AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

What Does It Mean to Be a British Isles Lupus Assessment Group–Based Composite Lupus Assessment Responder? Post Hoc Analysis of Two Phase III Trials

Richard Furie,[1] [ID] Eric F. Morand,[2] [ID] Ian N. Bruce,[3] David Isenberg,[4] [ID] Ronald van Vollenhoven,[5] Gabriel Abreu,[6] Lilia Pineda,[7] and Raj Tummala[7]

Objective. The British Isles Lupus Assessment Group–based Composite Lupus Assessment (BICLA) is a validated global measure of treatment response in systemic lupus erythematosus (SLE) clinical trials. To understand the relevance of BICLA in clinical practice, we investigated relationships between BICLA response and routine SLE assessments, patient-reported outcomes (PROs), and medical resource utilization.

Methods. This was a post hoc analysis of pooled data from the phase III, randomized, placebo-controlled, 52-week TULIP-1 (ClinicalTrials.gov identifier: NCT02446912; n = 457) and TULIP-2 (ClinicalTrials.gov identifier: NCT02446899; n = 362) trials of intravenous anifrolumab (150/300 mg once every 4 weeks) in patients with moderate-to-severe SLE. Changes from baseline to week 52 in clinical assessments, PROs, and medical resource use were compared in BICLA responders versus nonresponders, regardless of treatment assignment.

Results. BICLA responders (n = 318) achieved significantly improved outcomes compared with nonresponders (n = 501), including lower flare rates, higher rates of attainment of sustained oral glucocorticoid taper to ≤7.5 mg/day, greater improvements in PROs (Functional Assessment of Chronic Illness Therapy–Fatigue, Short Form 36 Health Survey), and fewer SLE-related hospitalizations/emergency department visits (all nominal $P < 0.001$). Compared with nonresponders, BICLA responders had greater improvements in global and organ-specific disease activity (Physician's Global Assessment, SLE Disease Activity Index 2000, Cutaneous Lupus Erythematosus Disease Area and Severity Index Activity, and joint counts; all nominal $P < 0.001$). BICLA responders had fewer lupus-related serious adverse events than nonresponders.

Conclusion. BICLA response is associated with clinical benefit in SLE assessments, PROs, and medical resource utilization, confirming its value as a clinical trial end point that is associated with measures important to patient care.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic rheumatic disease with several unmet needs, chief of which is the addition of safer and more efficacious therapies to available treatments. Recognizing the challenges of drug development in SLE that were facing the lupus community, the US Food and Drug Administration (FDA) issued draft guidance in 2005 and a final

guidance document in 2010 that emphasized that improvement in one domain of disease activity could not be accompanied by worsening in another (1). In response to the FDA draft guidance, the first composite index, the SLE Responder Index (SRI), was developed using data from the completed phase II belimumab study (2). The SRI comprises 3 components, with 1 component assessing improvement in disease activity (SLE Disease Activity Index 2000 [SLEDAI-2K] [3]) and the remaining 2 components assessing worsening (British Isles Lupus Assessment Group [BILAG] [4] and physician global assessment of disease activity [PhGA]). Shortly thereafter, the BILAG-based Composite Lupus Assessment (BICLA) was created based on similar principles (5,6).

The BICLA was developed following an expert panel review of disease activity indices used in SLE clinical trials (5,6). A BICLA response requires improvement in all domains affected at baseline, assessed by the BILAG 2004, no worsening of other BILAG 2004 domains, no worsening of SLEDAI-2K or PhGA scores compared with baseline, no initiation of non-protocol treatment or use beyond protocol-allowed thresholds, and no discontinuation of investigational product (5,6). Thus, in contrast to the SRI, the driver of improvement in the BICLA is the BILAG 2004, whereas worsening is assessed using the SLEDAI-2K and PhGA in addition to the BILAG 2004 (7). The BILAG 2004 weighs organ systems equally and distinguishes between inactive disease, partial or complete improvement, and deterioration of disease activity, while the SLEDAI-2K assigns weighted scores to its components and requires complete resolution of disease activity of the specific element to capture improvement (5,7,8).

BICLA response has been used as an end point in more than 20 SLE trials to date (5,9–14), including the phase II MUSE trial and the phase III TULIP-1 and TULIP-2 trials of anifrolumab, a human monoclonal antibody to type I interferon (IFN) receptor subunit 1 (12–14). BICLA response was a secondary end point in the MUSE and TULIP-1 studies and was the primary end point in the TULIP-2 study (12–14). BICLA response rate treatment differences of >16% between anifrolumab and placebo were observed at week 52 in all 3 studies.

Composite SLE assessment results incorporated as end points in clinical trials are not used in clinical practice, and thus the relevance of treatment response assessed in this way may not be appreciated by clinicians. We therefore investigated the relationship between BICLA response and other SLE disease measures that are meaningful in clinical practice, including flares, oral glucocorticoid daily dose and sustained oral glucocorticoid taper, patient-reported outcomes (PROs), medical resource utilization, and clinical and laboratory measures of global and organ-specific disease. These relationships were assessed between BICLA responders and nonresponders using pooled data from the phase III TULIP-1 and TULIP-2 trials of anifrolumab, regardless of treatment group assignment.

PATIENTS AND METHