular or quadratic loop in every network.[24,25] In addition, the design-by-treatment interaction model that provides a single inference and chi-square (χ^2) tests were used to assess the assumption of consistency in the entire network.[26]

Several sensitivity analyses were conducted to assess the generalizability of the findings; however, belimumab 200 mg SC was not included in any of the sensitivity analyses.

Results of ITC by Tian et al.[4]

Summary of Included Studies

Forty-four trials were included in the systematic review, with 9,898 patients identified and included. Of the RCTs included, nine were three-arm trials, one was a four-arm trial, and one was a five-arm trial. Of the RCTs included, 21 trials used IV cyclophosphamide (CYC) 0.5 g/m^2 to 1 g/m^2 body surface area monthly; 12 trials used azathioprine (AZA) 1 mg/kg to 4 mg/kg per day; two trials used mycophenolate mofetil (MMF) 500 mg to 3,000 mg per day; four trials used tacrolimus (TAC) 0.05 mg/kg to 0.1 mg/kg/day; two trials used oral CYC 1 mg/kg to 4 mg/kg per day; five trials used cyclosporine 1 mg/kg to 5 mg/kg per day; three trials used methotrexate 7.5 mg to 20 mg per week; two trials used rituximab 1 g per day; four trials used low-dose belimumab 1 mg/kg IV; one trial used moderate-dose belimumab 4 mg/kg IV; five trials used high-dose belimumab 10 mg/kg IV; one trial used SC belimumab 200 mg per week; one trial used leflunomide 1 mg/kg per day; one trial used chloroquine 150 mg per day; one trial used AZA combined with GC; two trials used MMF combined with TAC; one trial used IV CYC combined with MMF; two trials used CYC followed by AZA; and one trial used AZA combined with CYC.

The risk of bias in the studies, as assessed by the authors using the Cochrane Collaboration's tool for assessing risk of bias, was low for most criteria and unclear for some criteria.

The majority of patients included in the studies were women. The sample size in many of the studies was small. The duration of disease, SLEDAI score, and duration of study follow-up were substantially different among the studies. Characteristics of the included studies are presented in Table 16.

Table 16: Characteristics of Included Studies in Tian et al.[4]

Study	N	Interventions	Age (years)	Female, %	Types	Duration of disease (years)	SLEDAI	Follow-up (month)
Barile-Fabris et al. (2005)	19 13	IV CYC IV GC	33 26	93.7	NM	4.2 2.5	–	12
Hahn et al. (1975)	11 13	AZA Placebo	33 31	82 85	SLE, I	13.7 13.5	–	24
Fortin et al. (2008)	41 45	MTX Placebo	40 40	90 91	SLE	5.7 4.5	10.0 10.0	12
Griffiths et al. (2010)	47 42	CSA AZA	33 39	96 88	SLE	2 4	–	12
Merrill et al. (2010)	169 88	IV RTX Placebo	40.2 40.5	89.9 93.2	SLE	8.5 8.7	–	12
Wallace et al. (2013)	673 111 674 675	IV Belimumab LD IV Belimumab MD IV Belimumab HD Placebo	38.2 42.6 38.5 38.8	93.8 94.6 95.5 92.4	SLE	6.8 10.1 6.5 6.9	9.7 9.4 9.7 9.7	12
Navarra et al. (2011)	288 290 287	IV Belimumab LD IV Belimumab HD Placebo	35.0 35.4 36.2	94 97 94	SLE	5.0 5.0 5.9	9.6 10.0 9.7	12
Van Vollenhoven et al. (2012)	284 305 287	IV Belimumab LD IV Belimumab HD Placebo	–	–	SLE	–	10.8	12
Furie et al. (2011)	271 273 275	IV Belimumab LD IV Belimumab HD Placebo	40.0 40.5 40.0	93.4 94.9 91.6	SLE	7.9 7.2 7.4	9.7 9.5 9.8	17
Rovin et al. (2012)	73 71	IV RTX Placebo	31.8 29.4	87.5 93.1	LN, I	2.7 2.4 (LN)	–	12
Stohl et al. (2017)	556 280	SC Belimumab Placebo	38.1 39.6	93.7 95.7	SLE	4.3 4.6	10.5 10.3	12
Zhang et al. (2016)	446 217	IV Belimumab HD Placebo	–	–	SLE	–	>8	12
Contreras et al. (2004)	19 20 20	IV AZA IV CYC IV MMF	33 33 32	94.7 90 95	LN, M	5.4 2.9 3.9	7.5 8.4 9	72
Donadio et al. (1978)	24 26	CYC GC	30.2 32.3	79.2 84.6	LN, I	4.3 4.8	–	50
Chan et al. (2000)	21 21	MMF CYC-AZA	36 39	95.2 90.5	LN, I, and M	6 8.1	–	12

Study	N	Interventions	Age (years)	Female, %	Types	Duration of disease (years)	SLEDAI	Follow-up (month)
Zavada et al. (2010)	21 19	IV CYC CSA	30 28	71.4 73.7	LN, I, and M	–	19.9 19.3	18
Ginzler et al. (2005)	71 69	MMF IV CYC	32.5 31.0	86 94	LN, I	3.6 4.9	–	36.2 37.2
Gourley et al. (1996)	27 27 28	IV GC IV CYC IV GC + IV CYC	30 30 31	81.5 77.8 89.3	LN, I	2.6 2 3.3 LN	–	> 60
Moroni et al. (2006)	36 33	CSA AZA	31.7 31.2	91.7 87.9	LN, M	5.4 2.3	–	48
Ong et al. (2005)	25 19	IV CYC MMF	30.5 31.3	88 79	LN, I	2.7 4.1	14.8 15.8	6
Sesso et al. (1994)	14 15	IV CYC IV GC	30.0 24.3	85.7 86.7	LN, I	3.5 3.7	–	15
Steinberg and Steinberg (1991)	30 20 18 20 23	GC AZA CYC IV CYC CYC + AZA	27 26 34 31 29	–	LN, I	3 3 3 3 3	–	84
Mok et al. (2001)	22 21	IV pulse CYC Pulse CYC-AZA	33.5 30.1	95 95	LN, I, and M	2.8 2.6	–	24
Wang et al. (2008)	70 40	LEF IV CYC	31.3 33.1	85.7 92.5	LN, I	0.8 0.5	≥ 8	6
Appel et al. (2009)	184 180	MMF IV CYC	32.4 32.3	84.9 84.3	LN, I, and M	1.0 1.0 LN	–	6
Li et al. (2012)	20 20 20	MMF TAC IV CYC	26.5 29 33	86.7 86.7 90.0	LN, I	0.25 0.25 0.17 LN	18.2 18.3 18.9	6
Mok et al. (2016)	76 74	MMF TAC	36.1 36.2	89 95	LN, I	4.0 4.4	–	6
Ordi-Ros et al. (2017)	120 120	AZA MMF	40.9 42.1	92.5 90	SLE	5.1 6.2	9.5 9.9	24
Islam et al. (2012)	13 24	MTX Chloroquine	24.0 24.9	100 95.8	SLE	1.3 1.0	12.5 13.3	6
Carneiro and Sato (1999)	20 21	MTX Placebo	–	–	SLE	6.9	–	6
Cade et al. (1973)	13 13 15	AZA AZA + GC GC	30.0 26.1 22.4	76	LN, I	–	–	72
Boumpas et al. (1992)	40 25	IV CYC IV GC	29	92.3	LN, I	2.75	–	60
El-Sehemy et al. (2006)	7 7 8	IV CYC CSA AZA	25.6 22.0 21.4	100	LN, I	1.5 2.4 1.8	29.1 27.9 30.9	6

Study	N	Interventions	Age (years)	Female, %	Types	Duration of disease (years)	SLEDAI	Follow-up (month)
Grootscholten et al. (2006)	50 37	IV CYC AZA	30 33	88 84	LN, I	26 26	19 20	66.0-75.6
Bao et al. (2008)	20 20	MMF + TAC IV CYC	27.2 30.6	80 90	LN, I	3.0 3.7	14.9 14.0	6
El-Shafey et al. (2010)	24 23	MMF IV CYC	23.8 22.8	95.9 95.7	LN, I	0.3 0.2 LN	–	6
Chen et al. (2011)	39 34	TAC IV CYC	32.0 31.9	88.1 82.1	LN, I	3.8 3.3 LN	19.9 18.1	6
Liu et al. (2015)	181 181	MMF + TAC IV CYC	30.0 33.6	92.8 89.0	LN, I	0.2 0.3	16.0 15.0	6
Rathi et al. (2016)	50 50	IV CYC MMF	30.6 28.3	90 94	LN, I	–	18.1 17.9	6
Sun et al. (2015)	40 42	IV CYC IV CYC + MMF	33.3 31.9	92.5 90.5	LN, I	–	13.8 14.1	6
Dooley et al. (2011)	116 111	MMF AZA	31.8 31.0	85.3 86.5	LN, M	5.8 4.9	–	36
Houssiau et al. (2010)	52 53	AZA MMF	33 33	92.3 90.6	LN, M	–	17 19	110
Chen et al. (2012)	34 36	TAC AZA	30.7 33.1	85.3 88.9	LN, M	0.2 0.2	5.2 5.1	6
Austin et al. (2009)	15 15 12	GC IV CYC CSA	40 41 34	80 80 91.7	LN, I	1.25 1.67 1.25	1 0.6 0.5 LN	12

AZA = azathioprine; CSA = cyclosporine; CYC = cyclophosphamide; CYC-AZA = CYC followed by AZA; GC = glucocorticoid; HD = high-dose; I = induction therapy; LD = low-dose; LEF = leflunomide; LN = lupus nephritis; M = maintenance therapy; MD = moderate-dose; MMF = mycophenolate mofetil; MTX = methotrexate; NM = neurological manifestations; RTX = rituximab; SC = subcutaneous; SLE = systemic lupus erythematosus; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; TAC = tacrolimus.

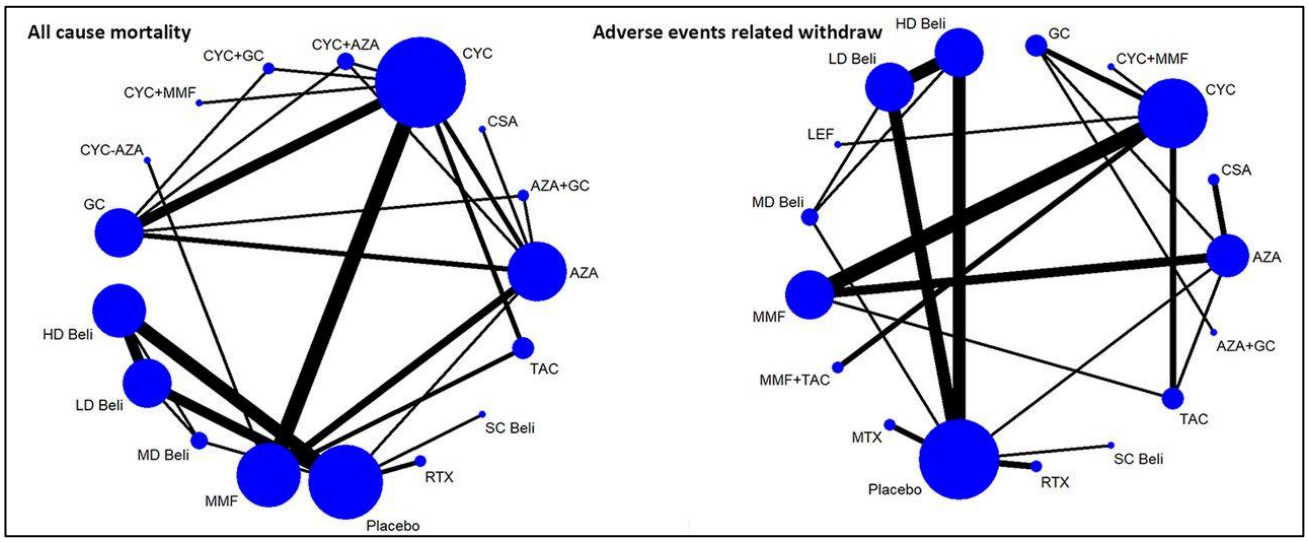
Note: Adapted according to the Creative Commons Licence BY-NC/4.0.

Source: Tian et al.[4]

Results

Figure 5, Figure 6, and Figure 7 present the evidence networks for all-cause mortality, withdrawals related to AEs, AEs, and SAEs. In all of these figures, the size of the nodes (blue circles) corresponds to the number of trials of the treatments. Comparisons are linked with a line, the thickness of which corresponds to the number of trials that assessed the comparison.

Figure 5: Evidence Networks of Treatment Comparisons for All-Cause Mortality and Withdrawal Due to Adverse Events

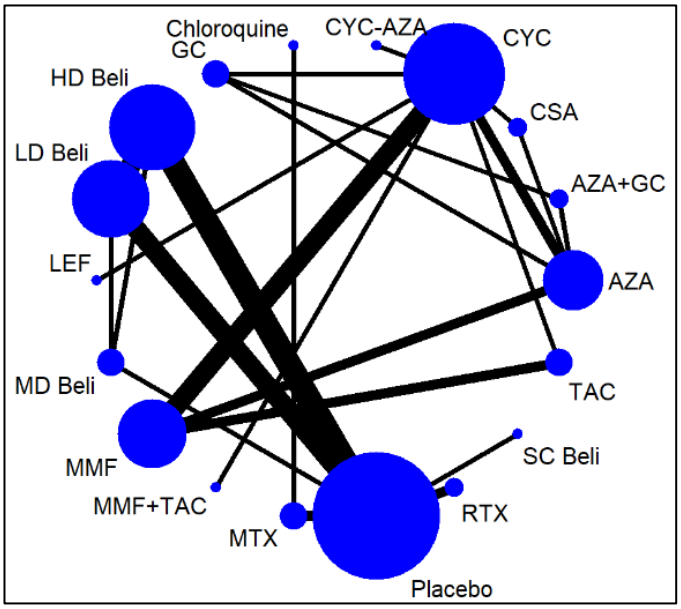


AZA = azathioprine; Beli = belimumab; CSA = cyclosporine; CYC = cyclophosphamide; CYC-AZA = CYC followed by AZA; GC = glucocorticoid; HD = high-dose; LD = low-dose; LEF = leflunomide; MD = moderate-dose; MMF = mycophenolate mofetil; MTX = methotrexate; RTX = rituximab; SC = subcutaneous; TAC = tacrolimus.

Note: Adapted according to the Creative Commons Licence BY-NC/4.0.

Source: Tian et al.[4]

Figure 6: Evidence Networks of Treatment Comparisons for Adverse Events

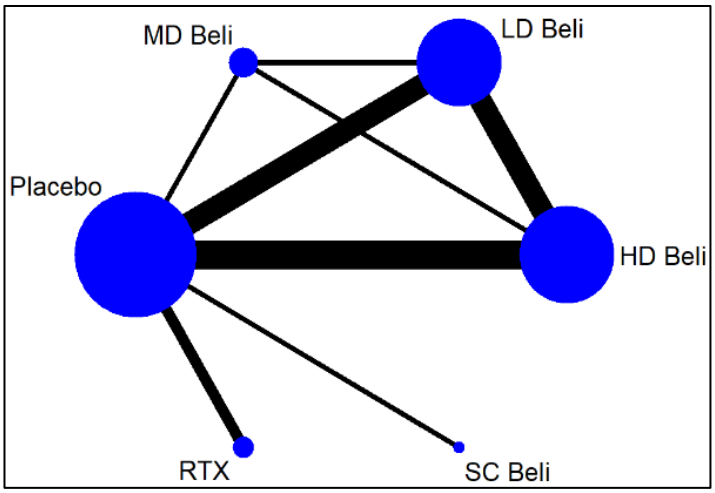


AZA = azathioprine; Beli = belimumab; CSA = cyclosporine; CYC = cyclophosphamide; CYC-AZA = CYC followed by AZA; GC = glucocorticoid; HD = high-dose; LD = low-dose; LEF = leflunomide; MD = moderate-dose; MMF = mycophenolate mofetil; MTX = methotrexate; RTX = rituximab; SC = subcutaneous; TAC = tacrolimus.

Note: Adapted according to the Creative Commons Licence BY-NC/4.0.

Source: Tian et al.[4]

Figure 7: Evidence Networks of Treatment Comparisons for Serious Adverse Events



Beli = belimumab; HD = high-dose; LD = low-dose; MD = moderate-dose; RTX = rituximab; SC = subcutaneous.

Note: Adapted according to the Creative Commons Licence BY-NC/4.0.

Source: Tian et al.[4]

For the outcomes (all-cause mortality, withdrawals related to AEs, and serious infection), belimumab 200 mg SC was compared with belimumab 10 mg/kg, rituximab, immunosuppressive drugs, GCs, and placebo. No treatment was favoured in any of these comparisons.

For the outcome SAEs, belimumab 200 mg SC was compared with belimumab 10 mg/kg, and rituximab. No treatment was favoured in any of these comparisons. When belimumab 200 mg SC was compared to placebo, it was found that there were significantly less SAEs with belimumab 200 mg SC than with placebo treatment.

For the outcome AEs, belimumab 200 mg SC was compared with belimumab 10 mg/kg, rituximab, immunosuppressive drugs, and placebo. No treatment was favoured in any of these comparisons.

The authors indicated that in the NMAs, statistical heterogeneity was low in most networks. However, there was substantial heterogeneity in the networks for AEs. Inconsistency was only noted in the network for AEs.

Table 17: Network Meta-Analysis Results for All-Cause Mortality, Withdrawal Due to Adverse Events, Adverse Events, Serious Adverse Events, and Serious Infection Reported by Tian et al.

Belimumab 200 mg (SC) versus	OR (95% CrI) random-effects model				
	All-cause mortality	Withdrawals related to adverse events	Adverse events	Serious adverse events	Serious infection
Belimumab 10 mg/kg (IV)	0.5 (0.07 to 3.57)	0.85 (0.47 to 1.54)	0.79 (0.52 to 1.22)	0.64 (0.41 to 1.00)	0.81 (0.4 to 1.67)
RTX	0.32 (0.02 to 4.17)	1.18 (0.48 to 2.86)	0.50 (0.25 to 1.02)	0.71 (0.39 to 1.29)	1.27 (0.53 to 3.03)
Placebo	0.76 (0.13 to 4.55)	0.79 (0.47 to 1.33)	0.78 (0.52 to 1.15)	**0.65 (0.43 to 0.99)[a]**	0.76 (0.39 to 1.49)
CYC	2.16 (0.47 to 99.19)	0.62 (0.20 to 19.08)	NR	NR	0.85 (0.12 to 6.08)
GC	1.35 (0.03 to 63.08)	0.13 (0.00 to 5.88)	NR	NR	1.07 (0.13 to 8.84)
AZA	2.08 (0.05 to 88.97)	0.21 (0.01 to 5.79)	NR	NR	0.90 (0.14 to 6.01)
MMF	2.06 (0.04 to 98.21)	0.43 (0.01 to 12.44)	NR	NR	1.08 (0.15 to 7.72)
CSA	0.45 (0.00 to 56.83)	0.19 (0.01 to 6.01)	NR	NR	1.03 (0.10 to 11.12)
TAC	4.89 (0.08 to 313.48)	0.75 (0.02 to 26.26)	NR	NR	2.66 (0.31 to 23.05)
CYC + GC	2.63 (0.04 to 100)	NR	NR	NR	1.08 (0.08 to 14.29)
CYC + AZA	1.64 (0.03 to 100)	NR	NR	NR	2.33 (0.13 to 50)
CYC + MMF	2.27 (0.02 to 1,000)	1.35 (0.02 to 100)	NR	NR	2.7 (0.06 to 100)
AZA+GC	1.67 (0.03 to 100)	1.06 (0.01 to 100)	NR	NR	0.66 (0.04 to 10.0)
CYC-AZA	0.37 (0.00 to 50)	NR	NR	NR	0.58 (0.06 to 5.56)
MMF + TAC	NR	0.22 (0.01 to 8.33)	NR	NR	0.41 (0.05 to 3.7)
MTX	NR	0.61 (0.17 to 2.17)	0.44 (0.17 to 1.10)	NR	NR
LEF	NR	0.54 (0.01 to 25)	NR	NR	0.74 (0.03 to 16.67)
Chloroquine	NR	NR	4.90 (0.78 to 30.85)	NR	NR

AZA = azathioprine; CrI = credible interval; CSA = cyclosporine; CYC = cyclophosphamide; CYC-AZA = CYC followed by AZA; GC = glucocorticoid; LEF = leflunomide; MMF = mycophenolate mofetil; MTX = methotrexate; NR = not reported; OR = odds ratio; RTX = rituximab; SC = subcutaneous; TAC = tacrolimus.

[a]Bold text indicates statistical significance

Source: Tian et al.[4]

Critical Appraisal of ITC by Tian et al.[4]

The article by Tian et al.[4] provided a description of the methods, which included a clear population, intervention, comparisons, and outcomes. The article by Tian et al.[4] conducted a literature search using several databases up to September 2017. As a result, the ITC may have missed several related trials that were published since that time. Screening and selection of literature was conducted appropriately; however, it was not reported whether

data extraction was undertaken by more than one reviewer and whether quality was checked by a second reviewer. The ITC description of the approach used to conduct the NMA was not clear, and additional details on the methods used would have been helpful. However, heterogeneity and inconsistency between direct and indirect evidence were appropriately assessed.

An important consideration when including a large number of trials in a network is whether these trials are sufficiently similar, methodologically and clinically, to warrant valid comparisons. Only a very high-level summary of trial and patient characteristics was provided; hence, it was not possible to assess fully how homogeneous the patients in the trials were. However, from the available data, it seems that the duration of disease, SLEDAI score, and duration of follow-up differed substantially among the studies, indicating that patients included in the trials were most likely heterogeneous. It was not clear whether meta-regression was used to adjust for potential effect modifiers. In addition, the sample size in many of the studies was small. All of these limitations might have yielded the wider CrIs reported. As a result, there is high uncertainty involved in the results presented by Tian et al.[4]

Summary

Three ITCs were summarized and critically appraised. One ITC was submitted by the sponsor,[18,19] and two additional ITCs (Lee and Song[3] and Tian et al.[4]) were identified in the literature.

The ITC submitted by the sponsor evaluated the comparative clinical efficacy of the IV and SC formulations of belimumab for the treatment of adult patients with active, autoantibody-positive SLE and I who are receiving standard therapy. Patient-level data from pivotal phase III trials were used to identify patients who meet the I criteria. When belimumab SC was compared with belimumab IV for the outcomes of SRI response at 52 weeks, patients with a reduction of four points or more in SELENA-SLEDAI score at week 52, and severe flare occurrences at week 52, no treatment administration route (either SC or IV) was favoured for any of the outcomes assessed. In comparison with placebo, belimumab SC was favoured for all of the outcomes assessed.

The ITC by Lee and Song[3] aimed to compare the efficacy and safety of belimumab 200 mg SC injection, and belimumab 1 mg/kg and 10 mg/kg IV administration with those of placebo in patients with active SLE despite having received standard therapy. Lee and Song[3] included all patients who were enrolled in the trials, not just a subgroup of patients as was done in the sponsor-submitted ITC. In the base-case analysis, when belimumab SC was compared with belimumab IV for the outcome SRI response at 52 weeks, belimumab IV was significantly better. This result was different than that found in the sponsor-submitted ITC, which reported that no treatment was favoured. However, it seems that some data used in the ITC were inaccurate. It was reported in Lee and Song[3] that the SRI response rate in BLISS 52[20] in the placebo group was 11%; however, Navarra et al.[20] reported that placebo SRI response rate was 44% at week 52, which is completely different than what was used in the base-case analysis. This raise questions about the accuracy of all the results reported by Lee and Song,[3] and the results are questionable. When a sensitivity analysis was conducted by omitting Navarra et al.,[20] no treatment was favoured, and results were in line with the sponsor-submitted ITC. Results for the SAEs indicated that no treatment was favoured when belimumab SC was compared with belimumab IV or with placebo.

The objective of the ITC by Tian et al.[4] was to assess the comparative safety of all available agents (immunosuppressive drugs, biologics, and GCs) used in the treatment of patients with SLE. When belimumab 200 mg SC was compared with belimumab 10 mg/kg, rituximab, immunosuppressive drugs, and glucocorticoids for the outcomes all-cause mortality, withdrawals related to AEs, AEs, SAEs, and serious infection, no treatment was favoured. However, there were many limitations to this ITC, such as the small number of patients included in most of the trials, different disease duration, SLEDAI score, and duration of follow-up, indicating significantly heterogenous patient population. As a result, there is high uncertainty involved in the results presented by Tian et al.[4]

No ITC was conducted that assessed quality of life.

Overall, findings from ITCs suggest there is no difference in efficacy between the SC and IV formulations of belimumab, and no difference between belimumab and other commonly used drugs for SLE (immunosuppressants, corticosteroids, or biologics) with respect to harms, although the latter finding has a high degree of uncertainty.

Other Relevant Studies

Health Canada provided a Notice of Compliance to SC belimumab in December 2017 based on the BLISS SC trial as well as data available for the IV formulation.[12] In this section, the pivotal studies for the IV formulation are reviewed, along with long-term data that are available for both the IV and SC formulations.

The following studies are summarized:

A. Two pivotal RCTs of IV belimumab (BLISS 52 and BLISS 76).[20,21,27,28]
B. Open-label extensions of IV belimumab:
   - Phase II extension (LBSL99)[29]
   - Phase III extension of BLISS 76 in patients from the US (BLISS 76 Extension US)[2,30,31]
   - Phase III extension of BLISS 52 and BLISS 76 in patients outside of the US (BLISS 52 and BLISS 76 Extension non-US)[32,33]
   - Phase III pooled extensions of US and non-US patients from BLISS 52 and BLISS 76 Extensions (pooled US and non-US extensions).[34,35]
C. Open-label extension of SC belimumab:
   - Phase III extension (BLISS SC Extension)36
D. A longitudinal propensity-score–matched study of IV belimumab + standard of care compared with standard of care alone.[37]
E. Pivotal RCTs of IV Belimumab.

A. BLISS 52 and BLISS 76

BLISS 52 (NCT00424476) was a phase III multi-centre (13 countries in Latin America, Asia–Pacific, and Eastern Europe), DB RCT in adults 18 years of age or older who met the ACR criteria for SLE and who had active disease (score ≥ 6 on SELENA-SLEDAI at screening).[20,27] BLISS 76 (NCT00410384) was a phase III, multi-centre (19 countries in Europe and North/Central America), DB RCT, also in adults 18 years of age or older who met the ACR criteria for SLE and who had active disease (SELENA-SLEDAI score ≥ 6 at screening).[21,28] In both studies, patients were also required to have positive ANA (titre

≥ 1:80) or anti-dsDNA (≥ 30 IU/mL), and a stable treatment regimen with fixed doses of one or more of prednisone (0 mg to 40 mg per day), NSAIDs, antimalarial drugs, or immunosuppressant drugs for at least 30 days before the first study dose. Patients with severe active lupus nephritis or CNS lupus were excluded. Randomization was in a 1:1:1 ratio to IV placebo, belimumab 1 mg/kg, or belimumab 10 mg/kg in addition to standard of care. The Health Canada–approved dose for belimumab is 10 mg/kg; therefore, the data presented here do not include the 1 mg/kg group.[6] Patients received infusion over one hour on days 0, 14, and 28, followed by every 28 days thereafter, until week 48 in BLISS 52 (patients were followed for a total of 52 weeks) or week 72 in BLISS 76 (patients were followed for a total of 76 weeks). Randomization was stratified by SELENA-SLEDAI score (6 to 9 versus ≥ 10), proteinuria concentration (< 2 g in 24 hours versus ≥ 2 g in 24 hours), and ethnic origin (African descent or Indigenous American versus other).

In both studies, the primary outcome was the proportion of SRI responders at week 52. An SRI responder was defined as meeting all of the following: a reduction of at least four points in the SELENA-SLEDAI score, no new BILAG A organ domain score, no more than one new BILAG B organ domain score, and no worsening in PGA score (increase > 0.3) at week 52 compared with baseline.

For BLISS 52, major secondary outcomes were proportion of patients with at least a four-point reduction from baseline in SELENA-SLEDAI score at week 52, mean change in PGA score at week 24, mean change in SF-36 PCS at week 24, and proportion of patients with average reduction in prednisone dosage of at least 25% from baseline to 7.5 mg per day or less during weeks 40 to 52. Other outcomes were SFI and BILAG scores during the 52 weeks, and harms. No adjustments were made to control for type I error arising from multiple statistical tests.

For BLISS 76, major secondary outcomes were the proportion of SRI responders at week 76, percentage of patients with reduction from baseline in SELENA-SLEDAI score of four points or more at week 52, change in PGA score at week 24, change in SF-36 PCS at week 24, and percentage of patients with mean prednisone dosage decrease of 25% or more from baseline to 7.5 mg per day or less during weeks 40 to 52. Other outcomes were disease activity assessed with the SFI and harms.

The primary outcome in both studies was evaluated with a logistic regression model that adjusted for baseline randomization stratification factors (i.e., SELENA-SLEDAI score ≤ 9 or ≥ 10, proteinuria < 2 g in 24 hours or ≥ 2 g in 24 hours, and race: African descent or Indigenous American descent, or other). The modified ITT population was analyzed, which included all randomized patients who received a dose of study drug. All analyses were adjusted for baseline stratification factors.

Baseline Characteristics

The baseline characteristics of patients in the placebo and belimumab 10 mg/kg groups of BLISS 52 and BLISS 76 are provided in

Table 18. In BLISS 52, patients were predominantly female and were primarily of Indigenous American (Alaska Native or American Indian from North, South, or Central America), white, Asian, or Hispanic/Latino ethnic origin. Patients had a high baseline level of disease activity, as shown by mean SELENA-SLEDAI score close to 10, with more than 50% of patients having a baseline SELENA-SLEDAI score of 10 points or more. Treatment groups were well balanced with respect to baseline SLE disease characteristics.

In BLISS 76, most patients were female and of white ethnic origin. Patients were on average slightly older than patients in BLISS 52 and had had disease for more years. As with BLISS 52, patients had a high baseline level of disease activity, as shown by mean SELENA-SLEDAI score close to 10, and close to 50% of patients had a baseline SELENA-SLEDAI score of 10 points or more. More patients in the placebo group (68%) had BILAG 1A or 2B score than the placebo group (59%).

Table 18: Baseline Patient Characteristics in BLISS 52 and BLISS 76

Characteristic	BLISS 52		BLISS 76	
	Placebo (N = 287)	Belimumab 10 mg/kg (N = 290)	Placebo (N = 275)	Belimumab 10 mg/kg (N = 273)
Age, years, mean (SD)	36.2 (11.8)	35.4 (10.8)	40.0 (11.9)	40.5 (11.1)
Female, n (%)	270 (94)	280 (97)	252 (91.6)	259 (94.9)
Ethnic origin, n (%)				
Indigenous American	89 (31)	92 (32)	36 (13.1)	34 (12.5)
White	82 (29)	71 (24)	188 (68.4)	189 (69.2)
Black American	11 (4)	11 (4)	39 (14.2)	40 (14.8)
Asian	105 (37)	116 (40)	11 (4.0)	11 (4.0)
Hispanic or Latino	143 (50)	136 (47)	55 (20.0)	56 (20.5)
Disease duration, years, mean (SD)	5.9 (6.2)	5.0 (5.1)	7.4 (6.7)	7.2 (7.5)
SELENA-SLEDAI score, mean (SD)	9.7 (3.6)	10.0 (3.9)	9.8 (4.0)	9.5 (3.6)
SELENA-SLEDAI score ≥ 10, n (%)	158 (55)	160 (55)	140 (50.9)	136 (49.8)
BILAG 1A or 2B score, n (%)	166 (58)	172 (59)	187 (68.0)	160 (58.6)
PGA score, mean (SD)	1.4 (0.5)	1.4 (0.5)	1.5 (0.5)	1.4 (0.5)
BILAG A or B organ domain, n (%)				
General	28 (10)	26 (9)	38 (13.8)	38 (13.9)
Mucocutaneous	172 (60)	174 (60)	178 (64.7)	141 (51.6)
Neurological	0 (0)	0 (0)	6 (2.2)	7 (2.6)
Musculoskeletal	147 (51)	160 (55)	195 (70.9)	179 (65.6)
Cardiovascular and respiratory	12 (4)	6 (2)	9 (3.3)	15 (5.5)
Vasculitis	22 (8)	33 (11)	30 (10.9)	18 (6.6)
Renal	38 (13)	34 (12)	21 (7.6)	24 (8.8)
Hematology	52 (18)	53 (18)	35 (12.7)	35 (12.8)
SDI score, mean (SD)	NR	NR	1.0 (1.5)	1.0 (1.4)
Proteinuria, g in 24 hours, mean (SD)	0.6 (1.2)	0.5 (0.9)	0.4 (0.8)	0.4 (0.7)
ANA ≥ 1:80, n (%)	264 (92)	276 (95)	253 (92.0)	245 (89.7)
Anti-dsDNA ≥ 30 IU/mL, n (%)	205 (71)	218 (75)	174 (63.3)	179 (65.6)
Prednisone, n (%)	276 (96)	278 (96)	212 (77.1)	200 (73.3)
Prednisone dosage, mg/day, mean (SD)	11.9 (7.9)	13.2 (9.5)	9.4 (8.9)	8.4 (7.9)

Characteristic	BLISS 52		BLISS 76	
	Placebo (N = 287)	Belimumab 10 mg/kg (N = 290)	Placebo (N = 275)	Belimumab 10 mg/kg (N = 273)
Immunosuppressants, n (%)	122 (43)	123 (42)	154 (56.0)	148 (54.2)
Antimalarial drugs, n (%)	201 (70)	185 (64)	180 (65.5)	168 (61.5)

ANA = antinuclear antibody; BILAG = British Isles Lupus Assessment Group; NR = not reported; PGA = Physician Global Assessment; SD = standard deviation; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index.

Source: Clinical Study Report for BLISS 52;[27] Navarra et al. (2011);[20] Clinical Study Report for BLISS 76;[28] Furie et al. (2011).[21]

Efficacy

Table 19 provides the results for efficacy outcomes in BLISS 52. The primary outcome, proportion of SRI responders at week 52, was statistically significantly greater in the belimumab 10 mg/kg group compared with placebo. For the major secondary outcomes, statistically significantly more patients in the belimumab group achieved a four-point or greater reduction in SELENA-SLEDAI score at week 52 and had a lower PGA score change from baseline at week 24. There was no statistical difference between belimumab and placebo in the SF-36 PCS change from baseline at week 24 or in prednisone dosage reduction of 25% or more to 7.5 mg per day or less during weeks 40 to 52.

Table 19: Efficacy in BLISS 52

	BLISS 52	
	Placebo (N = 287)	Belimumab 10 mg/kg (N = 290)
Primary outcome		
Proportion achieving SRI response at week 52	125 (44%)	167 (58%)
OR (95% CI) (belimumab versus placebo)	1.83 (1.30 to 2.59)	
P value	0.0006	
Major secondary outcomes		
≥ 4-point reduction in SELENA- SLEDAI at week 52		
n (%)	132 (46%)	169 (58%)
OR (95% CI) (belimumab versus placebo)	1.71 (1.21 to 2.41)	
P value	0.0024	
PGA score at week 24 (change from baseline)		
Least square mean change (SE)	–0.35 (0.04)	–0.50 (0.04)
Difference (95% CI) (belimumab versus placebo)	–0.15 (–0.23 to –0.07)	
P value	0.0003	
SF-36 PCS score at week 24 (change from baseline)		
Least square mean change (SE)	3.26 (0.54)	3.34 (0.55)
Difference (95% CI) (belimumab versus placebo)	0.08 (–1.00 to 1.15)	

	BLISS 52	
	Placebo (N = 287)	Belimumab 10 mg/kg (N = 290)
P value	0.8870	
Prednisone dosage reduced ≥ 25% to ≤ 7.5 mg/day during weeks 40 to 52		
n (%)	23/192 (12)	38/204 (19)
OR (95% CI) (belimumab versus placebo)	1.75 (0.99 to 3.08)	
P value	0.0526	
Other secondary outcomes		
Disease flares (all)		
With flare, n (%)	230 (80)	205 (71)
SFI time to first flare, days, median (range)	84 (1–368)	119 (1-367)
HR (95% CI) for time to flare (belimumab versus placebo)	0.76 (0.63 to 0.91)	
P value	0.0036	
Disease flares (severe)		
With flare, n (%)	66 (23)	40 (14)
OR (95% CI) (belimumab versus placebo)	0.57 (0.39 to 0.85)	
P value	0.0055	
BILAG New 1A or 2B		
n (%)	86 (30)	54 (19)
OR (95% CI) (belimumab versus placebo)	0.58 (0.41 to 0.81)	
P value	0.0016	

BILAG = British Isles Lupus Assessment Group; CI = confidence interval; HR = hazard ratio; OR = odds ratio; PCS = physical component summary; PGA = Physician Global Assessment; SE = standard error; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SF-36 = Short Form (36) Health Survey; SFI = System Lupus Erythematosus Flare Index; SRI = Systemic Lupus Erythematosus Responder Index.

Source: Clinical Study Report for BLISS 52;[27] Navarra et al. (2011).[20]

Table 20 provides efficacy results for BLISS 76. The primary outcome of proportion of SRI responders at week 52 was achieved by statistically more patients in the belimumab 10 mg/kg group compared with placebo. However, the SRI response rate at week 76, a major secondary outcome, was not statistically significant. Similarly, more patients achieved at least a four-point reduction in SELENA-SLEDAI score at week 52, but the result was not statistically significant at week 76. Other major secondary outcomes (i.e., change in PGA score at week 24, reduction in prednisone dosage of 25% or more to 7.5 mg per day during weeks 40 to 52, and SF-36 PCS at week 24) were also not statistically significant between placebo and belimumab. In addition, there was no difference between placebo and belimumab in the proportion of patients experiencing severe flares, no worsening by BILAG, or no worsening by PGA score, at week 76.

Table 20: Efficacy in BLISS 76

	BLISS 76	
	Placebo (N = 275)	Belimumab 10 mg/kg (N = 273)
Primary outcome		
Proportion achieving SRI response at week 52	92 (33.5)	118 (43.2)
P value	0.017	
Major secondary outcomes		
Proportion achieving SRI response at week 76	89 (32.4)	105 (38.5)
P value	0.13	
≥ 4-point reduction in SELENA- SLEDAI at week 52, n (%)	97 (35.3)	127 (46.5)
P value	0.006	
≥ 4-point reduction in SELENA- SLEDAI at week 76, n (%)	93 (33.8)	113 (41.4)
P value	NS	
PGA score at week 24		
Mean change	–0.49	–0.44
P value	NS	
Prednisone dosage reduced ≥ 25% to ≤ 7.5 mg/day during weeks 40 to 52	16/126 (12.7)	21/120 (17.5)
P value	NS	
SF-36 PCS score at week 24		
Mean change	3.35	3.21
P value	NS	
Other secondary outcomes		
Disease flares (severe)		
With flare (SFI), n (%)	73 (26.5)	56 (20.5)
P value	NS	
No worsening by BILAG at week 76, n (%)	162 (58.9)	173 (63.4)
P value	NS	
No worsening PGA at week 76, n (%)	160 (58.2)	172 (63.0)
P value	NS	

BILAG = British Isles Lupus Assessment Group; NS = not significant; PCS = physical component summary; PGA = Physician Global Assessment; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SF-36 = Short Form (36) Health Survey; SFI = System Lupus Erythematosus Flare Index; SRI = Systemic Lupus Erythematosus Responder Index.

Source: Clinical Study Report for BLISS 76;[28] Furie et al. (2011).[21]

Harms

Table 21 shows the harms data for BLISS 52 and BLISS 76. Most patients in both studies experienced at least one AE. WDAEs were similar between the groups. SAEs and infections were slightly higher in the belimumab groups of both studies. Infusion-related reactions were higher in the belimumab group of BLISS 76, but similar to placebo in BLISS 52. No malignant neoplasms occurred in BLISS 52; however, in BLISS 76 there were two cases in the belimumab group and one in placebo. In BLISS 52, there were four deaths in

the belimumab group and three deaths in the placebo group. In BLISS 76, there was one death in a patient who received belimumab and no deaths in patients who received placebo.

In BLISS 76, depression was more frequent in the belimumab group compared with placebo (6% to 7% versus 4%).[21] No suicides or suicides attempts occurred in any of the treatment groups of BLISS 76.[21]

Table 21: Harms in BLISS 52 and BLISS 76

	BLISS 52		BLISS 76	
	Placebo (N = 287)	Belimumab 10 mg/kg (N = 290)	Placebo (N = 275)	Belimumab 10 mg/kg (N = 273)
≥ 1 AE, n (%)	263 (92)	266 (92)	253 (92)	253 (93)
≥ 1 SAE, n (%)	36 (13)	41 (14)	54 (20)	61 (22)
WDAEs, n (%)	19 (7)	15 (5)	23 (8.4)	23 (8.4)
Deaths, n (%)	3 (1)	4 (1)	0 (0)	1 (0.4)
Malignant neoplasm, n (%)	0 (0)	0 (0)	1 (0.4)	2 (0.7)
Infections, n (%)	183 (64)	194 (67)	190 (69)	202 (74)
Infusion reactions, n (%)	49 (17)	48 (17)	27 (10)	37 (14)

AE = adverse event; SAE = serious adverse event; WDAE = withdrawal due to adverse event.

Source: Clinical Study Report for BLISS 52;[27] Navarra et al. (2011);[20] Clinical Study Report for BLISS 76;[28] Furie et al. (2011).[21]

Critical Appraisal of BLISS 52 and BLISS 76

There were notable between-trial differences in patient populations, in terms of geographic and ethnic origin, duration of disease, and baseline medication use. These differences may have accounted for the differences in the results of the studies; the proportion of responders was statistically significantly higher for belimumab compared with placebo at 52 weeks but not at 76 weeks (in BLISS 76). Furthermore, there were no between-treatment differences in outcomes that were important to patients, such as quality of life and reduction in prednisone dosage. As with BLISS SC, trials of IV belimumab excluded patients with severe active lupus nephritis and/or CNS lupus.

B. Open-Label Extensions of IV Belimumab

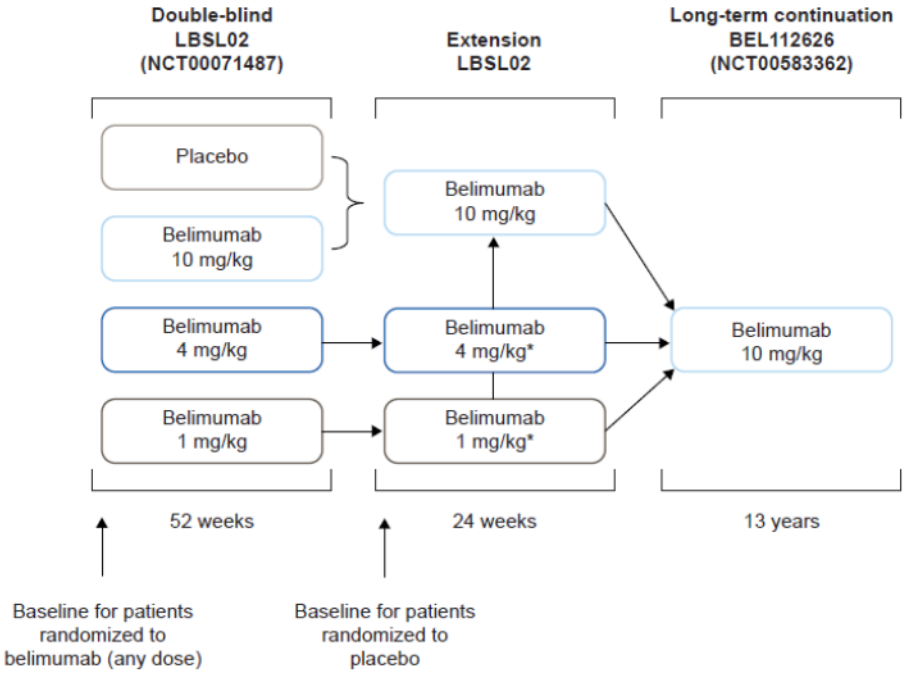
Phase II (LBSL99)

Study and Phase Design

The parent study of LBSL99 (LBSL02) was a multi-centre, DB RCT, 52-week dose-ranging study followed by a 24-week extension period, in adults with SLE (N = 449).[29] Patients were randomized to IV belimumab 1 mg/kg, 4 mg/kg, 10 mg/kg, or placebo, all administered with standard of care. LBSL99 followed patients for an additional 13 years after completion of LBSL02, for the primary objective of monitoring long-term safety of belimumab administration (Figure 8).[29] LBSL99 included patients who completed LBSL02 (52-week + 24-week extension), who tolerated belimumab, had a satisfactory response (improvement in PGA score from baseline), experienced no severe flare (as defined by the SFI) in the last 30 days, and who wished to continue on treatment. Patients were excluded if they received other biologics, IV cyclophosphamide, or corticosteroids of more than 100 mg per day

prednisone equivalent for reasons other than severe SLE flare. Patients received IV belimumab 10 mg/kg every 28 days, starting four weeks after the last dose in LBSL02.

Figure 8: Study Design of Phase II Open-Label Extension (LSBL99)



Source: Reprinted from Wallace et al. 2019 Supplementary Figure 1, under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).[29]

Safety was assessed from the first dose to 24 weeks after the final dose. Efficacy was assessed every 16 weeks with the SRI response, SELENA-SLEDAI score, BILAG, PGA score (every eight weeks), SFI, and corticosteroid use (every four weeks). Low disease activity was defined as SELENA-SLEDAI score of 2 or less and prednisone dosage of 5 mg per day or less. Analyses were conducted on patients who received at least one dose of belimumab in LSBL99. Baseline was defined as before the first dose of belimumab, which was in either the parent study or, for patients who had received placebo, the 24-week extension.

Results

A total of 298 patients (62.6% of patients randomized to LSBL02 and 92.8% of patients who completed the 24-week extension) were enrolled in LBSL99, of whom 296 received at least one dose of study drug. Ninety-six (32.2%) patients remained in the study until the end (13 years). Primary reasons for withdrawal from the study were decision by patient (28%), AE (22%), non-compliance with drug (14%), lack of efficacy (11%), and physician decision (9%).

Table 22 shows the baseline characteristics of the 296 patients who received at least one dose of belimumab. The mean age of patients was 43 years, and most were female and white. The mean duration of SLE was about nine years. Most patients had ANA titre of 1:80 or more and half had anti-dsDNA of 30 IU/mL or more. Nearly 36% had SELENA-SLEDAI

score of 10 points or more, and most had BILAG 1A or 2B score. Most patients were taking prednisone, with about a third taking more than 7.5 mg per day.

Table 22: Baseline Characteristics in LBSL99

Characteristic	LBSL99
	Belimumab IV (N = 296)
Age, years, mean (SD)	43.0 (11.6)
Female, n (%)	276 (93.2)
Ethnic origin, n (%)	
White	215 (72.6)
Black American	68 (23.0)
Other	13 (4.4)
Disease duration, years, mean (SD)	9.1 (7.8)
SELENA-SLEDAI score, mean (SD)	8.4 (4.7)
SELENA-SLEDAI score ≥ 10, n (%)	106 (35.8)
BILAG 1A or 2B score, n (%)	168 (56.8)
PGA score, mean (SD)	1.3 (0.6)
≥ 1 SFI flare, n (%)[a]	47 (15.9)
≥ 1 severe SFI flare, n (%)[a]	9 (3.0)
ANA ≥ 1:80, n (%)[b]	208 (81.3)
Anti-dsDNA ≥ 30 IU/mL, n (%)	149 (50.3)
Prednisone, n (%)	191 (64.5)
Prednisone > 7.5 mg/day, n (%)	92 (31.1)

ANA = antinuclear antibody; anti-dsDNA = anti–double-stranded DNA; BILAG = British Isles Lupus Assessment Group; PGA = Physician Global Assessment; SD = standard deviation; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SFI = System Lupus Erythematosus Flare Index.

[a] Any time before study entry for patients who received belimumab in the DB phase. Between the last visit in the DB phase and first dose of belimumab in the open-label phase for patients who received placebo in the DB phase.

[b] Data available for 256 patients.

Source: Reprinted from Wallace et al. 2019 Table 1, under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0), with changes made.[29]

Harms

Table 23 shows the harms in LSBL99, for the 13-year period overall and by each year. Almost all patients (99.3%) experienced an AE. The most frequent AEs (≥ 15/100 patient-years) were arthralgia (29.3/100 patient-years), upper respiratory tract infection (29.0/100 patient-years), sinusitis (16.9/100 patient-years), urinary tract infection (16.2/100 patient-years), and headache (15.0/100 patient-years). Forty-four patients (14.9%) discontinued drug or withdrew from the study due to an AE. SAEs were experienced by 61.5% of patients. The most frequent SAEs (≥ 0.5 events/100 patient-years) were pneumonia (0.9/100 patient-years), osteoarthritis (0.8/100 patient-years), noncardiac chest pain (0.7/100 patient-years), pyrexia (0.6/100 patient-years), cellulitis, chronic obstructive pulmonary disease, abdominal pain, viral gastroenteritis, and vomiting (all 0.5/100 patient-years). Serious infections or infestations occurred in 23.3% of patients and infections of special interest in 25.3%. Malignant neoplasms occurred in 4.7% patients. Depression was present in 136 patients (45.9%) and suicide or self-injury in six patients. One of these

suicide attempts resulted in death. A total of eight deaths occurred, with the reasons provided in Table 23.

Table 23: Harms in LBSL99

	LBSL99												
	Year Number of patients												
	All 296	**1 296**	**2 294**	**3 276**	**4 250**	**5 223**	**6 209**	**7 192**	**8 178**	**9 169**	**10 152**	**11 131**	**11+ 88**
≥ 1 AE, n (%)	294 (99.3)	291 (98.3)	283 (96.3)	260 (94.2)	239 (95.6)	203 (91.0)	190 (90.9)	182 (94.8)	162 (91.0)	157 (92.9)	137 (90.1)	105 (80.2)	45 (51.1)
WDAE, n (%)	44 (14.9)	2 (0.7)	3(1.0)	3 (1.1)	7 (2.8)	5 (2.2)	6 (2.9)	6 (3.1)	1 (0.6)	3 (1.8)	5 (3.3)	2 (1.5)	0 (0)
≥ 1 SAE, n (%)	182 (61.5)	41 (13.9)	43 (14.6)	50 (18.1)	30 (12.0)	40 (17.9)	33 (15.8)	35 (18.2)	34 (19.1)	28 (16.6)	25 (16.4)	14 (10.7)	7 (8.0)
Serious infections/ infestations, n (%)	69 (23.3)	11 (3.7)	13 (4.4)	9 (3.3)	9 (3.6)	6 (2.7)	6 (2.9)	12 (6.3)	10 (5.6)	8 (4.7)	5 (3.3)	5 (3.8)	4 (4.5)
Infections of special interest, n (%)[a]	75 (25.3)	11 (3.7)	13 (4.4)	7 (2.5)	11 (4.4)	8 (3.6)	6 (2.9)	8 (4.2)	9 (5.1)	4 (2.4)	8 (5.3)	5 (3.8)	3 (3.4)
Malignant neoplasms, n (%)[b]	14 (4.7)	0 (0)	0 (0)	1 (0.4)	1 (0.4)	4 (1.8)	1 (0.5)	1 (0.5)	0 (0)	2 (1.2)	3 (2.0)	0 (0)	0 (0)
Depression, n (%)	136 (45.9)	39 (13.2)	31 (10.5)	17 (6.2)	28 (11.2)	16 (7.2)	10 (4.8)	20 (10.4)	14 (7.9)	10 (5.9)	9 (5.9)	3 (2.3)	0 (0)
Suicide/self-injury, n (%)	6 (2.0)	1 (0.3)	1 (0.3)	1 (0.4)	0 (0)	2 (0.9)	1 (0.5)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Death, n (%)[c]	8 (2.7)	1 (0.3)	0 (0)	1 (0.4)	1 (0.4)	0 (0)	1 (0.5)	2 (1.0)	0 (0)	0 (0)	1 (0.7)	0 (0)	0 (0)

AE = adverse event; SAE = serious adverse event; WDAE = withdrawal due to adverse event.

[a] Infections of special interest are opportunistic infections, tuberculosis, herpes zoster, and sepsis.

[b] Excluding nonmelanoma skin cancer.

[c] Cause of deaths: Year 1 – coronary artery disease; Year 3 – suicide; Year 4 – pneumonia; Year 6 – cardiac arrest; Year 7 – acute respiratory distress syndrome and respiratory failure; Year 10 – retroperitoneal hemorrhage. One patient died of pneumonia during follow-up.

Source: Reprinted from Wallace et al. 2019 Supplementary Table 2, under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).[29]

Efficacy

The study demonstrated improvement in several efficacy outcomes over time; however, these data are of limited usefulness given that, with each year, fewer patients remained in the study and there was no comparator. Of 190 patients who received corticosteroids at baseline, 25 patients (13.2%) discontinued corticosteroids for the remainder of the study. Patients were followed for eight and 24 weeks after withdrawal from the study. The efficacy outcomes for the eight-week and 24-week follow-up times are provided in Table 24, and Figure 9 shows the SRI response over time. Outcomes remained stable from the eight-week to 24-week assessments, although fewer patients were assessed at 24 weeks.

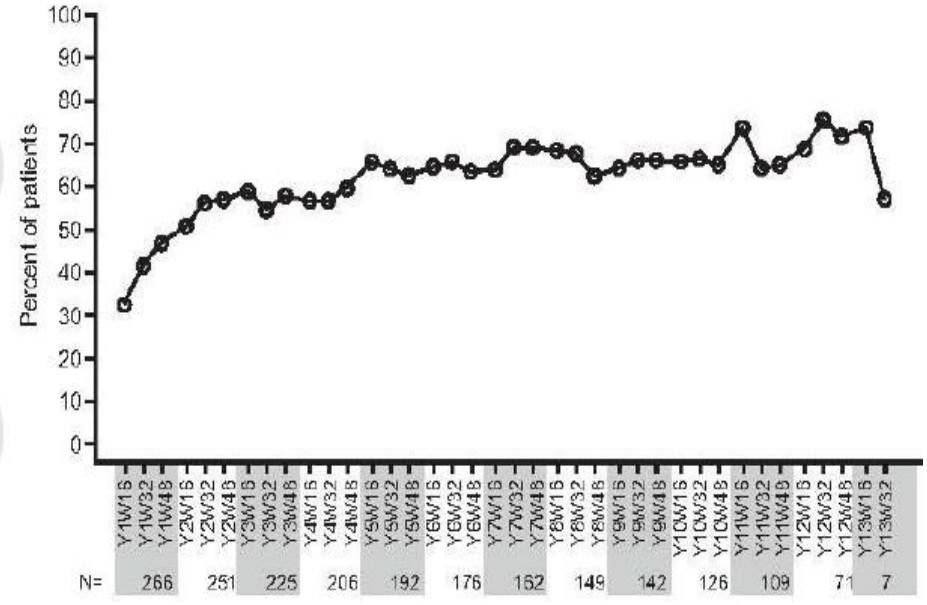
Table 24: Disease Activity During Follow-Up in LBSL99

	LBSL99	
	Week 8	**Week 24**
SRI responders, n (%)	122/197 (61.9)	114/178 (64.0)
≥ 4-point increase in SELENA-SLEDAI score from baseline	137/220 (62.3)	120/199 (60.3)
PGA score, n (%)		
0 to 1	158/216 (73.1)	148/197 (75.1)
> 1 to 2.5	57/216 (26.4)	49/197 (24.9)
> 2.5	1/216 (0.5)	0 (0)
Low disease activity, n (%)	79/220 (35.9)	77/199 (38.7)
SFI flare	45/219 (20.5)	42/199 (21.1)
Severe SFI flare	6/219 (2.7)	6/199 (3.0)
Reduction in prednisone from ≥ 7.5 mg/day to < 7.5 mg/day, n (%)	23/70 (32.9)	18/63 (28.6)
Increase in prednisone from ≤ 7.5 mg/day to > 7.5 mg/day, n (%)	26/152 (17.1)	23/137 (16.8)

PGA = Physician Global Assessment; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SFI = System Lupus Erythematosus Flare Index; SRI = Systemic Lupus Erythematosus Responder Index.

Source: Reprinted from Wallace et al. 2019 Supplementary Table 5, under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).[29]

Figure 9: SRI Response Over Time in LBSL99



SRI = Systemic Lupus Erythematosus Responder Index.

Source: Reprinted from Wallace et al. 2019 Supplementary Figure 3, panel A under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).[29]

Phase III (BLISS 76 Extension US)

Study and Phase Design

This extension study was conducted in patients in the US who were enrolled in BLISS 76.[2,30] Patients were eligible to enrol in BLISS 76 Extension if they completed 72-week BLISS 76. The same dose of belimumab as the randomized phase was administered every 28 days plus standard of care. Patients who had been randomized to placebo received belimumab 10 mg/kg. After a protocol amendment on March 9, 2011, patients who were receiving belimumab 1 mg/kg were switched to 10 mg/kg. For analyses, data for all patients who received belimumab were pooled.

The primary objective of BLISS 76 Extension was to evaluate long-term safety, and it included eight years of data (study is completed). Other exploratory outcomes were SDI, SRI, SELENA-SLEDAI score, BILAG, PGA score, flare rates (modified SFI and BILAG), and prednisone use. Assessments were performed at week 24 and then every 48 weeks. Adverse events were monitored throughout the study and for eight weeks following the last dose. Organ damage was assessed every 48 weeks. Health-related quality of life, as measured with the SF-36 and FACIT, have also been reported.[31]

Results

Of 826 patients randomized in the parent study, 268 (32%) continued to the extension. Of these, 128 (47.8%) discontinued the study. Primary reasons for withdrawal were patient decision (24.2%), AEs (19.5%), other (17.2%), physician decision (13.3%), lack of efficacy (10.9%), and loss to follow-up (9.4%). The mean duration of exposure to belimumab was approximately five years (range 0.08 to 7.6 years).

Most patients were white and female, with an average age of approximately 43 years at baseline (Table 25). At baseline, the mean duration of disease was 7.7 years, and approximately one-third of patients had SELENA-SLEDAI scores of 10 or more, while slightly more than half had BILAG organ domain involvement. The mean SDI score was 1.2 points. About 44% had low complement levels. About a third of patients were on a prednisone dosage of 7.5 mg per day or less.

Table 25: Baseline Characteristics in BLISS 76 Extension (Patients From the US)

Characteristic	BLISS 76 Extension (US)
	Belimumab IV (N = 268)
Age, years, mean (SD)	42.8 (11.3)
Female, n (%)	250 (93.3)
Ethnic origin, n (%)	
White	186 (69.4)
Black American	57 (21.3)
Asian	13 (4.9)
American Indian or Alaska Native	8 (3.0)
Mixed	4 (1.5)
Disease duration, years, mean (SD)	7.7 (6.8)
SELENA-SLEDAI score, mean (SD)	7.8 (3.9)

Characteristic	BLISS 76 Extension (US)
SELENA-SLEDAI score ≥ 10, n (%)	80 (29.9)
BILAG organ domain involvement, n (%)	137 (51.1)
PGA score, mean (SD)	1.2 (0.6)
SDI score (SD)	1.2 (1.5)
≥ 1 SFI flare, n (%)[a]	65 (24.3)
≥ 1 severe SFI flare, n (%)[a]	2 (0.7)
Low C3 and/or low C4	119 (44.4)
Corticosteroids only, n (%)	21 (7.8)
Prednisone ≤ 7.5 mg/day	87 (32.5)
Immunosuppressants only, n (%)	11 (4.1)
Antimalarial drugs only, n (%)	41 (15.3)

BILAG = British Isles Lupus Assessment Group; C = complement; PGA = Physician's Global Assessment; SD = standard deviation; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SFI = System Lupus Erythematosus Flare Index.

[a] At baseline, defined as the last available value before the initiation of belimumab. For patients randomized to belimumab in the DB phase, baseline was taken from the parent study. For patients randomized to placebo in the DB phase, baseline was the last value available from the parent study.

Source: Clinical Study Report for BLISS 76 Extension (US).[2]

Harms

Table 26 shows the harms in BLISS 76 Extension in US patients, overall and by year. Nearly all patients experienced at least one AE (99.6%). SAEs were experienced by about 42%. Serious infections and infections of special interest were about 16% each. Malignant neoplasms occurred in almost 4% and depression, suicide, or self-injury in 27%. There were three suicide or self-injury events.[2] A total of two deaths occurred, one due to hypertensive heart disease and the other to polydrug toxicity, which was later adjudicated as a suicide.

Table 26: Harms in BLISS 76 Extension (Patients From the US)

	BLISS 76 Extension (US)								
	Year								
	All N = 268	1 N = 268	2 N = 259	3 N = 244	4 N = 219	5 N = 202	6 N = 192	7 N = 130	7+ N = 65
≥ 1 AE, n (%)	267 (99.6)	260 (97.0)	235 (90.7)	206 (84.4)	184 (84.0)	167 (82.7)	145 (75.5)	87 (66.9)	31 (47.7)
WDAE, n (%)	26 (9.7)	3 (1.1)	4 (1.5)	7 (2.9)	8 (3.7)	0 (0)	2 (1.0)	2 (1.5)	0 (0)
≥ 1 SAE, n (%)	112 (41.8)	33 (12.3)	30 (11.6)	25 (10.2)	22 (10.0)	24 (11.9)	16 (8.3)	13 (10.0)	3 (4.6)
Serious infections/ infestations, n (%)	44 (16.4)	13 (4.9)	9 (3.5)	4 (1.6)	6 (2.7)	8 (4.0)	7 (3.6)	2 (1.5)	1 (1.5)
Infections of special interest, n (%)	43 (16.0)	14 (5.2)	13 (5.0)	8 (3.3)	6 (2.7)	7 (3.5)	11 (5.7)	6 (4.6)	0 (0)
Malignant neoplasms, n (%)[a]	10 (3.7)	0 (0)	1 (0.4)	4 (1.6)	2 (0.9)	0 (0)	0 (0)	2 (1.5)	1 (1.5)
Depression/suicide/self-injury, n (%)	73 (27.2)	25 (9.3)	22 (8.5)	17 (7.0)	6 (2.7)	8 (4.0)	3 (1.6)	4 (3.1)	1 (1.5)

	BLISS 76 Extension (US)								
Death, n (%)	2 (0.7)	0 (0)	1 (0.4)	0 (0)	1 (0.5)	0 (0)	0 (0)	0 (0)	0 (0)

AE = adverse event; SAE = serious adverse event; WDAE = withdrawal due to adverse event.

[a] Excluding nonmelanoma skin cancer.

Source: Clinical Study Report for BLISS 76 Extension (US);[2] Furie et al. (2018).[30]

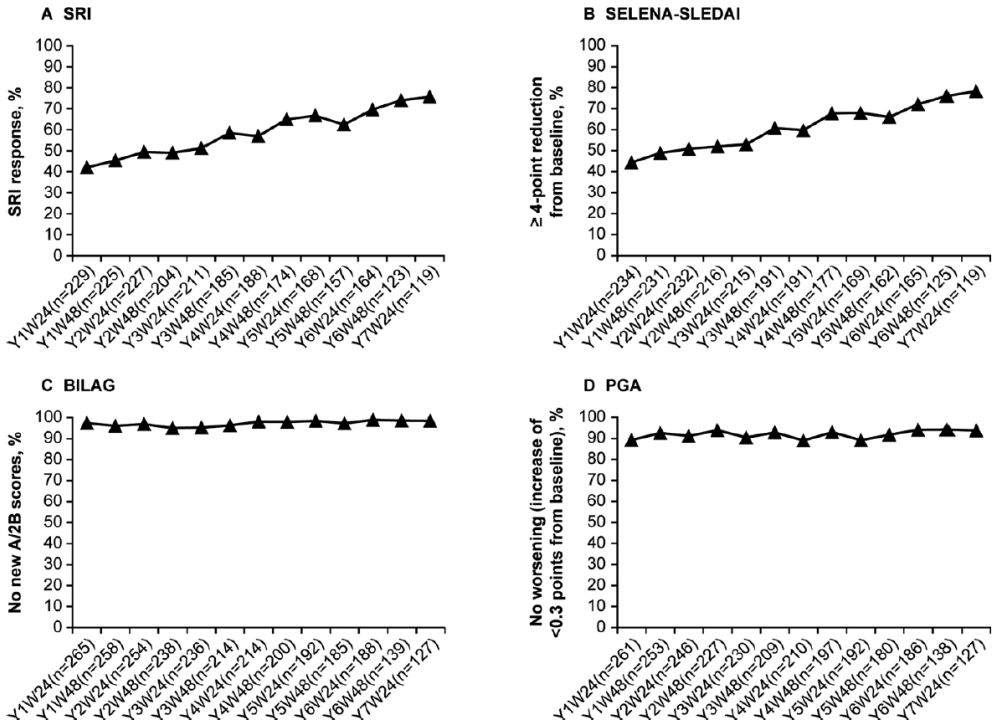
Efficacy

Figure 10 shows data for SRI, SELENA-SLEDAI score, BILAG score, and PGA score over time in the BLISS 76 Extension in US patients.[30] At the seven-year midpoint of the 13-year study, SRI response had been achieved by 90 of 119 (75.6%) patients (panel A). A reduction from baseline in the SELENA-SLEDAI score of four points or more was achieved by 93 of 119 (78.2%) patients at the seven-year midpoint (panel B). While panels A and B show that the percentage of patients achieving SRI and SELENA-SLEDAI score of four points or more increased with time, this must be interpreted with caution because there were fewer patients remaining in the study after each year, and patients who discontinued may have had a worse response than patients who remained in the study. The percentage of patients (among those were still in the study, which was fewer at each time point) with no new BILAG A or no new 2 BILAG B organ domain scores was stable throughout the study (panel C). The percentage of patients with no new worsening in PGA score (increase of < 0.3 points) was 93.7% (119 or 127 patients) at the seven-year midpoint and was stable overall during the study (panel D).

Figure 11 shows the mean change in prednisone dosage.[30] At the seven-year midpoint, among 77 patients who were using prednisone, the mean decrease in dosage was 31.4%. A baseline dosage of more than 7.5 mg per day was reduced to 7.5 mg per day or less in 15 of 38 patients (39.5%) at year 7.[30] Figure 11 shows that the mean prednisone dosage decreased with time. However, as with the SRI and SELENA-SLEDAI responses, this must be interpreted with caution due to the omission of patients after each year of the study.

Figure 12 shows the SF-36 (panel A) and FACIT-Fatigue (panel B) mean changes from baseline up to year 6.[31] Among 185 patients remaining in the study at year 6, the mean change (SD) from baseline in the PCS was 4.8 (9.4) points and the change in the mental component summary (MCS) was 2.7 (11.3) points. For the FACIT-Fatigue, the mean change (SD) at year 6 was 3.7 (11.8) points. Table 27 shows the percentage of patients who exceeded the sponsor-defined minimal important difference (MID) for the SF-36 domains (≥ 5 points) and the sponsor-defined MID for the FACIT-Fatigue (≥ 4 points) for each year. Nearly half exceeded the FACIT-Fatigue MID at year 6. Fifty percent or greater exceeded the MID for the SF-36 domains of bodily pain, general health, mental health, physical functioning, role physical, and vitality at year 6. Again, these data must be interpreted with caution because of the smaller number of patients in the study with each passing year.

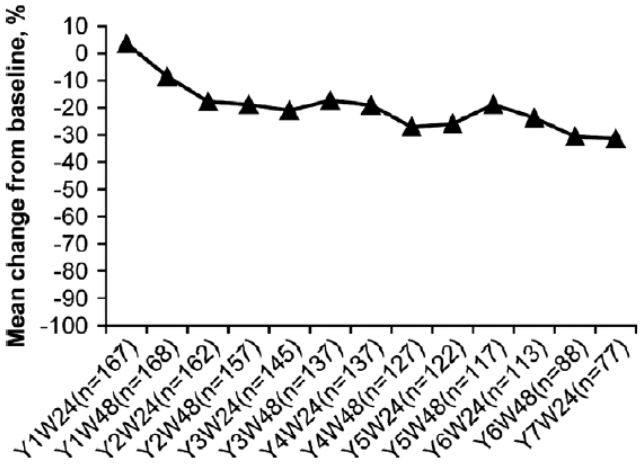
Figure 10: Efficacy Outcomes in BLISS 76 Extension (Patients From the US)



W = week; Y = year.

Source: Reprinted from Furie 2018 Figure 2, under licence Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).[30]

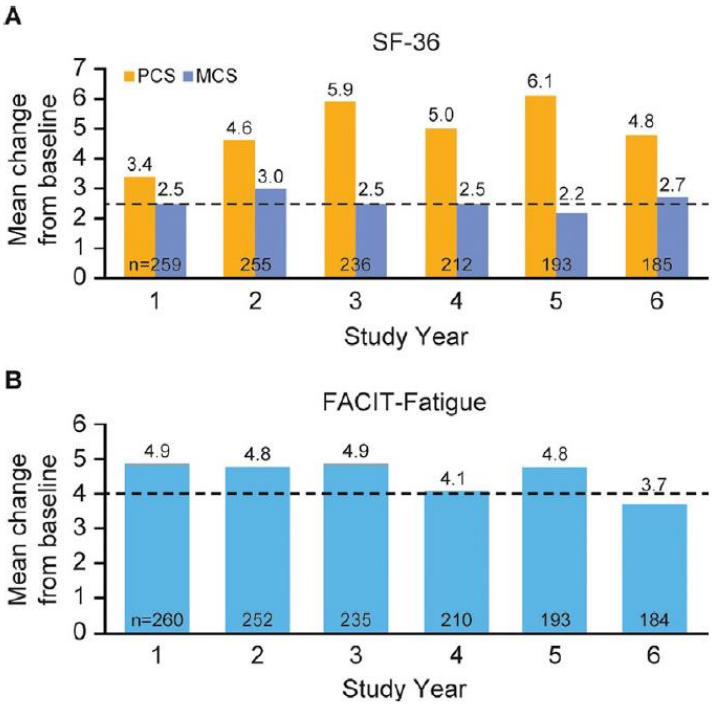
Figure 11: Change in Prednisone Dosage in BLISS 76 Extension (Patients From the US)



W = week; Y = year.

Source: Reprinted from Furie et al. (2018) Figure 3, under licence Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).[30]

Figure 12: Health-Related Quality of Life in BLISS-76 Extension (Patients From the US)



FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy–Fatigue Scale; SF-36 = Short Form (36) Health Survey;

Source: Reprinted from Strand et al. (2019) Figure 1, under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).[31]

Table 27: Improvements Greater Than Minimal Important Difference for SF-36 Domains and FACIT-Fatigue

	BLISS 76 Extension (US)					
	Year					
	1 N = 268	2 N = 259	3 N = 244	4 N = 219	5 N = 202	6 N = 192
SF-36, n (%)[a]						
Bodily pain	148 (55.2)	137 (52.9)	135 (55.3)	117 (53.4)	109 (54.0)	100 (52.1)
General health	151 (56.3)	148 (57.1)	143 (58.6)	127 (58.0)	126 (62.4)	100 (52.1)
Mental health	139 (51.9)	133 (51.4)	116 (47.5)	102 (46.6)	100 (49.5)	97 (50.5)
Physical functioning	141 (52.6)	129 (49.8)	124 (50.8)	112 (51.1)	105 (52.0)	100 (52.1)
Role emotional	111 (41.4)	114 (44.0)	98 (40.2)	91 (41.6)	82 (40.6)	72 (37.5)
Role physical	142 (53.0)	139 (53.7)	131 (53.7)	113 (51.6)	109 (54.0)	96 (50.0)
Social functioning	127 (47.4)	125 (48.3)	124 (50.8)	103 (47.0)	93 (46.0)	88 (45.8)
Vitality	152 (56.7)	144 (55.6)	134 (54.9)	122 (55.7)	112 (55.4)	98 (51.0)
FACIT-Fatigue, n (%)[b]	135 (50.4)	123 (47.5)	119 (48.8)	97 (44.3)	108 (53.5)	89 (46.4)

FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy–Fatigue; SF-36 = Short Form (36) Health Survey.

[a] MID for SF-36 domains was five points or greater.

[b] MID for FACIT-Fatigue was four points or greater.

Source: Reprinted from Strand et al. (2019) Table 2, under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0), with changes made.[31]

Phase III (BLISS 52 and BLISS 76 Extension Non-US)

Patients from outside of the US who completed the 52-week randomized BLISS 52 or the 76-week randomized BLISS 76 were eligible to enrol in this long-term extension.[32,33] Patients continued to receive the same dose of belimumab as in the randomized phase every 28 days plus standard of care. Patients who had been randomized to placebo received belimumab 10 mg/kg. A July 2011 amendment switched patients who were receiving belimumab 1 mg/kg to 10 mg/kg. In addition, a Mexico National Amendment allowed for the enrolment of five patients from the 52-week BLISS SC study. Baseline was defined as the last available value before the start of belimumab treatment. Patients were followed for eight years (study is completed).

Results

A total of 738 patients from outside of the US were enrolled in the open-label extension, and 735 (99.6%) received at least one dose of belimumab. The study was completed by 368 patients (49.9%). Common reasons for withdrawal were patient decision (40.8%), AEs (18.6%), and other (18.6%). The mean duration of exposure was 4.5 years (range 0.08 to 8.8 years). Table 28 provides the baseline characteristics of patients in the non-US BLISS 52 and BLISS 76 Extension who received at least one dose of belimumab.

Table 28: Baseline Characteristics in BLISS 52 and 76 Extension (Non-US Patients)

Characteristic	BLISS 52 and 76 extension (non-US)
	Belimumab IV (N = 735)
Age, years, mean (SD)	37.2 (11.2)
Female, n (%)	695 (94.6)
Ethnic origin, n (%)	
White	278 (37.8)
Native American or Alaska Native	225 (30.6)
Asian	214 (29.1)
Black or African-American/African heritage	18 (2.4)
Mixed	2 (0.3)
Disease duration, years, median (range)	4.5 (0 – 37)
SELENA-SLEDAI score, mean (SD)	8.3 (4.3)
SELENA-SLEDAI score ≥ 10, n (%)	284 (38.6)
BILAG, n (%)	
≥ 1A or 2B	324 (44.1)
≥ 1A	107 (14.6)
≥ 2B	531 (72.2)
No A or B	171 (23.3)
PGA score, mean (SD)	1.2 (0.6)
SDI score (SD)	0.6 (1.0)
≥ 1 SFI flare, n (%)[a]	107 (14.6)
≥ 1 severe SFI flare, n (%)[a]	4 (0.5)
Anti-dsDNA ≥ 30 IU/mL, n (%)	528 (71.8)

Characteristic	BLISS 52 and 76 extension (non-US)
ANA ≥ 1:80, n (%)	683 (93.6)
Low C3 and/or low C4	488 (66.4)
Corticosteroids, n (%)	690 (93.9)
Prednisone, n (%)	
≤ 7.5 mg/day	227 (30.9)
> 7.5 to ≤ 40 mg/day	462 (62.9)
> 40 mg/day	1 (0.1)
Immunosuppressants, n (%)	323 (43.9)
Antimalarial drugs, n (%)	474 (64.5)

ANA = antinuclear antibody; anti-dsDNA = anti–double-stranded DNA; BILAG = British Isles Lupus Assessment Group; PGA = Physician Global Assessment; SD = standard deviation; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SFI = System Lupus Erythematosus Flare Index.

[a] Assessed during screening and day 0 of the parent study for patients randomized to belimumab in the DB phase and between the last two visits in the parent study for patients randomized to placebo in the DB phase.

Source: Clinical Study Report for BLISS 52 and BLISS 76 Extension (non-US);[32] van Vollenhoven (2019).[33]

Harms

Table 29 provides data on harms in the open-label extension overall and by study year. Almost all patients experienced at least one AE (96.1%). The most frequent AEs were headache (28%), nasopharyngitis (21%), diarrhea (20%), arthralgia (19%), and influenza (18%).[32] Slightly more than a third experienced an SAE (31.4%), and 9.4% discontinued the drug due to an AE. The most common SAEs were bacterial pneumonia (1.9%), cellulitis (1.6%), and lupus nephritis (1.6%).[32] Serious infections or infestations occurred in 31 patients. Ninety-five patients (12.9%) developed an infection of special interest (i.e., opportunistic infections, tuberculosis, herpes zoster, and sepsis).[33] Malignant neoplasms occurred in a total of six patients. Depression was experienced by 3.5% and suicide/self-injury by three patients. Eleven deaths occurred over the study period, and two additional deaths occurred more than eight weeks after the last belimumab dose. Causes of death were ischemic stroke (n = 1), cardiac arrest (n = 1), cardiogenic shock (n = 1), pancreatitis (n = 1), thrombotic thrombocytopenic purpura (n = 1), pulmonary hemorrhage (n = 1), intracranial hemorrhage (n = 1), respiratory distress (n = 1), pneumonia (n = 2), sepsis (n = 2), septic shock (n = 1).[32]

Table 29: Harms in BLISS 52 and BLISS 76 Extension (Non-US Patients)

	BLISS 52 and 76 extension (non-US)								
	Year								
	All N = 735	1 N = 735	2 N = 701	3 N = 620	4 N = 514	5 N = 442	6 N = 345	7 N = 219	8 N = 65
≥ 1 AE, n (%)	706 (96.1)	617 (83.9)	502 (71.6)	441 (71.1)	344 (66.9)	261 (59.0)	181 (52.5)	92 (42.0)	26 (40.0)
WDAE, n (%)	69 (9.4)	13 (1.8)	13 (1.9)	20 (3.2)	10 (1.9)	7 (1.6)	5 (1.4)	0 (0)	0 (0)
≥ 1 SAE, n (%)	231 (31.4)	78 (10.6)	58 (8.3)	66 (10.6)	44 (8.6)	27 (6.1)	16 (4.6)	11 (5.0)	1 (1.5)
Serious infections/ infestations, n (%)	31 (0.9)	11 (1.5)	3 (0.5)	5 (0.9)	5 (1.0)	2 (0.5)	1 (0.3)	0 (0)	0 (0)
Malignant neoplasms, n (%)[a]	6 (0.8)	1 (0.1)	1 (0.1)	1 (0.2)	2 (0.4)	1 (0.2)	0 (0)	0 (0)	0 (0)

	BLISS 52 and 76 extension (non-US)								
Depression, n (%)	117 (3.5)	48 (6.7)	27 (4.1)	14 (2.5)	14 (2.9)	4 (1.0)	6 (2.0)	4 (2.7)	0 (0)
Suicide/self-injury, n (%)	3 (< 0.1)	0 (0)	0 (0)	1 (0.2)	1 (0.2)	0 (0)	0 (0)	1 (0.7)	0 (0)
Death, n (%)[b]	11 (1.5)	2 (0.3)	1 (0.1)	2 (0.3)	2 (0.4)	2 (0.5)	0 (0)	0 (0)	0 (0)

AE = adverse event; SAE = serious adverse event; WDAE = withdrawal due to adverse event.

[a] Excluding nonmelanoma skin cancer.

[b] An additional two deaths occurred eight weeks after the last dose of belimumab.

Source: Clinical Study Report for BLISS 52 and BLISS 76 Extension (non-US).[32]

Efficacy

At year 8, mean change (SD) in SDI score was 0.2 (0.6) points.[33] By organ involvement, only the musculoskeletal domain demonstrated a change from baseline to year 8 (increase of 0.1 [0.24] points). Among patients on a prednisone dosage of more than 7.5 mg per day at baseline, 59 of 87 patients (67.8%) experienced a reduction in dosage at year 7.

Phase III (Pooled US and Non-US Extensions of BLISS 52 and BLISS 76)

Study and Phase Design

This study pooled data for patients from the US and outside of the US, who completed the open-label extension of BLISS 52 and BLISS 76.[34,35] As with the non-US studies, patients continued on the same dose of belimumab, as received in the randomized phase, every 28 days plus standard of care. Patients who had received placebo were switched to belimumab 10 mg/kg. After study protocol amendments in March and July 2011, patients who were receiving belimumab 1 mg/kg were switched to 10 mg/kg. For the pooled analysis, both belimumab dose groups were combined. The primary end point was change in SDI score from baseline to year 5 to 6. Baseline was defined as the last assessment before receiving belimumab. A key secondary outcome was time to first SDI worsening. Other outcomes were change from baseline in SDI for subgroups (baseline SDI 0 or ≥ 1, baseline SELENA-SLEDAI score ≤ 9 or ≥ 10), total SDI score, change from baseline in SDI organ damage system, and harms. The SDI score was assessed every 48 weeks, and AEs were collected throughout the study. Analyses were performed on all patients who were enrolled into the US or non-US extension and who received at least one dose of study drug.

Results

A total of 998 patients were enrolled in one of the extension studies (US or non-US) and received at least one dose of study drug. About 43% (n = 427) of patients withdrew, primarily due to patient decision (n = 168), AEs (n = 85), other (n = 70), investigator decision (n = 48), loss to follow-up (n = 25), lack of compliance (n = 12), and lack of efficacy (n = 16). Table 30 provides the baseline characteristics of this cohort.

Table 30: Baseline Characteristics in Pooled US and Non-US Extension of BLISS 52 and BLISS 76

Characteristic	Pooled US and non-US extensions
	Belimumab IV (N = 998)
Age, years, mean (SD)	38.7 (11.5)
Female, n (%)	940 (94.2)
Disease duration, years, mean (SD)	6.7 (6.2)
SELENA-SLEDAI score, mean (SD)	8.2 (4.2)
SELENA-SLEDAI score ≥ 10, n (%)	364 (36.5)
BILAG score, n (%)	
≥ 1A or 2B	462 (46.3)
≥ 1A	128 (12.8)
≥ 1B	738 (73.9)
No A or B	224 (22.4)
PGA score, mean (SD)	1.2 (0.6)
SDI score (SD)	0.7 (1.2)
≥ 1 SFI flare, n (%)	186 (18.6)
≥ 1 severe SFI flare, n (%)	8 (0.8)

BILAG = British Isles Lupus Assessment Group; PGA = Physician Global Assessment; SD = standard deviation; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SFI = System Lupus Erythematosus Flare Index.

Source: Bruce et al. (2016).[35]

Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index

Table 31 provides data for SDI score changes for each year of the study.[34] Most patients experienced no change in SDI score. At year 5 to 6, the mean (SD) score change from baseline was 0.2 (0.5). Among 35 patients, the SDI score decreased and, since the score represents irreversible damage and cannot decrease, these patients were omitted from primary analyses. A post hoc sensitivity analysis was conduced by imputing the highest SDI score for omitted patients to subsequent assessments. In this sensitivity analysis, the change in SDI from baseline was similar to that in the primary analysis.[35] In patients without organ damage at baseline, the mean SDI (SD) score change from baseline to year 5 to 6 was 0.2 (0.4).[35] In patients with organ damage at baseline, the change from baseline was 0.2 (0.5).[35] In patients with baseline SELENA-SLEDAI score of 10 or more, the mean change from baseline was 0.2 (0.5).[35] At year 5 to 6, organ systems with the highest rates of new damage were ocular (4.7%) and musculoskeletal (3.7%), with the third-highest rate due to diabetes (1.7%).[35] The mean (interquartile range) for time to first SDI score worsening was 677 (364 to 1,045) days.[35]

Table 31: Change in SDI Score in Pooled US and Non-US Extensions of BLISS 52 and BLISS 76

	Pooled US and non-US extensions						
	Year						
	0 to 1 N = 941	1 to 2 N = 887	2 to 3 N = 785	3 to 4 N = 677	4 to 5 N = 565	5 to 6 N = 403	Year 6+ N = 354
No change, n (%)	896 (95.2)	821 (92.6)	702 (89.4)	591 (87.3)	488 (86.4)	343 (85.1)	104 (82.5)
+1, n (%)	40 (4.3)	58 (6.5)	69 (8.8)	68 (10.0)	59 (10.4)	46 (11.4)	16 (12.7)
+2, n (%)	3 (0.3)	6 (0.7)	13 (1.7)	18 (2.7)	16 (2.8)	13 (3.2)	5 (4.0)
+3, n (%)	2 (0.2)	2 (0.2)	1 (0.1)	0 (0)	2 (0.4)	1 (0.2)	1 (0.8)
Mean change (SD)	0.06 (0.3)	0.09 (0.3)	0.12 (0.4)	0.15 (0.4)	0.17 (0.5)	0.19 (0.5)	0.23 (0.6)

SD = standard deviation.

Source: Clinical Study Report for pooled extension.[34]

C. Open-Label Extension of Subcutaneous Belimumab

Phase III (BLISS SC Extension)

Study and Phase Design

This study was a six-month, multi-centre (30 countries in North America), extension of BLISS SC that was conducted from December 17, 2012, to October 1, 2015.[36] Patients were from US or Canada (24.9%), Eastern Europe (24.2%), Asia (21.6%), Americas (excluding US or Canada) (21.1%), and Western Europe/Australia/Israel (8.2%). The primary objective was to evaluate safety, followed by efficacy, of SC belimumab. Patients who completed the 52-week randomized phase of BLISS SC were eligible to enrol in the open-label extension study. Patients were required to be on a stable treatment regimen for at least 30 days beforehand, consisting of either prednisone equivalent 0 to 40 mg/d in combination with other SLE treatments, or 7.5 to 40 mg/d if used alone, antimalarial drugs, NSAIDs, or immunosuppressants. The use of other biologics or IV cyclophosphamide were not permitted. In the open-label extension, patients received belimumab SC 200 mg weekly plus standard of care. In addition to harms and SDI, the following efficacy outcomes were assessed: FACIT-Fatigue, SELENA-SLEDAI score, PGA score, BILAG score, modified SFI, and prednisone dosage. The analysis population consisted of all patients who received at least one dose of study drug. Baseline was defined as the last available value before starting belimumab.

Results

In the parent study, 839 patients were randomized, and 677 completed the DB phase to week 52. Of the 677 patients, 662 (97.8%) were enrolled in the open-label extension study and received at least one dose of study drug (206 switched from placebo and 456 continued on belimumab). Thirty-seven patients (5.6%) withdrew from the extension study, due to AE (n = 18), other (n = 7), patient decision (n = 5), disease progression/lack of efficacy (n = 4), loss to follow-up (n = 2), or protocol violation (n = 1). The mean (SD) duration of exposure to belimumab in the open-label extension was 162.9 (SD = 25.2) days.

The baseline characteristics of patients in BLISS SC Extension are provided in Table 32 for the total population, patients who switched from placebo, and patients who continued on belimumab. The mean age of the cohort was close to 39 years. The majority were women and white. Patients had had disease for a mean of about seven years. Several disease

parameters indicated more severe disease in the patients who continued on belimumab compared with patients who were switched from placebo, such as the SELENA-SLEDAI score, BILAG organ involvement, anti-dsDNA of 30 IU/mL or more, and use of prednisone at a dosage of more than 7.5 mg per day. This difference was due to the definition of baseline (i.e., last value before starting belimumab), which was at week 52 of the DB phase of BLISS SC for the placebo group and before the start of the DB phase for patients who continued on belimumab.

Table 32: Baseline Patient Characteristics in BLISS SC Extension

Characteristic	BLISS SC extension		
	Belimumab SC		
	Placebo to belimumab (N = 206)	Belimumab to belimumab (N = 456)	Total (N = 662)
Age, years, mean (SD)	39.4 (12.0)	38.3 (11.8)	38.7 (11.9)
Female, n (%)	196 (95.1)	430 (94.3)	626 (94.6)
Ethnic origin, n (%)			
White	126 (61.2)	277 (60.7)	403 (60.9)
Asian	49 (23.8)	98 (21.5)	147 (22.2)
African-American/African heritage	16 (7.8)	40 (8.8)	56 (8.5)
American Indian or Alaska Native	15 (7.3)	40 (8.8)	55 (8.3)
Native Hawaiian or other Pacific islander	0 (0)	1 (0.2)	1 (0.2)
Mixed	2 (1.0)	6 (1.3)	8 (1.2)
Disease duration, years, mean (SD)	7.2 (6.6)	6.5 (6.6)	6.7 (6.6)
SELENA-SLEDAI score, mean (SD)	5.8 (3.9)	10.4 (3.2)	9.0 (4.0)
SELENA-SLEDAI score ≥ 10, n (%)	27 (13.1)	285 (62.5)	322 (48.6)
BILAG, n (%)			
≥ 1A or 2B	42 (20.4)	314 (68.9)	356 (53.8)
≥ 1A	5 (2.4)	69 (15.1)	74 (11.2)
≥ 1B	123 (59.7)	405 (88.8)	528 (79.8)
No A or B	79 (38.3)	26 (5.7)	105 (15.9)
PGA score, mean (SD)	0.8 (0.6)	1.6 (0.4)	1.3 (0.6)
≥ 1 SFI flare, n (%)	21 (10.2)	51 (11.2)	72 (10.9)
≥ 1 severe SFI flare, n (%)	2 (1.0)	4 (0.9)	6 (0.9)
SDI score, mean (SD)	0.7 (1.2)	0.6 (1.0)	0.6 (1.1)
Proteinuria ≥ 2 g in 24 hours, n (%)	9 (4.4)	11 (2.4)	20 (3.0)
Anti-dsDNA ≥ 30 IU/mL, n (%)	126 (61.8)	333 (73.0)	459 (69.5)
Prednisone, n (%)	174 (84.5)	398 (87.3)	572 (86.4)
Prednisone > 7.5 mg/day, n (%)	102 (49.5)	275 (60.3)	377 (56.9)
Antimalarial drugs, n (%)	138 (67.0)	327 (71.7)	465 (70.2)
Immunosuppressants, n (%)	97 (47.1)	203 (44.5)	300 (45.3)

Characteristic	BLISS SC extension		
Aspirin, n (%)	34 (16.5)	76 (16.7)	110 (16.6)
NSAIDs, n (%)	52 (25.2)	101 (22.1)	153 (23.1)

Anti-dsDNA = anti–double-stranded DNA; BILAG = British Isles Lupus Assessment Group; NSAID = nonsteroidal anti-inflammatory drug; PGA = Physician Global Assessment; SC = subcutaneous; SD = standard deviation; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SFI = System Lupus Erythematosus Flare Index.

Source: Clinical Study Report for BLISS SC Extension.[36]

Harms

Nearly 50% of patients experienced at least one AE (Table 33). The most frequent AEs (≥ 3%) were viral upper respiratory tract infection (3.9%), arthralgia (2.7%), bacterial urinary tract infection (2.6%), and nasopharyngitis (2.4%). At least one SAE was experienced by 6% of patients. The most frequent SAEs (two or more patients) were acute kidney injury (n = 3), deep vein thrombosis (n = 3), abdominal abscess (n = 2), dyspnea (n = 2), herpes zoster (n = 2), and pneumonia (n = 2). The study drug was discontinued by 17 (2.6%) patients due to an AE; five patients switched from placebo to belimumab; and 12 patients continued on study drug. Local injection-site reactions (i.e., erythema, pain, induration, and urticaria) occurred in four patients, and post-injection systemic reactions in 21 (3.2%). Seventeen (2.6%) patients experienced an infection of special interest (e.g., opportunistic infections, tuberculosis, herpes zoster), and five had a serious infection. No malignant neoplasms were observed. Depression was present in 12 (1.8%) patients, and suicide or self-injury in one patient. There were two deaths, due to metabolic acidosis and acute respiratory failure.

Table 33: Harms in BLISS SC Extension

	BLISS SC extension		
	Placebo to belimumab (N = 206)	Belimumab to belimumab (N = 456)	Total (N = 662)
≥ 1 AE, n (%)	106 (51.5)	220 (48.2)	326 (49.2)
≥ 1 SAE, n (%)	14 (6.8)	25 (5.5)	39 (5.9)
WDAE, n (%)	5 (2.4)	12 (2.6)	17 (2.6)
Death, n (%)	1 (0.5)	1 (0.2)	2 (0.3)

AE = adverse event; SAE = serious adverse event; SC = subcutaneous; WDAE = withdrawal due to adverse event.

Source: Clinical Study Report for BLISS SC Extension.[36]

Efficacy

Table 34 shows the efficacy results observed in BLISS SC Extension. An SRI response was achieved by 61.4% of patients. Patients who switched from placebo to belimumab had lower SRI response than patients who continued on belimumab (16.1% and 76.3%, respectively). These two groups of patients are difficult to compare, as baseline was defined differently (i.e., for placebo to belimumab, baseline was before the first dose of belimumab at the end of the DB 52-week phase, whereas, for patients who continued on belimumab, baseline was before the start of the DB phase). Therefore, at baseline of the extension study, patients who switched from placebo to belimumab had lower disease activity (due to standard of care in the DB phase) than patients who continued on belimumab. For the components of the SRI response, a reduction in SELENA-SLEDAI score of four points or more was achieved by 64% of patients, no worsening in PGA score

(increase of < 0.3 points) by 95.3%, and no new BILAG 1A or 2B by 95.3%. The mean (SD) SELENA-SLEDAI score decrease from baseline was 4.6 (4.5) points. Any flare was experienced by 14.5% of patients and severe flare by 2.1%. Worsening in SDI was experienced by 25 (3.8%) patients. Among those on prednisone at a dosage of more than 7.5 mg per day, about 20% were able to reduce the dose to 7.5 mg per day or less by the end of the six-month open-label extension. The mean change in prednisone dosage from baseline was 1.6 mg per day. The FACIT-Fatigue score mean (SD) change from baseline was 4.0 (9.9) points. Improvement in FACIT-Fatigue above the sponsor-defined MID (≥ 4) occurred in 43.4% of patients.

Table 34: Efficacy in BLISS SC Extension

	BLISS SC extension		
	Placebo to belimumab (N = 206)	**Belimumab to belimumab (N = 456)**	**Total (N = 662)**
SRI response, n/N (%)	23/143 (16.1)	332/435 (76.3)	355/578 (61.4)
≥ 4-point reduction in SELENA-SLEDAI score, n/N (%)	25/143 (17.5)	345/435 (79.3)	370/578 (64.0)
No worsening in PGA, n/N (%)	125/143 (87.4)	426/435 (97.9)	551/578 (95.3)
No new BILAG 1A or 2B, n/N (%)	134/143 (93.7)	417/435 (95.9)	551/578 (95.3)
SELENA-SLEDAI score change from baseline, mean (SD)	–0.7 (2.8)	–6.3 (4.0)	–4.6 (4.5)
SFI any flare, n (%)	38 (18.4)	58 (12.7)	96 (14.5)
Severe flare, n (%)	2 (1.0)	12 (2.6)	14 (2.1)
SDI score change from baseline, mean (SD)	0.0 (0.2)	0.0 (0.2)	0.0 (0.2)
SDI worsening	6 (2.9)	19 (4.2)	25 (3.8)
Prednisone			
Any decrease in dose from baseline, n/N (%)	20/174 (11.5)	125/398 (31.4)	145/572 (25.3)
Reduced to ≤ 7.5 mg/day, n/N (%)	10/102 (9.8)	67/275 (24.4)	77/377 (20.4)
FACIT-Fatigue change from baseline, mean (SD)	0.7 (7.1)	5.6 (10.6)	4.0 (9.9)

BILAG = British Isles Lupus Assessment Group; FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy–Fatigue; PGA = Physician Global Assessment; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SFI = System Lupus Erythematosus Flare Index; SRI = Systemic Lupus Erythematosus Responder Index.

Source: Clinical Study Report for BLISS SC extension.[36]

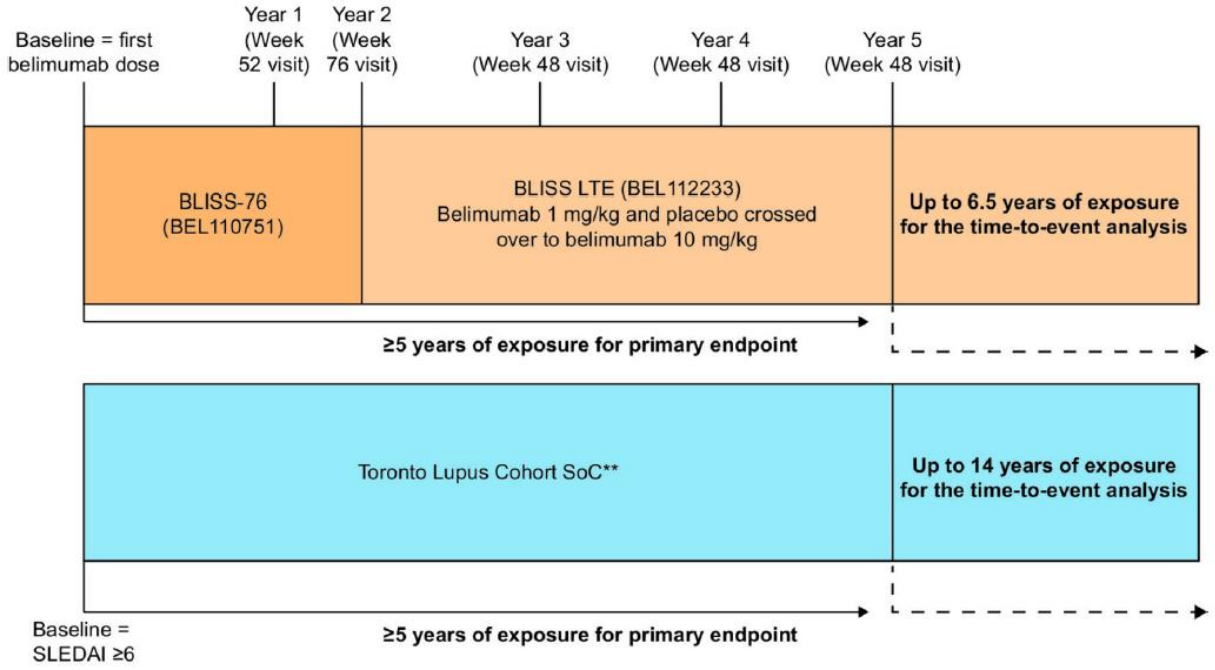
D. Longitudinal Propensity-Score–Matched Study of IV Belimumab + Standard of Care Versus Standard of Care Alone

Study Design

A post hoc, observational, longitudinal, propensity-score–matched study was conducted to compare IV belimumab plus standard of care with standard of care alone, using patient cohorts from the US BLISS Extension and the Toronto Lupus Cohort (TLC) (Figure 13).[37] The two cohorts were matched 1:1 with a propensity score, with a caliper of 20% of the SD of the propensity-score distribution. Before matching, the eligibility criteria of the US BLISS

Extension (i.e., ≥ 18 years of age, diagnosis of SLE based on ≥ 4 of 11 ACR criteria, SELENA-SLEDAI or SLEDAI-2K score ≥ 6, and autoantibody-positive) were applied to patients in the TLC. Baseline was defined as the date of first exposure to belimumab for the BLISS Extension cohort, or date of obtaining an SLEDAI-2K score of 6 or more for the TLC (SLEDAI-2K ≥ 6 was an inclusion criterion of the BLISS Extension parent study). The primary outcome was the mean difference in SDI change from baseline to five years. Secondary outcomes were time to first worsening of SDI score in patients with one year or more of follow-up, and magnitude of SDI score worsening in patients with one year or more and five years of follow-up. Mean difference in SDI score change was estimated with a linear regression model. To account for loss of patients after matching, the difference in change in SDI score was also estimated with a sensitivity analysis that used the entire patient sample and the propensity score to weight the observations (an inverse propensity-score–weighted analysis), with the addition of unbalanced matching variables as covariates. Time to first worsening of SDI included the full longitudinal data available for patients in BLISS Extension (up to 6.5 years) and the TLC (up to 14 years) and was analyzed with a parametric survival model.

Figure 13: Study Design of Propensity-Score–Matched Analysis



LTE = long-term extension; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SoC = standard of care.

Source: Reprinted from Urowitz et al. (2019) Figure 1, under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).[37]

Results

Of 195 patients in BLISS Extension and 372 patients in the TLC with at least five years of follow-up, 99 patients in each cohort could be matched 1:1 for the primary outcome.

Table 35 shows the baseline characteristics that were used in the calculation of the propensity score, before and after matching. For each characteristic, bias was calculated as the standardized distance between the cohorts. After matching, bias for most

characteristics was considerably reduced, and the cohorts were well balanced. Of 259 patients in BLISS Extension and 706 patients in the TLC with at least one year of follow-up, 179 patients could be matched (Table 36). Before matching, there were baseline differences between the cohorts; however, characteristics were well balanced after matching.

Table 35: Baseline Characteristics Before and After Matching in Patients With Five Years of Follow-Up

Characteristic	Propensity-score–matched study					
	Before propensity-score–matching			After propensity-score–matching		
	Belimumab (N = 195)	SoC (N = 372)	Bias (%)	Belimumab (N = 99)	SoC (N = 99)	Bias (%)
Age, years, mean	42.8	37.3	45.5	40.0	39.0	8.4
Female, %	92.8	89.5	11.6	92.9	91.9	3.8
Ethnic origin, %						
Black	23.1	15.3	19.7	21.2	23.2	–4.8
Asian/Other	9.2	23.4	–39.0	14.1	12.1	6.0
Disease duration, years, mean	7.9	5.8	30.0	7.4	7.6	–2.6
Smoker, %	3.6	23.7	–61.1	7.1	7.1	0.0
Hypertension, %	67.7	37.6	63.0	54.5	53.5	2.0
Dyslipidemia, %	22.6	58.1	–77.5	28.3	31.3	–6.6
Proteinuria, %	12.3	31.7	–48.1	20.2	18.2	5.1
Number of ACR criteria satisfied	5.9	5.7	19.8	6.0	5.9	6.5
Baseline SLEDAI	7.8	10.1	–48.4	8.5	8.5	–2.2
Corticosteroid, %	63.6	60.8	5.8	64.6	66.7	–4.2
Antimalarial drug, %	73.8	51.9	46.6	69.7	68.7	2.2
Immunosuppressant, %	53.8	31.5	46.4	45.5	44.4	2.0
SDI score = 1, %	27.2	14.8	30.7	24.2	27.3	–6.9
SDI score ≥ 2, %	28.7	10.8	46.2	15.2	18.2	–8.1

ACR = American College of Rheumatology; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SoC = standard of care; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index.

Source: Reprinted from Urowitz et al. (2019) Table 1, under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0), with changes.[37]

Table 36: Baseline Characteristics Before and After Matching in Patients With One Year or More of Follow-Up

Characteristic	Propensity-score–matched study					
	Before propensity-score–matching			After propensity-score–matching		
	Belimumab (N = 259)	SoC (N = 706)	Bias (%)	Belimumab (N = 179)	SoC (N = 179)	Bias (%)
Age, years, mean	42.6	36.9	46.0	40.4	40.7	–2.4
Female, %	93.4	88.8	16.3	91.6	91.6	0.0
Ethnic origin, %						

Characteristic	Propensity-score–matched study					
	Before propensity-score–matching			After propensity-score–matching		
	Belimumab (N = 259)	SoC (N = 706)	Bias (%)	Belimumab (N = 179)	SoC (N = 179)	Bias (%)
Black	21.6	14.6	18.3	22.3	23.5	–2.7
Asian/Other	9.3	28.0	–49.6	12.8	12.8	0.0
Disease duration, years, Mean	7.7	6.2	21.5	7.5	7.7	–3.2
Smoker, %	3.9	24.2	–61.2	5.6	6.7	–4.6
Hypertension, %	53.3	38.0	31.1	45.8	45.8	0.0
Dyslipidemia, %	22.8	34.7	–26.5	25.1	22.9	5.2
Proteinuria, %	13.5	33.0	–47.4	16.8	17.9	–2.9
Number of ACR criteria satisfied	6.0	5.7	22.0	6.0	5.9	1.9
Baseline SLEDAI	7.9	10.0	–49.0	8.4	8.5	–3.7
Corticosteroid, %	64.9	62.5	5.0	68.2	69.3	–2.4
Antimalarial drug, %	71.8	56.4	32.6	65.9	67.0	–2.4
Immunosuppressant, %	55.2	34.4	42.7	45.8	46.4	–1.1
SDI score = 1, %	27.8	14.2	33.9	24.6	25.7	–2.6
SDI score ≥ 2, %	27.8	10.2	46.0	16.8	16.8	0.0

ACR = American College of Rheumatology; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SoC = standard of care; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index.

Source: Reprinted from Urowitz et al. (2019) Table 1, under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0), with changes.[37]

Primary outcome: Among the 99 matched patients, the change in SDI score from baseline to 5 years was statistically lower in patients who received belimumab compared with standard of care (mean difference in change for belimumab versus standard of care = –0.4; 95% CI, –0.7 to –0.2) (Table 37). In the inverse propensity-score–weighted sensitivity analysis, the difference in SDI change between belimumab and standard of care was similar (–0.4; 95% CI, –0.7 to –0.2).[37] The cohorts in the sensitivity analysis, however, were not balanced. In another sensitivity analysis that included characteristics with greater than 10% bias as covariates, the difference in change in SDI score was also similar (–0.45; 95% CI, –0.7 to –0.2).[37]

Secondary outcomes: Among the 179 matched patients with one year or more of follow-up, patients who received belimumab were less likely to progress to organ damage than patients who received standard of care alone (HR = 0.4; 95% CI, 0.3 to 0.6; P < 0.001).[37] An SDI score increase of 1 or more occurred 33 times in patients treated with belimumab and 72 times in patients treated with standard of care alone. Of these patients, two patients (6.1%) on belimumab and 22 patients (30.6%) on standard of care alone had SDI score increases of 2 or more.

Table 37: Change in SDI Score in Longitudinal Propensity-Score–Matched Study

	Propensity-score–matched Study	
	Belimumab (N = 99)	**SoC (N = 99)**
5-year change in SDI score, mean	0.3	0.7
Difference (95% CI) (belimumab versus SoC)	–0.4 (–0.7 to –0.2)	
P value	< 0.001	

CI = confidence interval; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SoC = standard of care.

Source: Urowitz et al. (2019) Table 3, under licence Attribution-NonCommercial 4.0 International (CC BY-NC 4.0), with changes.[37]

Critical Appraisal of Propensity-Score–Matched Study

The methods for selecting a comparison cohort and developing the propensity scores, were rigorous. To identify an appropriate comparator, a systematic review of cohorts, registries, and other databases was conducted to identify patient populations with similar characteristics as the BLISS Extension and with at least five years of follow-up.[37] Those studies with at least 400 patients and three peer-reviewed publications were reviewed in detail.[37] Of 21 such cohorts identified, the TLC was chosen as the most appropriate comparator for BLISS Extension.[37] The propensity score was also based on a systematic review of the literature for predictors of organ damage and progression in SLE.[37] The predictors identified from the literature were then reviewed by a clinical expert and limited to those that were available in both BLISS Extension and the TLC. The following 14 predictors went into the calculation of the propensity score: age; female gender; race/ethnicity; SLE duration; history of hypertension; dyslipidemia; proteinuria; current smoker; number of ACR criteria satisfied; baseline SLEDAI score; use of corticosteroids, antimalarial drugs, or immunosuppressants; and baseline SDI score (1 or ≥ 2).[37] These formed the independent variables of a logistic regression model, and the propensity score was the estimated log-odds (calculated for each patient). Although baseline differences were present in the cohorts at the start, after matching on the propensity score (± caliper), the cohorts were well balanced.

A large percentage of patients in both cohorts were excluded from the matched analyses. For patients with at least five years of follow-up, the matched analyses excluded 96/195 (49%) of patients in BLISS Extension and 273/372 (73%) in the TLC. For patients with at least one year of follow-up, the matched analyses excluded 80/259 (31%) of patients in BLISS Extension and 527/706 (75%) in the TLC. For the primary outcome of difference in change in SDI score from baseline to five years, sensitivity analyses were conducted that included the full set of patients, and results were similar to the matched analysis. This provides some degree of confidence in the stability of the primary outcome and its applicability to the entire cohorts. However, the secondary outcomes are less certain. The exclusion of such a large percentage of patients from analyses may impact the generalizability of results. In addition, patients with active severe lupus nephritis or CNS lupus were excluded, and, therefore, the comparative effects of belimumab and standard of care alone in these patients are unknown.

The outcome examined in the study, organ damage with the SDI score over five years, is important in the assessment of SLE progression. No MID is available for the SDI, so it is unclear whether the observed difference in scores was clinically meaningful. Also, no other outcomes of disease activity, such as the SLEDAI or BILAG, were examined. The BLISS

Extension cohort received belimumab + standard of care, and the TLC cohort received standard of care alone. It is unclear how similar the standard of care treatments were in both cohorts and whether the results can be ascribed to the effect of belimumab only. The intervals of assessment and follow-up times were different for BLISS Extension and the TLC. In BLISS Extension, assessments were conducted every year (Figure 13), whereas, in the TLC assessments, they were conducted at three- to four-month intervals.[37] For the outcome of time to progression of organ damage, there was up to 6.5 years of follow-up data available for BLISS Extension, and up to 14 years of data in the TLC (Figure 13). The decade on entry into the study was added as a covariate in models of the primary and secondary outcomes (time to first SDI worsening) to account for potential changes in standard of care over time.

This study provides some preliminary comparative data over five years for IV belimumab + standard of care versus standard of care alone for organ damage in patients with SLE. The study suggested that belimumab is associated with lower organ damage progression compared with standard of care alone over five years. The cohorts chosen were appropriate for comparison, and matching was adequate. However, a large percentage of patients were excluded from the main analyses, which may affect the generalizability of the results, and organ damage was assessed in isolation, without any other outcomes of disease activity. While the study does fill a gap in the evidence base for long-term comparative data, the results require confirmation by additional studies.

Discussion

Summary of Available Evidence

One multinational manufacturer-sponsored DB RCT, BLISS SC, met the inclusion criteria for this systematic review. BLISS SC randomized 836 patients with active SLE (SELENA-SLEDAI score ≥ 8), 2:1, to either belimumab or placebo over a treatment course of 52 weeks. The primary outcome of this study was patients with an SRI response at 52 weeks, and secondary outcomes included time to severe flare and the percentage of patients who could reduce their prednisone dosage by 25% or more to 7.5 mg daily or less.

Indirect evidence was available from an ITC submitted by the sponsor, as well as two published ITCs. The sponsor-submitted ITC compared the IV and SC formulations of belimumab in patients with SLE and HDA, using patient-level data from phase III trials. Trials included were BLISS SC, BLISS 52, BLISS 76, and BEL 113750, and the primary outcome was SRI response at 52 weeks. One of the published ITCs compared the efficacy and safety of belimumab IV or SC to placebo, in patients with active SLE, and the efficacy data were based on many of the same trials as the sponsor-submitted ITC. The other published ITC compared safety of various drugs for SLE (immunosuppressants, corticosteroids, and biologics) in patients with SLE, and included 44 trials and 9,898 patients with SLE.

Other studies reviewed included the two pivotal multinational manufacturer-sponsored phase III DB RCTs of the IV formulation of belimumab, approved by Health Canada in 2011. Both BLISS 52 and BLISS 76 enrolled patients with active SLE (SELENA-SLEDAI score of ≥ 6) and compared belimumab to placebo over 52 and 76 weeks, respectively. The primary outcome of both BLISS 52 and BLISS 76 was the percentage of patients with an SRI response at week 52. Extensions were also reviewed, including a six-month extension to BLISS SC and multi-year extensions to studies involving the IV formulation, with follow-up of up to 13 years. Additionally, a long-term propensity-score–matched study compared belimumab plus standard of care to standard of care alone with respect to organ damage; the primary outcome was the change in SDI score and the secondary outcome was time to worsening in SDI.

Interpretation of Results

Efficacy

Results from the primary outcome of BLISS SC suggest a treatment effect of about 13% for an SRI response at 52 weeks. The limitations of this composite, including questions as to how the cut points for PGA and SELENA-SLEDAI were derived, introduce some uncertainty into these findings, especially in light of the 13% treatment effect. The clinical experts consulted by CDR on this review believed this to be a clinically significant response in a disease with limited therapeutic options and no new approved therapies in approximately 50 years. Aside from antimalarial drugs, which are reasonably safe and well tolerated, all other options for treating SLE have significant tolerability issues and a number of serious harms. Most notable among these harmful drugs that form the cornerstone of SLE therapies are corticosteroids.

The clinical experts consulted by CDR on this review noted the importance of corticosteroids such as prednisone in the management of SLE, including the significant

burden of side effects associated with this class. In their input to CDR, patients noted the burden of side effects with corticosteroids as a limitation of treatments for SLE. Treatment with belimumab for 52 weeks did not elicit a statistically significant reduction in patients' ability to reduce their prednisone dosage over the study period, although there was a numerical improvement in the percentage of belimumab-treated patients who could reduce their prednisone dosage. Interpretation of this finding is complicated by the smaller sample size of patients upon which this outcome was based; however, similar findings occurred in the studies of IV belimumab, in which statistical improvement in SRI response was seen after 52 weeks but a reduction in corticosteroid dosage was not. It is not clear why a reduction in disease activity with belimumab might not translate into a reduction in corticosteroid use; however, this is unfortunate given the devastating impact of chronic corticosteroid use and how common it is for physicians to have to resort to chronic corticosteroids in managing patients with active SLE. The clinical expert consulted by CDR on this review suggested that the lack of ability to reduce corticosteroid dose below 5 mg to 7.5 mg per day may be related to adrenal suppression, rather than to disease flares due to SLE.

The lack of data on health-related quality of life in BLISS SC is a limitation of this review. In their input to CDR, patients state that SLE has a significant impact on quality of life. Given that SLE is a multi-system disorder, the impact on quality of life and activities of daily living are also variable, although fatigue and pain are common themes. The SF-36 was assessed in the extension to BLISS SC; however, the lack of a control group limits any conclusions that can be drawn from these data. The FACIT-Fatigue instrument was assessed as an exploratory outcome; however, without controlling for multiple statistical comparisons, limited conclusions can be drawn from these data.

SLE also causes significant damage to many vital organs and tissues, most notably the kidneys and those in the CNS. These effects of the disease take longer to develop. According to the clinical experts consulted by CDR on this review, it is unlikely that a 52-week study like BLISS SC would be able to demonstrate a reduction in the accumulation of organ damage. Thus, in their opinion, it is not surprising that there was no statistically significant difference in the SDI, used to assess organ damage, between belimumab and placebo, and this was an exploratory outcome of this study. It should also be noted that BLISS SC excluded patients with severe renal or CNS involvement, and, thus, the effects of belimumab cannot be ascertained in this population. As noted, there are longer-term extensions available, particularly involving the IV formulation. However, these are again limited by the lack of a comparator. A published propensity-score–matched analysis found patients treated with belimumab IV plus standard of care had reduced organ damage progression versus a matched cohort of those receiving standard of care alone, over a five-year treatment period.[37] Although this analysis does provide longer-term comparative data, there are limitations to such an observational analysis. More study is needed before concluding that belimumab has a positive impact on organ damage in SLE.

There are no direct comparisons of the addition of belimumab versus other treatments such as antimalarial drugs or immunosuppressants for SLE. The available ITCs do not add much to the understanding of the relative efficacy of belimumab versus these drugs, as the two ITCs that assessed efficacy merely compared belimumab formulations to each other, and these ITCs included the same studies that are already described in the Other Relevant Studies section of this review. The only ITC that compared belimumab to other drugs focused on safety. This ITC, by Tian et al., suggest there is no difference between belimumab IV or SC and comparators such as rituximab, immunosuppressants, or

corticosteroids for outcomes such as mortality, AEs, SAEs, serious infections, and WDAEs. There were a number of limitations of this ITC and, thus, uncertainty concerning the results.

The sponsor's request is for belimumab to be reimbursed for adults with active SLE, currently receiving standard therapy, and with SELENA-SLEDAI score of at least 8, and the population enrolled in BLISS SC appears to be consistent with this request, although it is unclear to what extent standard of care had been optimized. In the protocol for this systematic review, CDR sought subgroup data based on baseline disease activity (SELENA-SLEDAI score) and prior therapies. Although these subgroup data were available from the included study, there was no adjustment for multiple comparisons; thus, limited conclusions can be drawn from these data. Of note, in the subgroup of patients who were not taking corticosteroids at baseline, there was no difference in SRI response between belimumab and placebo.

Harms

Based on its mechanism of action, targeting B cells, infection is one of the notable harms that should be monitored with belimumab. The risk of serious infections, including progressive multifocal leukoencephalopathy, is noted as a black box warning in the product monograph for belimumab.[6] As noted previously, belimumab has a history dating back to its approval as an IV formulation by Health Canada in 2011. There has been no indication from BLISS SC of an increased risk of infection or of serious infection from belimumab treatment, although all three deaths in the belimumab group were related to infections (sepsis, urosepsis, and tuberculosis). Longer-term extensions, with up to 13 years' follow-up, have followed patients on the original IV formulation. There is no clear signal of high risk of opportunistic infections from these studies, although the conclusions that can be drawn are limited by the lack of control group and attrition over the many years of the study. Concerns over infection risk with belimumab also need to be weighed against the risk of infection of many of its comparators, including immunosuppressants and corticosteroids. Both groups of drugs are known for their increased infection risk and potential for increased malignancy risk. As noted, an ITC that focused on comparative safety of various drugs for SLE found no differences for a variety of safety and tolerability outcomes, although there were methodological issues that limit the conclusions that can be drawn from this study.

The product monograph also notes that depression, suicidality, and suicides have been reported in studies of belimumab, although it is not known whether belimumab is the cause. Psychiatric AEs were a notable harm of this review, and there was no indication of a numerical increase in risk of these events with belimumab compared to placebo. Patients judged recently to be at high risk of suicide were excluded from BLISS SC; thus, the safety of belimumab in these patients is unknown.

Conclusions

One multinational manufacturer-sponsored DB RCT was included in this review. In a population of patients with active SLE (SELENA-SLEDAI score of 8 or more), belimumab reduced disease activity after 52 weeks compared to placebo, as measured by SRI responses. Although belimumab reduced the risk of a severe flare over this time period, it did not elicit a reduction in prednisone dosage. Chronic use of corticosteroids contributes to morbidity in patients with SLE due to the severe adverse effects of these drugs. Health-related quality of life was not studied, and this is an important gap in a severe, chronic multi-system disorder like SLE. The duration of the study was not sufficient to study the effects of belimumab on organ damage. Findings from ITCs suggest there is no difference in efficacy between the SC and IV formulations of belimumab, and no difference between belimumab and other commonly used drugs for SLE (immunosuppressants, corticosteroids, and biologics) with respect to harms, although the latter finding has a high degree of uncertainty. Data from the DB phase as well as the extensions to the SC and IV formulations do not suggest issues with tolerability or safety, although the extensions are limited by a lack of control group.

Appendix 1: Literature Search Strategy

Clinical Literature Search

OVERVIEW	
Interface:	Ovid
Databases:	MEDLINE All (1946 to present) Embase (1974 to present) **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	June 26, 2019
Alerts:	Biweekly search updates until project completion
Study Types:	Search filters were applied to limit retrieval to randomized controlled trials or controlled clinical trials.
Limits:	Retrieval was not limited by publication date or by language. Conference abstracts were excluded.

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
MeSH	Medical Subject Heading
exp	Explode a subject heading
*	Before a word, indicates that the marked subject heading is a primary topic; or, after a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
.ti	Title
.ab	Abstract
.hw	Heading word; usually includes subject headings and controlled vocabulary
.kf	Author keyword heading word (MEDLINE)
.kw	Author keyword (Embase)
.pt	Publication type
.dq	Candidate term word
adj#	Requires terms to be adjacent to each other within # number of words (in any order)
.rn,nm	CAS registry number (MEDLINE)
.ot	Original title
medall	Ovid database code: MEDLINE All, 1946 to present, updated daily
oemezd	Ovid database code; Embase, 1974 to present, updated daily

MULTI-DATABASE STRATEGY	
Line #	**Search Strategy**
1	(benlysta* or belimumab* or benlista* or benlystia* or LimphoStat-B or LimphoStatB or LymphoStat-B or LymphoStatB or 73B0K5S26A or L04AA26 or HGS-1006 or HGS1006).ti,ab,kf,ot,hw,rn,nm.
2	1 use medall
3	*belimumab/
4	(benlysta* or belimumab* or benlista* or benlystia* or LimphoStat-B or LimphoStatB or LymphoStat-B or LymphoStatB or L04AA26 or HGS-1006 or HGS1006).ti,ab,kw,dq.
5	3 or 4
6	5 use oemezd

MULTI-DATABASE STRATEGY	
Line #	**Search Strategy**
7	2 or 6
8	(conference abstract or conference review).pt.
9	7 not 8
10	(Randomized Controlled Trial or Controlled Clinical Trial or Pragmatic Clinical Trial or Equivalence Trial or Clinical Trial, Phase III).pt.
11	Randomized Controlled Trial/
12	exp Randomized Controlled Trials as Topic/
13	"Randomized Controlled Trial (topic)"/
14	Controlled Clinical Trial/
15	exp Controlled Clinical Trials as Topic/
16	"Controlled Clinical Trial (topic)"/
17	Randomization/
18	Random Allocation/
19	Double-Blind Method/
20	Double-Blind Procedure/
21	Double-Blind Studies/
22	Single-Blind Method/
23	Single Blind Procedure/
24	Single-Blind Studies/
25	Placebos/
26	Placebo/
27	Control Groups/
28	Control Group/
29	(random* or sham or placebo*).ti,ab,hw,kf,kw.
30	((singl* or doubl*) adj (blind* or dumm* or mask*)).ti,ab,hw,kf,kw.
31	((tripl* or trebl*) adj (blind* or dumm* or mask*)).ti,ab,hw,kf,kw.
32	(control* adj3 (study or studies or trial* or group*)).ti,ab,kf,kw.
33	(Nonrandom* or non random* or non-random* or quasi-random* or quasirandom*).ti,ab,hw,kf,kw.
34	allocated.ti,ab,hw.
35	((open label or open-label) adj5 (study or studies or trial*)).ti,ab,hw,kf,kw.
36	((equivalence or superiority or non-inferiority or noninferiority) adj3 (study or studies or trial*)).ti,ab,hw,kf,kw.
37	(pragmatic study or pragmatic studies).ti,ab,hw,kf,kw.
38	((pragmatic or practical) adj3 trial*).ti,ab,hw,kf,kw.
39	((quasiexperimental or quasi-experimental) adj3 (study or studies or trial*)).ti,ab,hw,kf,kw.
40	(phase adj3 (III or "3") adj3 (study or studies or trial*)).ti,hw,kf,kw.
41	or/10-40
42	9 and 41
43	remove duplicates from 42

CLINICAL TRIAL REGISTRIES		
ClinicalTrials.gov	Produced by the US National Library of Medicine. Targeted search used to capture registered clinical trials. Search terms: (benlysta* OR belimumab*) AND systemic lupus erythematosus	
WHO ICTRP	International Clinical Trials Registry Platform, produced by the World Health Organization. Targeted search used to capture registered clinical trials. Search terms: (benlysta* OR belimumab*) AND systemic lupus erythematosus	

OTHER DATABASES		
PubMed	Searched to capture records not found in MEDLINE. Same MeSH, keywords, limits, and study types used as per MEDLINE search, with appropriate syntax used.	

Grey Literature

Dates for search:	June 18–20, 2019
Keywords:	(benlysta* OR belimumab* OR LimphoStat-B OR LimphoStatB OR LymphoStat-B OR LymphoStatB OR 73B0K5S26A OR L04AA26 OR HGS-1006 OR HGS1006) AND/OR systemic lupus erythematosus
Limits:	No limits

Relevant websites from the following sections of the CADTH grey literature checklist *Grey Matters: A Practical Tool For Searching Health-Related Grey Literature* (https://www.cadth.ca/grey-matters) were searched:

- Health Technology Assessment Agencies
- Health Economics
- Clinical Practice Guidelines
- Drug and Device Regulatory Approvals
- Advisories and Warnings
- Drug Class Reviews
- Clinical Trial Registries
- Databases (free)
- Health Statistics
- Internet Search
- Open Access Journals.

Appendix 2: Excluded Studies

Table 38: Excluded Studies

Reference	Reason for exclusion
Doria, Lupus 2018 27(9):1489–1498 Ginzler, Journal of Rheumatology 2014 41(2):300–9	Study design
Lee, Lupus 2018 27(1):112–119 Ramachandran, Journal of Comparative Effectiveness Research 2018 7(6):581–593	Systematic review
Gamble, Archives of Dermatology 2012 148(3):376–8 Furie, Arthritis and Rheumatism 2011 63(12):3918–30 Navarra, Lancet 2011 377(9767):721–31	Intervention

Appendix 3: Detailed Outcome Data

Table 39: Subgroup Analyses

	BLISS SC	
	Belimumab N = 556	**Placebo N = 280**
Subgroups, SRI responses based on:		
Baseline SELENA-SLEDAI score		
≤ 9	98/202 (48.5)	46/111 (41.4)
OR [95% CI]	1.33 [0.83 to 2.13]	
≥ 10	242/352 (68.8)	89/168(53.0)
OR [95% CI]	1.95 [1.34 to 2.85]	
Interaction P value	P = 0.2120	
Baseline C3 or C4 levels		
No low C3 or C4	167/285 (58.6)	74/153(48.4)
OR [95% CI]	1.51 [1.02 to 2.24]	
Low C3 and/or C4	173/269 (64.3)	61/126(48.4)
OR [95% CI]	1.92 [1.25 to 2.95]	
Interaction P value	P = 0.4206	
Baseline C3/C4 levels and anti-dsDNA		
Not ≥ 1 C3/C4 low and anti-dsDNA ≥ 30 IU/mL	181/308 (58.8)	84/171 (49.1)
OR [95% CI]	1.46 [1.00 to 2.14]	
≥ 1 C3/C4 low and anti-dsDNA ≥ 30 IU/mL	159/246 (64.6)	51/108 (47.2)
OR [95% CI]	2.23 [1.36 to 3.64]	
Interaction P value	P = 0.2393	
Baseline anti-dsDNA		
< 30 IU/mL	90/152 (59.2)	40/86 (46.5)
OR [95% CI]	1.67 [0.98 to 2.84]	
≥ 30 IU/mL	250/402 (62.2)	95/193 (49.2)
OR [95% CI]	1.70 [1.20 to 2.40]	
Interaction P value	P = 0.9601	
Baseline prednisone use		
Non-use	36/74 (48.6)	18/39 (46.2)
OR [95% CI]	1.11 [0.51 to 2.40]	
Use	304/480 (63.3)	117/240 (48.8)
OR [95% CI]	1.82 [1.33 to 2.49]	
Interaction P value	P = 0.2457	
Baseline average daily prednisone dosage		
0 mg/day	36/74 (48.6)	18/39 (46.2)
OR [95% CI]	1.11 [0.51 to 2.40]	
> 0 to ≤ 7.5 mg/day	98/146 (67.1)	35/72 (48.6)
OR [95% CI]	2.16 [1.21 to 3.84]	

	BLISS SC	
	Belimumab N = 556	**Placebo N = 280**
> 7.5 mg/day	206/334 (61.7)	82/168 (48.8)
OR [95% CI]	1.69 [1.16 to 2.45]	
Interaction P value	P = 0.3992	
Baseline medications		
Other	261/421 (62.0)	108/212 (50.9)
OR [95% CI]	1.57 [1.13 to 2.19]	
Steroid, antimalarial drug, and immunosuppressant	79/133 (59.4)	27/67 (40.3)
OR [95% CI]	2.17 [1.19 to 3.94]	
Interaction P value	P = 0.3571	
Baseline MMF		
No MMF	306/484 (63.2)	120/245 (49.0)
OR [95% CI]	1.79 [1.31 to 2.44]	
MMF	34/70 (48.6)	15/34 (44.1)
OR [95% CI]	1.20 [0.53 to 2.73]	
Interaction P value	P = 0.3691	

CI = confidence interval; dsDNA = double-stranded DNA; MMF = mycophenolate mofetil; OR = odds ratio; SELENA = Safety of Estrogen in Lupus Erythematosus National Assessment; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SRI = Systemic Lupus Erythematosus Responder Index.

Source: Clinical Study Report for BLISS SC.[2]

Appendix 4: Description and Appraisal of Outcome Measures

Aim

To describe the following outcome measures and review their validity, reliability, responsiveness to change, and MID:

- SLE Responder Index (SRI)
- Safety of Estrogens in Lupus Erythematosus National Assessment–SLE Disease Activity Index (SELENA-SLEDAI)
- British Isles Lupus Assessment Group (BILAG)
- Physician Global Assessment (PGA)
- SLE Flare Index (SFI)
- Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index (SDI)
- Functional Assessment of Chronic Illness Therapy–Fatigue Score (FACIT-Fatigue).

Table 40: Outcome Measures Included in Each Study

Outcome measure	BLISS SC
SRI	Primary
SELENA-SLEDAI	Subcomponent of primary[a]
BILAG	Subcomponent of primary[b]
PGA	Subcomponent of primary[c]
SFI	Secondary
SDI	Other secondary
FACIT-Fatigue	Exploratory

BILAG = British Isles Lupus Assessment Group; PGA = Physician Global Assessment; FACIT = Functional Assessment of Chronic Illness Therapy; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SFI = System Lupus Erythematosus Flare Index; SRI = Systemic Lupus Erythematosus Responder Index.

[a] Four point or greater reduction from baseline in SELENA-SLEDAI score.

[b] No new BILAG A organ domain score or two new BILAG B organ domain scores compared with baseline.

[c] No worsening (increase of < 0.3 points) in PGA.

In addition, a brief review of the application of the SF-36 to patients with SLE is included at the end of this section. The SF-36 was not assessed in BLISS SC; however, it was an outcome in the studies for IV belimumab (i.e., BLISS 52, BLISS 76, and the extensions of these studies).

Findings

Table 41: Summary of Outcome Measures and Their Measurement Properties

Outcome measure	Type	Conclusions about measurement properties	MID
SELENA-SLEDAI	A measure of disease activity at time of visit or in the preceding 10 days; consists of 24 weighted clinical and laboratory variables, with total possible score of 105 (higher scores represent greater disease activity)	**Validity:** ≥ 3-point reduction correlated with clinically meaningful change in BILAG and a reduction in therapy; ≥ 3-point increase associated with disease worsening and new/ increased therapy; correlated with BILAG, PGA, and HRQoL scores **Reliability:** Almost-perfect agreement between raters (ICC 0.87) **Responsiveness:** Less responsive to change than other disease measures	Anchor-based MIDs; clinically meaningful improvement: –7 points Clinically meaningful worsening: +8 points Associated with flare (SLEDAI): +4 points
BILAG	Scoring of eight organ domains on an ordinal scale of A to E over past 4 weeks. A total score is usually not calculated. A = most active B = moderate active C = minor activity D = stable E = never present	**Validity:** Sensitivity and specificity demonstrated with gold standard of new or increase in disease-modifying therapy; correlated with SLEDAI and PGA **Reliability:** Substantial to almost-perfect agreement between raters **Responsiveness:** Sensitive to change over time	Not identified[b]
PGA	Measure of current disease activity on a VAS with equal markings between 0 to 3, with higher scores representing worse disease activity 0 = none 1 = mild 2 = moderate 3 = severe	**Validity:** Correlated with SLEDAI 2K **Reliability:** Substantial agreement among raters **Responsiveness:** Limited data available	Not identified[c]
SRI	A composite outcome based on SELENA-SLEDAI score, BILAG, and PGA score Rated dichotomously as achieved or not achieved	**Validity[a]:** Correlated with measures of disease activity, biomarkers, and HRQoL **Reliability:** No information **Responsiveness[a]:** Less responsive than BILAG or physician VAS for musculoskeletal SLE	N/A
SFI	Classifies flares as mild/moderate or severe, based on criteria of clinical activity,	**Validity:** Criteria were developed based on consensus of the SELENA trial investigators. With the BILAG	N/A

Outcome measure	Type	Conclusions about measurement properties	MID
	need for additional treatment, or PGA score	2004 flare index, 52% agreement observed. Most variability is in the scoring of mild and moderate flares (less so for severe flares). **Reliability:** Fair agreement among raters[d] **Responsiveness:** No information	
SDI	Score of organ damage Damage is defined as irreversible change in an organ system, regardless of cause, that has occurred since the onset of SLE, and is present for at least 6 months Consists of 42 items in 12 domains, with a maximum score of 46 (higher scores denote more damage) At SLE diagnosis, the SDI score is 0 by definition. Damage is considered if the SDI score is ≥ 1.	**Validity:** Higher scores found in patients with damage versus stable disease and in patients with active versus inactive disease; predictor of mortality; low correlation observed with SLEDAI and BILAG, although one study found strong correlation with SLEDAI **Reliability:** Moderate agreement among raters **Responsiveness:** Scores have been shown to increase with disease duration.	Not identified
FACIT-Fatigue	Questionnaire completed by patients to assess fatigue during the past 7 days; consists of 13 statements, each rated on a 4-point Likert scale	**Validity:** Differentiated between groups defined on BILAG General domain and BILAG Musculoskeletal domain. Correlated with SF-36, Brief Pain Inventory, and patient global assessment. Weak to moderate correlation with PGA score. Weak correlation with BILAG and SELENA-SLEDAI score. **Reliability:** No information **Responsiveness:** Responsive to clinical improvement but not clinical deterioration	Anchor-based MIDs: 2.5 to 8.4 points. Distribution-based MIDs: 1/3 SD: 3.8 to 4.6 ½ SD: 5.8 to 6.8 SEM: 2.7 to 2.9

BILAG = British Isles Lupus Assessment Group; FACIT-Fatigue = Functional Assessment of Chronic Illness Therapy–Fatigue Scale; HRQoL = health-related quality of life; ICC = intraclass correlation coefficient; MID = minimal important difference; N/A = not applicable; PGA = Physician Global Assessment; SELENA-SLEDAI = Safety of Estrogens in Lupus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SD = standard deviation; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SEM = standard error of the mean; SF-36 = Short Form (36) Health Survey; SFI = System Lupus Erythematosus Flare Index; SLE = systemic lupus erythematosus; SRI = Systemic Lupus Erythematosus Responder Index; VAS = visual analogue scale.

[a] The SRI in the studies used the SELENA 2K rather than SELENA-SLEDAI, or used the SELENA-SLEDAI with proteinuria scored according to the SELENA 2K.

[b] A major clinical response has been defined as BILAG C score or better at six months, with no new BILAG A or B scores, and maintenance of response as no new BILAG A or B scores at six to 12 months. However, it is unclear as to how this definition was developed.

[c] The following definitions are found in the literature: No worsening of PGA: increase of less than 0.3 points (based on patients with rheumatoid arthritis); Mild or moderate flare: increase in PGA score of 1 or greater; Severe flare: increase in PGA score of greater than 2.5.

[d] Unclear whether this was based on the SFI or the revised SFI, which is based on organ system and does not include the SLEDAI. (Note: the revised SFI was not used in BLISS SC.)

Safety of Estrogens in Lupus Erythematosus National Assessment–SLE Disease Activity Index

The SLEDAI is a measure of disease activity that was derived by consensus among experts in rheumatology, followed by regression models to assign relative weights to each parameter.[38] The SELENA-SLEDAI is a modified version of the original SLEDAI. The modifications were made by a committee of the SELENA trials (Safety of Estrogen in Lupus Erythematosus National Assessment, which were DB randomized trials of hormone replacement therapy in post-menopausal women with lupus, and oral contractive therapy in pre-menopausal women).[39] No changes were made to the actual descriptors or weights; however, the definitions for rash, mucosal ulcers, and alopecia were changed to capture ongoing disease activity, whereas in the original version they were scored only if they presented as new or recurrent issues.[38] A second change was that either objective or subjective findings could be used to score a descriptor as present.[38]

The SELENA-SLEDAI consists of 24 items (nine organ systems) that are scored based on presence at the time of visit or in the preceding 10 days.[12] The scoring is additive across items, with a possible total score range of 0 to 105 (higher scores represent greater disease activity).[38] The tool can be used to calculate overall disease activity, as well as disease activity in a specific organ system.[38] A physician is required to complete the assessment. The following 24 items are included in the tool and are weighted in the total score:[12]

- Seizure, psychosis, organic brain syndrome, visual disturbance, cranial nerve disorder, lupus headache, cerebrovascular accident, vasculitis, arthritis, myositis, urinary casts, hematuria, proteinuria, pyuria, rash, alopecia, mucosal ulcers, pleurisy, pericarditis, low complement, increased DNA binding, fever, thrombocytopenia, and leukopenia.

Observational studies have demonstrated the validity, reliability, and sensitivity to change of the original SLEDAI.[38] It is predictive of mortality within a six-month period and has been shown to have substantial inter-rater reliability (correlation 0.61 to 0.80).[38] Another version of SLEDAI, the SLEDAI 2K, has been used in clinical trials. This version was introduced in 2002 as a modification of the original SLEDAI, to allow for persistent disease activity in the descriptors of rash, alopecia, mucosal ulcers, and proteinuria.[38] The SLEDAI 2K has been validated against the original SLEDAI.[38] There is less psychometric performance data available for the SELENA-SLEDAI.[38] Lower scores for health-related quality of life at baseline were correlated with higher disease activity on the SELENA-SLEDAI.[38] In a phase IIb trial that randomized 547 patients with SLE to blisibimod or placebo, SELENA-SLEDAI was moderately correlated with PGA (Spearman correlation coefficient 0.394, $P < 0.001$) and strongly correlated with BILAG (Spearman correlation coefficient 0.595, $P < 0.001$).[40] Based on patients in the SELENA trials (10 patients and six physician raters), the reliability (i.e., agreement among raters) of the SELENA-SLEDAI was almost perfect (intraclass correlation coefficient [ICC] = 0.87).[39]

A reduction from baseline in the SELENA-SLEDAI score of four points or more has been defined as clinically meaningful; however, this is not supported by the literature.[41] In a study that evaluated 230 patients with SLE from the University of Toronto Lupus Clinic, a clinician who was not familiar with the SLEDAI assigned a clinical activity score based on each patient's record, as 0 = no activity; 1 = mild activity with no therapeutic intervention; 2 = activity, but improvement from a previous visit; 3 = persistent activity/refractory to treatment; and 4 = flare.[42] The SLEDAI score was also calculated for each patient based on clinical and laboratory data.[42] In this cohort of patients, a median SLEDAI score of 2 corresponded to no activity, 4 to mild activity, 6 to activity but improvement from a previous

visit, and 8 to persistent activity/refractory to treatment and flare.[42] A median decrease of two points in the SLEDAI score corresponded to improvement from a previous visit.[42] A median increase of four points between assessments corresponded to flare.[42] In a study by the ACR, a pool of 310 patients were assessed with various SLE disease activity instruments, including the SELENA-SLEDAI.[43] From this pool, 15 cases were chosen and presented to 88 international experts on SLE, who were blinded to the disease activity assessments. The experts rated the patient cases as worsened, improved, or unchanged relative to the previous visit.[43] A clinically meaningful difference was estimated based on the change in disease activity score that corresponded to a 0.70 or greater probability that experts would rate the patient as improved or worsened. For the SELENA-SLEDAI, a clinically meaningful improved score was estimated to be a decrease of seven points, and a clinically meaningful worsened score as an increase of eight points.[43] The first and second study differed in quantifying a meaningful improvement (a decrease of two points in the first study and of seven points in the second study). The decrease of seven points is based specifically on the SELENA-SLEDAI and more clinical experts, whereas the decrease of two points is based on the original SLEDAI with what appears to be a single clinical expert. Therefore, a decrease of seven points for a clinically meaningful improvement is most evidence-based for the SELENA-SLEDAI.

Some issues have been identified with the SLEDAI. In comparison with BILAG, the SLEDAI is less responsive to change; it does not capture improvement or worsening; and it does not assess severity in an organ system.[38] A summary score to describe disease activity masks the underlying organ systems that are contributing to the score (i.e., the same score could represent multiple mild disease in many organs or severe disease is a single organ; or a score may not change despite worsening in one organ system if there is also improvement in another system).[16]

British Isles Lupus Assessment Group

The BILAG measures disease activity due to SLE across eight organ systems.[38] A physician is required to complete the assessment, and formal training is needed to ensure optimal performance.[38] Each organ system is scored on an ordinal scale from A to E, where A = most active, B = moderate activity, C = minor activity, D = stable, and E = never present.[39] A computer program is needed to conduct the scoring (e.g., The British Lupus Integrated Prospective System [BLIPS] is a computerized program that calculates BILAG scores).[38] Usually, a total score is not calculated, although a score can be calculated by assigning points to the ordinal scale as follows: A = 9, B = 3, C = 1, D = 0, and E = 0 (total minimum score 0 and maximum score 72).[39] The BILAG can be used to identify flares, with severe flare defined as a new organ domain score of A and moderate flare as a new organ domain score of B.[41]

The following organ systems are evaluated as part of the BILAG assessment, involving symptoms over the previous four weeks:[12]

- General: pyrexia, unintentional weight loss > 5%, lymphadenopathy/splenomegaly, fatigue/malaise/lethargy, anorexia/nausea/vomiting
- Mucocutaneous: maculopapular eruption, active discoid lesions, alopecia, panniculitis, angioedema, mucosal ulceration, malar erythema, SC nodules, perniotic skin lesions, periungual erythema, swollen fingers, sclerodactyly, calcinosis, telangiectasis
- Neurological: deteriorating level of consciousness, psychosis or confusional state, seizures, stroke or stroke syndrome, aseptic meningitis, mononeuritis multiplex, ascending or transverse myelitis, peripheral or cranial neuropathy, disc swelling/cytoid

bodies, chorea, cerebellar ataxia, severe unremitting headache, organic depressive illness, organic brain syndrome, episodic migraine headache

- Musculoskeletal: definite myositis, severe polyarthritis with loss of function, arthritis, tendonitis, mild chronic myositis, arthralgia, myalgia, tendon contractures and fixed deformity, aseptic necrosis
- Cardiorespiratory: pleuropericardial pain, dyspnea, cardiac failure, friction rub, pericardial or pleural effusion, chest pain, progressive chest X-ray changes in lung fields or heart size, evidence on electrocardiogram of pericarditis or myocarditis, cardiac arrhythmias, pulmonary function failure of more than 20%, cytohistological evidence of inflammatory lung disease
- Vasculitis: major cutaneous vasculitis, major abdominal crisis due to vasculitis, first or recurrent thromboembolism, excluding stroke, Raynaud's syndrome, livedo reticularis, superficial phlebitis, minor cutaneous vasculitis
- Renal: increased systolic and diastolic blood pressure, accelerated hypertension, positive dipstick value, urinary protein, proteinuria (note, in BLISS SC, spot urine protein:creatinine ratio was used to determine proteinuria), nephrotic syndrome, plasma or serum creatinine, creatinine clearance or glomerular filtration rate, active urinary sediment, histological evidence of active nephritis within three months
- Hematology: abnormal hemoglobin, total white cell count, neutrophils, lymphocytes, platelets, evidence of active hemolysis, positive Coomb's test, evidence of circulating anticoagulant

The BILAG has been found to be valid, reliable, and sensitive to change over time.[38] It is correlated with disease activity measures, especially the SLEDAI.[38] A major clinical response has been defined as BILAG C score or better at six months, with no new BILAG A or B scores, and maintenance of response as no new BILAG A or B scores at six to 12 months.[38] Hay et al. conducted validity and inter-rater reliability studies of the BILAG.[44] In the validity study, 353 patients with SLE were included.[44] Patients were assessed at intervals of at least one month apart over a 12-month period, and at least two BILAG assessments were conducted on each patient. Criterion validity was based on the gold standard of initiation or increase in disease-modifying therapy (i.e., corticosteroids or immunosuppressants). Construct validity was tested by comparing BILAG assessment with erythrocyte sedimentation rate (ESR), dsDNA antibody titres, and need for hospitalization. Compared with the gold standard, BILAG had 87% sensitivity, 99% specificity, and 80% positive predictive value for BILAG A score in any system.[44] The positive predictive values for a BILAG A score by organ system were the following: general = 83%, mucocutaneous = 82%, neurological = 30%, musculoskeletal = 81%, cardiorespiratory = 100%, vasculitis = 100%, renal = 100%, and hematology = 50%.[44] Construct validity was also demonstrated. The inter-rater reliability study included 82 patients with SLE treated at outpatient clinics. Two rheumatologists who were experienced with the BILAG assessed each patient (renal and hematological systems were not scored because they are based on laboratory results and not prone to inter-rater measurement error). The weighted statistics for inter-rater reliability (kappa) showed substantial to almost-perfect agreement between assessors (general = 0.79, mucocutaneous = 0.80, neurological = 0.72, musculoskeletal = 0.85, cardiorespiratory = 0.97, and vasculitis = 0.76).[44]

In a cross-sectional study of 141 patients with SLE in outpatient clinics between July 1994 and February 1995, assessments were conducted with the BILAG, Patient and Physician

Global Assessment, and blood sampling (i.e., ESR, C3, and anti-dsDNA antibodies).[45] The BILAG (individual domains and total) was found to be weakly to moderately[46] correlated with patient global assessment (Spearman rank correlation coefficients: general = 0.50, neurological = 0.23, musculoskeletal = 0.27, vasculitis = 0.23, number of A scores = 0.25, number A and B scores = 0.26, total score = 0.40, P < 0.01).[45] The BILAG was also weakly to moderately correlated with PGA (general = 0.43, neurological = 0.24, musculoskeletal = 0.27, number of A scores = 0.25, number A and B scores = 0.37, total score = 0.47, P < 0.01).[45] The BILAG hematological system was moderately correlated with ESR (Spearman rank order correlation coefficient = 0.34).[45] The BILAG renal system was weakly correlated with C3 (–0.26).[45] There were no statistically significant correlations with dsDNA.[45] In a phase IIb trial that randomized 547 patients with SLE to blisibimod or placebo, BILAG was found to be moderately correlated with PGA (Spearman correlation coefficient 0.447, P < 0.001).[40]

The BILAG assessment may be affected by the presence of sleep disorders, depression, and fibromyalgia.[38] Since the original BILAG, an update (BILAG 2004) is available (the updated version was not used in BLISS SC). In the new version, vasculitis is removed and two systems, ophthalmic and abdominal, are added.[38]

Physician Global Assessment

The PGA is a visual analogue scale (VAS) with equal markings between 0 to 3 (on a 10 cm scale).[12] In response to the question, "How do you assess your patient's current disease activity?," a mark is placed on the line, with 0 = none, 1 = mild, 2 = moderate, and 3 = severe.[12] The scoring physician should score the PGA before scoring the SELENA-SLEDAI and BILAG; otherwise, the PGA becomes less global and less sensitive.[47]

In 201 patients with SLE, there was strong correlation between a 0 to 3 PGA scale and the SLEDAI 2K (Pearson correlation coefficient, r = 0.553, P < 0.0001).[48] The PGA has also been shown to correlate with SELENA-SLEDAI.[41] There was substantial inter-rater agreement when the PGA was applied to patients in the SELENA trials (ICC = 0.75).[39] Limited data were available on the responsiveness of the PGA. In a small study of 20 patients with SLE who presented with inflammatory musculoskeletal symptoms, clinical assessments included PGA (a musculoskeletal VAS) and 28 tender and swollen joint counts (by ultrasound).[49] Effect sizes from baseline to two or four weeks were calculated from paired nonparametric tests (effect size r = Z statistic/sqrt[2N]). Physician musculoskeletal VAS had large effect sizes (i.e., large degree of change) at week 0 to 2 (r = –0.603, P = 0.001) and week 0 to 4 (–0.593, P < 0.001), suggesting improvement.[49] The ultrasound assessments also demonstrated improvement at weeks 0 to 2 and 0 to 4, with large effect sizes.[49]

The PGA is part of the SRI and SFI. In the SRI, no worsening of PGA is defined as an increase of less than 0.3 points.[41] The change of 0.3 points on the PGA is based on patients with rheumatoid arthritis.[41] In the SFI, mild or moderate flare can occur with an increase in PGA score of 1 or greater, and a severe flare with an increase in PGA score of greater than 2.5.[39] Through consensus, the Hopkins Lupus Center chose a one-point change on the PGA over the last 93 days as a gold standard definition of flare.[39] Based on this definition, moderate flares were defined as a score of 2 to 2.5, and severe flares as a score of 3.[39]

Systemic Lupus Erythematosus Responder Index

The SRI is a composite outcome that is rated dichotomously, as to whether a patient has achieved or not achieved response. In BLISS SC, SRI was achieved if all of the following criteria were met:[15]

- four point or greater reduction from baseline in SELENA-SLEDAI score, and
- no worsening (increase of < 0.3 points from baseline) in PGA, and
- no new BILAG A organ domain score or two new BILAG B organ domain scores compared with baseline.

The SRI was developed from an exploratory analyses of a phase II belimumab trial (LBSL99), which included 449 patients with SLE over 56 weeks.[41] According to the developers of the SRI, the SELENA-SLEDAI component was incorporated to capture global improvement, the BILAG domain to ensure no significant worsening in unaffected organ systems, and the PGA to ensure that improvements in disease activity are not at the expense of a patient's overall condition that are not captured with the SELENA-SLEDAI or BILAG.[41] It is unclear how these particular outcomes for the composite were chosen amid other outcomes available for SLE. In the article that describes the development of the SRI, it is stated that components of disease-activity indices were evaluated in a subset of 71.5% patients from the phase II trial who responded better to belimumab than placebo (ANA ≥ 1:80 and/or anti-dsDNA antibodies ≥ 30 IU/mL at screening and baseline).[41] Positivity for ANA and/or anti-dsDNA were also part of the inclusion criteria for BLISS SC. The restriction of the trial population and development of the composite outcome involving these patients may have increased the likelihood of a positive effect of belimumab.

Studies have found that the SRI is correlated with other clinical parameters of disease activity. In a post hoc analysis of pooled data from BLISS 52 and BLISS 76, responders and nonresponders on the SRI were compared on SELENA-SLEDAI score, BILAG score, changes in corticosteroid dose, normalization of anti-dsDNA, C3 and C4 biomarkers, PGA, risk of SFI flares, FACIT-Fatigue, and SF-36.[50] Of 1,684 patients, 761 were SRI responders and 923 were nonresponders at week 52. As expected based on the definition of the SRI, more responders than nonresponders achieved a reduction in SELENA-SLEDAI score of four points or more (100% versus 3.8%), a reduction in SELENA-SLEDAI score of seven points or more (40.3% versus 1.3%), higher number of improved organ domains per patient using the SELENA-SLEDAI and BILAG, and greater improvement in PGA scores; as well, more responders had no worsening of PGA scores.[50] More responders had one or less BILAG B score compared with nonresponders (91.9% versus 35.9%).[50] The risk of any flare or severe flare, based on the SFI, was lower in responders than nonresponders (any flare HR = 0.58 [95% CI, 0.52 to 0.65] and severe flare HR = 0.13, [95% CI, 0.09 to 0.17]).[50] In addition, more responders had reductions in prednisone (or equivalent) dosage of 25% or more to less than 7.5 mg per day (25.5% versus 16.4%), and fewer had dose increases to more than 7.5 mg per day (4.1% versus 21.3%).[50] Normalization of biomarkers occurred in a slightly larger percentage of responders than nonresponders (normalization of anti-dsDNA: 14.4% versus 10.8%, normalization of C3: 30.5% versus 25.3%, and normalization of C4: 31.7% versus 29.6%).[50] For the measures of health-related quality of life, SRI responders had greater improvements on the SF-36 PCS and MCS (4.9 versus 2.6 and 4.4 versus 1.7, respectively).[50] SRI responders also had greater improvements on the FACIT-Fatigue score compared with nonresponders (5.2 versus 3.0).[50]

In another post hoc analysis of pooled data from two 52-week phase IIb trials of sifalimumab and anifrolumab in 736 patients with SLE, changes in disease measures according to SRI responder status were assessed.[51] The SRI in this study used the SLEDAI 2K rather than the SELENA-SLEDAI. Compared with nonresponders, more SRI responders demonstrated a reduction in SLEDAI 2K of seven points or more (P < 0.001); had a greater mean change from baseline in SLEDAI 2K score (P < 0.001), PGA score (P = 0.019), FACIT-Fatigue score, and SF-36 score (P < 0.001); had more organ domains with improvement in SLEDAI 2K (P < 0.0001); experienced reduction in prednisone equivalent dosage to 7.5 mg per day or less (P < 0.001); had 50% or greater improvement in swollen and tender joint counts (P < 0.001), and 50%or more improvement in the Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI) (P < 0.001).[51] In addition, fewer SRI responders experienced one or more flares, as measured by BILAG A or 2B, compared with nonresponders (P < 0.001).[51] SRI responders had greater mean change from baseline in anti-dsDNA compared with nonresponders (P = 0.051), although no statistical difference was observed for C3 and C4 concentrations.[51]

Among 91 patients from the Oklahoma Lupus Cohort study, SRI was compared with a physician's assessment of improvement.[52] The SRI in this study used the SELENA-SLEDAI, except that the scoring for proteinuria was based on the SLEDAI 2K. Physicians rated patient's disease as either clinically significant improvement, worsening, or no change. In relation to these assessments, the SRI had a sensitivity of 85% and specificity of 74%.[52] In a small study of 20 patients with SLE who presented with inflammatory musculoskeletal symptoms, clinical and ultrasound parameters were compared at two and four weeks from baseline.[49] The SRI in this study used the SLEDAI 2K rather than SELENA-SLEDAI. Effect sizes from baseline to two or four weeks were calculated from paired nonparametric tests (effect size r = Z statistic/sqrt[2N]).[49] Among SRI responders, large effect sizes were observed for tender joint counts and swollen joint counts (r = –0.505 and –0.492, P = 0.024 and 0.028, respectively) and smaller, nonsignificant, effect sizes in nonresponders (r = –0.365 and –0.331, and P = 0.122 and 0.160, respectively). However, the SRI was found to be less responsive to musculoskeletal SLE (e.g., SRI underestimated response as there was objective improvement in synovitis among patients classified as nonresponders), than the BILAG or a physician VAS.[49]

Modified SLE Flare Index

The SFI is used to identify and classify flares as mild/moderate or severe, based on clinical activity, need for additional treatment, or PGA score.[41] The original definitions of mild/moderate and severe flares were reached by consensus of the investigators of the SELENA trials.[39] In BLISS SC, a modified form of the SFI was used, in which the modification excluded SELENA-SLEDAI score greater than 12 from the definition of severe flare, because this may indicate only modest increase in disease activity.[15]

In BLISS SC, mild/moderate flare and severe flare were defined according to the following criteria:[12]

- Mild or moderate flare:
  - change in SELENA-SLEDAI score of three points or more (not up to 12 points); or
  - new or worse discoid, photosensitive, profundus, cutaneous vasculitis, bullous lupus, nasopharyngeal ulcers, pleuritis, pericarditis, arthritis, or SLE fever; or
  - increase in prednisone dosage (not greater than 0.5 mg/kg per day); or
  - added NSAID or hydroxychloroquine for SLE activity; or

- or more increase in PGA score (not greater than 2.5).

- Severe flare:
  - change in SELENA-SLEDAI score greater than 12 points; or
  - new or worse CNS SLE, vasculitis, nephritis, myositis, platelet count less than 60,000, hemolytic anemia (Hb < 70 g/L or decrease in Hb > 30 g/L) that requires a doubling of prednisone dose, prednisone increase greater than 0.5 mg/kg/day, or hospitalization; or
  - increase in prednisone dosage to 0.5 mg/kg per day; or
  - new cyclophosphamide, azathioprine, methotrexate, or mycophenolate for SLE; or
  - hospitalization for SLE; or
  - increase in PGA score to 2.5 or higher.

A study evaluated the SFI using paper-based cases of patients with SLE.[53] Initially, 988 cases were assessed by three physicians for degree of flare or presence of disease activity and rated as severe, moderate, or mild flare, or persistent/ongoing disease. Cases for which there was agreement by the three physicians (N = 451 cases) were moved on the second part of the study and assessed by 18 pairs of physicians with three instruments, BILAG 2004 flare index, SFI, and revised SFI. (Note: the revised SFI is based on organ system and does not include the SLEDAI[38]; the instrument of relevance to this discussion is the SFI.) The assessments based on these instruments were compared with the assessments conducted initially in the first stage of the study by the three physicians. For the SFI, assessments matched the conclusions of the three physicians in 72% of cases (weighted kappa 0.59).[53] The discrepancies were concentrated in classifying moderate flares as severe flares, and identifying persistent activity as a flare.[53] There was also an issue of over-scoring due to classifying treatment change as a flare, even when there were no new or worsening clinical features.[53] The authors of this study indicate that "the problem of capturing lupus flare accurately is not completely solved."[53]

In a small study of 16 patients who were each evaluated by four physicians, there was 52% agreement between the SFI and BILAG 2004 flare index in classifying patients as having no flare, or mild, moderate, or severe flare.[54] It was unclear, however, whether this study used the SFI or the revised SFI. The agreement among raters on the SFI was fair (ICC = 0.21 [95% CI, 0.08 to 0.48]), and lower than the BILAG 2004 assessment of flares.[54]

Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index

The SDI was developed by the international collaboration Systemic Lupus International Collaborating Clinics (SLICC).[16] The purpose of the assessment is to score irreversible damage, regardless of cause. Damage is defined as irreversible change in an organ system that has occurred since the onset of SLE, and is present for at least six months.[16] The tool is completed by a physician and consists of 42 items in 12 domains, with a maximum score of 46 (higher scores denote more damage).[16,17] The items are rated as present or absent and, in the case of recurring events, a rating of two or three points to an item can be provided.[16] At diagnosis of SLE, the SDI score is 0 by definition.[17] Damage is considered if the SDI score is 1 or more.[17]

In the BLISS SC, the SDI score was based on the following items (each scored one point, except where indicated):[12]

- Ocular (any cataract ever, retinal change, or optic atrophy)
- Neuropsychiatric (cognitive impairment, seizures requiring therapy for six months, cerebrovascular accident ever [up to two points], cranial or peripheral neuropathy excluding optic, transverse myelitis)
- Renal (estimated or measured glomerular filtration rate < 50%, proteinuria > 3.5 g in 24 hours, or end-stage renal disease, regardless of dialysis or transplantation [up to three points])
- Pulmonary (pulmonary hypertension, pulmonary fibrosis, shrinking lung, pleural fibrosis, pulmonary infarction)
- Cardiovascular (angina or coronary artery bypass, myocardial infarction ever [up to two points], cardiomyopathy, valvular disease, pericarditis for six months or pericardiectomy)
- Peripheral vascular (claudication for six months, minor tissue loss, significant tissue loss ever [up to two points], venous thrombosis with swelling or ulceration or venous stasis)
- Gastrointestinal (infarction or resection of bowel below duodenum, spleen, liver, or gallbladder ever for any cause [up to two points], mesenteric insufficiency, chronic peritonitis, stricture, or upper gastrointestinal tract surgery ever)
- Musculoskeletal (muscle atrophy or weakness, deforming or erosive arthritis, osteoporosis with fracture or vertebral collapse, avascular necrosis [up to two points], osteomyelitis)
- Skin (scarring chronic alopecia, extensive scarring or panniculus other than scalp and pulp space, skin ulceration excluding thrombosis for more than six months)
- Premature gonadal failure
- Diabetes
- Malignancy excluding dysplasia (up to two points)

The SDI has been demonstrated to be reliable and sensitive to change (i.e., SDI values increase with disease progression), and to predict mortality.[16] To assess the validity of the SDI, centres that treated patients with SLE submitted two assessments, five years apart, on two patients with active disease (one patient with increase in damage over the five years and one patient with stable damage) and two patients with inactive disease (one patient with increase in damage and one patient with stable damage).[55] The cases (14 cases in three separate packages) were written up in a uniform format and sent back out, in mixed order, to the centres, where the SDI was completed by 20 physicians (two assessments per patient at time 1 and time 2). The SDI scores of patients with damage after five years were increased by a greater degree than those of patients with stable disease (2.08 points versus 0.24 points).[55] The SDI scores of patients with active disease also increased more than those of patients with inactive disease (1.48 points versus 0.83 points).[55]

In a small study (N = 10 patients with SLE), six physicians from five countries assessed patients on the SDI and SLEDAI, to determine the correlation between the instruments and the reliability of the SDI for physicians from different regions of the world.[56] The SDI used in this study was modified in three aspects: addition of “lung resection not for malignancy” rather than “pulmonary infarction;” addition of “pancreatic insufficiency requiring enzyme replacement,” and addition of “ruptured tendons.”[56] No correlation was found between the SDI and SLEDAI (correlation coefficient 0.05), indicating that disease damage is not

correlated with disease activity (although other studies have found an association, as described in this appendix).[56] Also, no statistically significant differences were found between the ratings of physicians.[56] In a cross-sectional study of 80 patients, statistically significant associations were found between SDI and disease flares, as assessed by an increase in SLEDAI 2K of more than three points; disease severity, as assessed by the presence of class II/IV glomerulonephritis, CNS involvement, or administration of IV cyclophosphamide; and antiphospholipid antibodies.[57] A study of 71 patients found that the SDI was associated with SLEDAI 2K ($r = 0.742$, $P < 0.001$) and disease duration ($P = 0.007$).[58] The SDI and BILAG have been found to have weak correlation (Spearman correlation coefficient 0.19).[59]

The SDI is a statistically significant predictor of clinically important outcomes. In a 10-year retrospective study of 80 patients with SLE, the mean SDI renal damage score at one year after diagnosis was a significant predictor of end-stage renal failure (at one year: renal failure versus no renal failure, SDI renal damage score 0.33 versus 0.03; at five years: SDI renal damage score 1.33 versus 0.14; at 10 years: SDI renal damage score 2.80 versus 0.35).[60] The total SDI score was also associated with end-stage renal failure at five and 10 years.[60] The SDI pulmonary damage score at one year after diagnosis was a significant predictor of death within 10 years; however, total SDI score was not associated with death.[60] More recent studies with larger cohorts of patients have shown that the SDI is a predictor of mortality. Patients with SLE (N = 1,297) were identified within two years of a first clinical visit from eight centres, and followed for two, five to 10, and more than 10 years.[17] The SDI increased over time and was found to be higher among patients who died.[17] In the University of Toronto Lupus Clinic, 263 patients were followed for 10 years.[61] Within 10 years, 25% of patients who exhibited damage at the first SDI assessment (i.e., one year after diagnosis) died, compared with 7.3% of patients who had no early signs of damage.[61]

Among 20 SLICC members who completed the SDI on 42 cases, there was moderate agreement between raters (ICC = 0.553).[17] Similarly, when the SDI was completed by another physician based on retrospective review of patient cases, interobserver reliability was moderate (kappa 0.47 [95% CI, 0.28 to 0.66]).[59] An MID was not identified for the SDI.

Functional Assessment of Chronic Illness Therapy–Fatigue Score

The FACIT-Fatigue is completed by patients to assess fatigue. In BLISS SC, FACIT-Fatigue version 4 was used.[12] Patients were presented with a list of 13 statements and asked to rate each on a four-point Likert scale (0 = not at all, 1 = a little bit, 2 = somewhat, 3 = quite a bit, and 4 = very much), to indicate how true the statement was during the past seven days.[12] Examples of statements were "I feel fatigued" and "I feel weak all over." In the scoring, the numbers are reversed so that higher scores denote better quality of life (i.e., 4 = not at all, 3 = a little bit, 2 = somewhat, 1 = quite a bit, and 0 = very much). For statements 7 ("I have energy") and 8 ("I am able to do my usual activities"), the scores are not reversed. The total score is a simple sum across the statements, with a possible range of 0 to 52.

The FACIT-Fatigue was validated in patients with SLE by Lai et al.[62] Patients with moderately to severely active extrarenal SLE (N = 254) completed the FACIT-Fatigue, SF-36, Brief Pain Inventory, and a patient global assessment VAS at baseline, week 12, week 24, and week 52.[62] Physicians also completed the BILAG and PGA at the same visits. The FACIT-Fatigue could differentiate between groups that were defined by BILAG General domain ratings at 12 weeks (A/B score: mean [SD] FACIT-Fatigue 17.5 [10.3] versus C/D/E

score: 25.9 [13.0], P = 0.001).[62] The FACIT-Fatigue scores also differed by BILAG musculoskeletal domain ratings at 12 weeks (A score: mean [SD] FACIT-Fatigue 18.8 [12.5] versus B score: 21.4 [11.5] versus C/D/E score: 26.8 [13.2], P = 0.003).[62] The FACIT-Fatigue was moderately to strongly correlated with the SF-36 (Spearman correlation coefficient 0.69 to 0.87 at week 52, and 0.42 to 0.70 for change from baseline to week 52), Brief Pain Inventory (–0.72 to –0.82 at week 52, and –0.47 to –0.62 for change from baseline to week 52) and patient global assessment (–0.76 at week 52 and –0.56 for change from baseline to week 52).[62] However, the correlations of FACIT-Fatigue with total BILAG score and PGA were weak (Spearman correlation coefficient for BILAG: –0.25 at week 52 and –0.15 for change from baseline to week 52; Spearman correlation coefficient for PGA: –0.21 at week 52 and –0.12 for change from baseline to week 52).[62] In a phase IIb trial that randomized 547 patients with SLE to blisibimod or placebo, FACIT-Fatigue was weakly to moderately correlated with PGA (Spearman correlation coefficient –0.32, P < 0.001), SELENA-SLEDAI (–0.13, P = 0.006), and BILAG (–0.18, P < 0.001).[40] The FACIT-Fatigue was responsive to clinical improvement but not clinical deterioration.[63]

The study by Lai et al. included estimation of MIDs for the FACIT-Fatigue with anchor and distribution-based techniques.[62] The anchors were based on the general and musculoskeletal domains of the BILAG. These were selected as anchors for the FACIT-Fatigue because the general domain contains physician assessment of fatigue and malaise, and the musculoskeletal domain contains assessment of pain, which is associated with fatigue.[62] The anchor-based MIDs were estimated from cross-sectional (i.e., comparing mean FACIT-Fatigue scores across groups defined by BILAG disease activity at each assessment) and longitudinal analyses (i.e., changes in FACIT-Fatigue with changes in BILAG disease activity between consecutive assessments).[62] Changes in BILAG disease activity were classified as more active, less active, or stable (with stable defined as change from BILAG D/E to C or vice versa).[62] The anchor-based MIDs ranged from 2.5 to 8.4 points.[62] The distribution-based MIDs fell within this range (based on 1/3 SD: 3.8 to 4.6 points; ½ SD: 5.8 to 6.8 points; SEM: 2.7 to 2.9 points).[62]

IV Belimumab Studies (BLISS 52, BLISS 76, and Extensions)

Short Form (36) Health Survey

Although several measures of health-related quality of life have been studied in SLE, the most commonly used and accepted measure is the SF-36, a generic measure that is applicable to a variety of conditions and chronic diseases, including SLE.[63,64] The SF-36 differs across disease activity categories; however, studies have suggested that it has poor responsiveness in SLE.[63] Minimum important differences that are specific to SLE have been estimated.[63] For the summary scores, anchor-based MIDs range from 2.1 to 2.4. Based on an anchor of patient-self report as "better" or "worse" in response to the question, "How would you describe your overall status since your last visit?" the MID for the PCS of the SF-36 was 2.1 (for better) and –2.2 (for worse), and for the MCS the MID was 2.4 (for better) and –1.2 (for worse).[65]

References

1. Center for Drug Evaluation Research. Medical review(s). Benlysta (belimumab) subcutaneous injection. Company: Human Genome Sciences, Inc. Application No.: 761043. Approval date: 07/20/2017 (FDA approval package). Rockville (MD): U. S. Food and Drug Administration (FDA); 2017: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/761043Orig1s000MedR.pdf. Accessed 2019 Jul 3.
2. Clinical Study Report: BEL112233. A multi-center, continuation trial of belimumab (HGS1006, LymphoStat-B), a fully human monoclonal anti-BLyS antibody, in subjects with systemic lupus erythematosus (SLE) who completed the phase III protocol HGS1006-C1056 in the United States - interim. [**CONFIDENTIAL** internal sponsor's report]. Mississauga (ON): GlaxoSmithKline Inc.; 2015 Sep 18.
3. Lee YH, Song GG. Comparative efficacy and safety of intravenous or subcutaneous belimumab in combination with standard therapy in patients with active systemic lupus erythematosus: a Bayesian network meta-analysis of randomized controlled trials. *Lupus.* 2018;27(1):112-119.
4. Tian J, Luo Y, Wu H, Long H, Zhao M, Lu Q. Risk of adverse events from different drugs for SLE: a systematic review and network meta-analysis. *Lupus Sci Med.* 2018;5(1):e000253.
5. Lupus Canada. Lupus Canada. 2019; https://www.lupuscanada.org/. Accessed 2019 Aug 1.
6. Benlysta (belimumab): lyophilized powder for intravenous infusion 120 mg/5 mL, 400 mg/20 mL, solution for subcutaneous injection 200 mg/1 mL [product monograph]. Mississauga (ON): GlaxoSmithKline Inc.; 2017 Dec 7.
7. e-CPS. Ottawa (ON): Canadian Pharmacists Association; 2019: https://www.e-therapeutics.ca. Accessed 2019 Oct 17.
8. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline explanation and elaboration (PRESS E&E). *(CADTH methods and guidelines)*. Ottawa (ON): CADTH; 2016: https://www.cadth.ca/sites/default/files/pdf/CP0015_PRESS_Update_Report_2016.pdf. Accessed 2019 Jun 26.
9. Grey matters: a practical tool for searching health-related grey literature. Ottawa (ON): CADTH; 2019: https://www.cadth.ca/resources/finding-evidence/grey-matters. Accessed 2019 Jun 26.
10. 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(5):1016-1027.
11. Doria A, Stohl W, Schwarting A, et al. Efficacy and safety of subcutaneous belimumab in anti-double-stranded DNA-positive, hypocomplementemic patients with systemic lupus erythematosus. *Arthritis Rheumatol.* 2018;70(8):1256-1264.
12. CDR submission: Benlysta (belimumab), subcutaneous (SC) autoinjector [**CONFIDENTIAL** sponsor's submission]. Mississauga (ON): GlaxoSmithKline Inc.; 2019 May 29.
13. Health Canada reviewer's report: Benlysta (belimumab) [**CONFIDENTIAL** internal report]. Ottawa (ON): Therapeutics Products Directorate, Health Canada; 2018 Jan 4.
14. Center for Drug Evaluation Research. Statistical review(s). Benlysta (belimumab) subcutaneous injection. Company: Human Genome Sciences, Inc. Application No.: 761043. Approval date: 07/20/2017 (FDA approval package). Rockville (MD): U. S. Food and Drug Administration (FDA); 2017: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/761043Orig1s000StatR.pdf. Accessed 2019 Jul 3.
15. Clinical Study Report: BEL112341. A phase 3, multi-center, randomized, double-blind, placebo-controlled, 52-week study to evaluate the efficacy and safety of belimumab (HGS1006) administered subcutaneously (SC) to subjects with systemic lupus erythematosus (SLE) – double-blind endpoint analysis. [**CONFIDENTIAL** internal sponsor's report]. Mississauga (ON): GlaxoSmithKline Inc.; 2016 Mar 8.
16. Castrejon I, Tani C, Jolly M, Huang A, Mosca M. Indices to assess patients with systemic lupus erythematosus in clinical trials, long-term observational studies, and clinical care. *Clin Exp Rheumatol.* 2014;32(5 Suppl 85):S-85-95.
17. Romero-Diaz J, Isenberg D, Ramsey-Goldman R. Measures of adult systemic lupus erythematosus: updated version of British Isles Lupus Assessment Group (BILAG 2004), European Consensus Lupus Activity Measurements (ECLAM), Systemic Lupus Activity Measure, Revised (SLAM-R), Systemic Lupus Activity Questionnaire for Population Studies (SLAQ), Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K). *Arthritis Care Res.* 2011;63(Suppl 11):S37-S46.
18. Indirect treatment comparison of IV formulation and SC formulation of belimumab for the treatment of adult patients with active, autoantibody-positive systemic lupus erythematosus (SLE) who are receiving standard therapy. In: CDR submission: Benlysta (belimumab), subcutaneous (SC) autoinjector. Company: GlaxoSmithKline Inc. [**CONFIDENTIAL** sponsor's submission]. Mississauga (ON): GlaxoSmithKline Inc.; 2019 May 29.
19. Ramachandran S, Parks D, Kurtinecz M, Roth DA, Alfonso-Cristancho R. An indirect comparison of intravenous and subcutaneous belimumab efficacy in patients with SLE and high disease activity. *J Comp Eff Res.* 2018;7(6):581-593.
20. 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(9767):721-731.
21. 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(12):3918-3930.
22. 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(3):355-363.

23. 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(9):1168-1178.

24. Salanti G, Marinho V, Higgins JP. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. *J Clin Epidemiol.* 2009;62(8):857-864.

25. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *Int J Epidemiol.* 2013;42(1):332-345.

26. Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods.* 2012;3(2):98-110.

27. Clinical Study Report: HGS1006-C1057. A phase 3, multi-center, randomized, double-blind, placebo-controlled, 52-week study to evaluate the efficacy and safety of belimumab (HGS1006, LymphoStat-B), a fully human monoclonal anti-BLyS antibody, in subjects with systemic lupus erythematosus (SLE) Rockville (MD): Humn Genome Sciences, Inc.; 2010 Jan 21.

28. Clinical Study Report: HGS1006-C1056. A phase 3, multi-center, randomized, double-blind, placebo-controlled, 76-week study to evaluate the efficacy and safety of belimumab (HGS1006, LymphoStat-B), a fully human monoclonal anti-BLyS antibody, in subjects with systemic lupus erythematosus (SLE) Rockville (MD): Human Genome Sciences, Inc.; 2010 Mar 10.

29. Wallace DJ, Ginzler EM, Merrill JT, et al. Safety and efficacy of belimumab plus standard therapy for up to thirteen years in patients with systemic lupus erythematosus. *Arthritis Rheumatol.* 2019;16:16.

30. Furie RA, Wallace DJ, Aranow C, et al. Long-term safety and efficacy of belimumab in patients with systemic lupus erythematosus: a continuation of a seventy-six-week phase III parent study in the United States. *Arthritis Rheumatol.* 2018;70(6):868-877.

31. Strand V, Berry P, Lin X, Asukai Y, Punwaney R, Ramachandran S. Long-term impact of belimumab on health-related quality of life and fatigue in patients with systemic lupus erythematosus: six years of treatment. *Arthritis Care Res (Hoboken).* 2019;71(6):829-838.

32. Clinical Study Report: BEL112234. A multi-center, continuation trial of belimumab (HGS1006, LymphoStat-B), a fully human monoclonal anti-BLyS antibody, in subjects with systemic lupus erythematosus (SLE) who completed the phase 3 protocol HGS1006-C1056 or HGS1006-C1057 [**CONFIDENTIAL** internal sponsor's report]. Mississauga (ON): GlaxoSmithKline; 2017 Jul 31.

33. van Vollenhoven RF, Navarra SV, Levy RA, et al. Long-term safety and limited organ damage in patients with systemic lupus erythematosus treated with belimumab: a phase III study extension. *Rheumatology (Oxford).* 2019.

34. Clinical Study Report: 201223. Results report for 201223, a pooled analysis of BEL112233 and BEL112234, two long-term, extension studies of BEL110751 and BEL110752, to investigate long-term safety and organ damage accrual with belimumab treatment in systemic lupus erythematosus. [**CONFIDENTIAL** internal sponsor's report]. Mississauga (ON): GlaxoSmithKline Inc.; 2015 Nov 4.

35. Bruce IN, Urowitz M, van Vollenhoven R, et al. Long-term organ damage accrual and safety in patients with SLE treated with belimumab plus standard of care. *Lupus.* 2016;25(7):699-709.

36. Clinical Study Report: BEL112341 [extension]. Phase 3, multi-center, randomized, double-blind, placebo-controlled, 52-week study to evaluate the efficacy and safety of belimumab (HGS1006) administered subcutaneously (SC) to subjects with systemic lupus erythematosus (SLE) - open-label phase - end of study report [**CONFIDENTIAL** internal sponsor's report]. Mississauga (ON): GlaxoSmithKline Inc.; 2016 Sep 26.

37. Urowitz MB, Ohsfeldt RL, Wielage RC, Kelton KA, Asukai Y, Ramachandran S. Organ damage in patients treated with belimumab versus standard of care: a propensity score-matched comparative analysis. *Ann Rheum Dis.* 2019;78(3):372-379.

38. Mikdashi J, Nived O. Measuring disease activity in adults with systemic lupus erythematosus: the challenges of administrative burden and responsiveness to patient concerns in clinical research. *Arthritis Res Ther.* 2015;17:183.

39. Petri M, Buyon J, Kim M. Classification and definition of major flares in SLE clinical trials. *Lupus.* 1999;8(8):685-691.

40. Petri MA, Martin RS, Scheinberg MA, Furie RA. Assessments of fatigue and disease activity in patients with systemic lupus erythematosus enrolled in the phase 2 clinical trial with blisibimod. *Lupus.* 2017;26(1):27-37.

41. Furie RA, Petri MA, Wallace DJ, et al. Novel evidence-based systemic lupus erythematosus responder index. *Arthritis Rheum.* 2009;61(9):1143-1151.

42. Gladman DD, Urowitz MB, Kagal A, Hallett D. Accurately describing changes in disease activity in systemic lupus erythematosus. *J Rheumatol.* 2000;27(2):377-379.

43. The American College of Rheumatology response criteria for systemic lupus erythematosus clinical trials: measures of overall disease activity. *Arthritis Rheum.* 2004;50(11):3418-3426.

44. Hay EM, Bacon PA, Gordon C, et al. The BILAG index: a reliable and valid instrument for measuring clinical disease activity in systemic lupus erythematosus. *Q J Med.* 1993;86(7):447-458.

45. Stoll T, Stucki G, Malik J, Pyke S, Isenberg DA. Further validation of the BILAG disease activity index in patients with systemic lupus erythematosus. *Ann Rheum Dis.* 1996;55(10):756-760.

46. Cohen J. A power primer. *Psychol Bull.* 1992;112(1):155-159.

47. Luijten KM, Tekstra J, Bijlsma JW, Bijl M. The Systemic Lupus Erythematosus Responder Index (SRI); a new SLE disease activity assessment. *Autoimmun Rev.* 2012;11(5):326-329.

48. Narayanan S, Collins CE, Busch HM, Gaylis N, Hautamaki E, Chatham WW. Convergent validity of new disease assessment instruments (Dai) in systemic lupus erythematosus (SLE) in relation to Sledai-2k. *Value Health.* 2015;18(7):A677.

49. Mahmoud K, Zayat AS, Yusof Y, et al. Responsiveness of clinical and ultrasound outcome measures in musculoskeletal systemic lupus erythematosus. *Rheumatology.* 2019;03:03.

50. Furie R, Petri MA, Strand V, et al. Clinical, laboratory and health-related quality of life correlates of Systemic Lupus Erythematosus Responder Index response: a post hoc analysis of the phase 3 belimumab trials. *Lupus Sci Med.* 2014;1(1):e000031.

51. Furie R, Wang L, Illei G, Drappa J. Systemic Lupus Erythematosus (SLE) Responder Index response is associated with global benefit for patients with SLE. *Lupus.* 2018;27(6):955-962.

52. Thanou A, Chakravarty E, James JA, Merrill JT. Which outcome measures in SLE clinical trials best reflect medical judgment? *Lupus Sci Med.* 2014;1(1):e000005.

53. Isenberg D, Sturgess J, Allen E, et al. Study of flare assessment in systemic lupus erythematosus based on paper patients. *Arthritis Care Res (Hoboken).* 2018;70(1):98-103.

54. Isenberg DA, Allen E, Farewell V, et al. An assessment of disease flare in patients with systemic lupus erythematosus: a comparison of BILAG 2004 and the flare version of SELENA. *Ann Rheum Dis.* 2011;70(1):54-59.

55. Gladman D, Ginzler E, Goldsmith C, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheum.* 1996;39(3):363-369.

56. Gladman DD, Urowitz MB, Goldsmith CH, et al. The reliability of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index in patients with systemic lupus erythematosus. *Arthritis Rheum.* 1997;40(5):809-813.

57. Bonakdar ZS, Mohtasham N, Karimifar M. Evaluation of damage index and its association with risk factors in patients with systemic lupus erythematosus. *J Res Med Sci.* 2011;16 Suppl 1:S427-432.

58. Shariati-Sarabi Z, Monzavi SM, Ranjbar A, Esmaily H, Etemadrezaie H. High disease activity is associated with high disease damage in an Iranian inception cohort of patients with lupus nephritis. *Clin Exp Rheumatol.* 2013;31(1):69-75.

59. Castrejon I, Rua-Figueroa I, Rosario MP, Carmona L. Clinical composite measures of disease activity and damage used to evaluate patients with systemic lupus erythematosus: a systematic literature review. *Reumatol Clin.* 2014;10(5):309-320.

60. Stoll T, Seifert B, Isenberg DA. SLICC/ACR Damage Index is valid, and renal and pulmonary organ scores are predictors of severe outcome in patients with systemic lupus erythematosus. *Br J Rheumatol.* 1996;35(3):248-254.

61. Feld J, Isenberg D. Why and how should we measure disease activity and damage in lupus? *Presse Med.* 2014;43(6P2):e151-e156.

62. Lai JS, Beaumont JL, Ogale S, Brunetta P, Cella D. Validation of the functional assessment of chronic illness therapy-fatigue scale in patients with moderately to severely active systemic lupus erythematosus, participating in a clinical trial. *J Rheumatol.* 2011;38(4):672-679.

63. Izadi Z, Gandrup J, Katz PP, Yazdany J. Patient-reported outcome measures for use in clinical trials of SLE: a review. *Lupus Sci Med.* 2018;5(1):e000279.

64. Touma Z, Gladman DD, Ibanez D, Urowitz MB. Is there an advantage over SF-36 with a quality of life measure that is specific to systemic lupus erythematosus? *J Rheumatol.* 2011;38(9):1898-1905.

65. Colangelo KJ, Pope JE, Peschken C. The minimally important difference for patient-reported outcomes in systemic lupus erythematosus including the HAQ-DI, pain, fatigue, and SF-36. *J Rheumatol.* 2009;36(10):2231-2237.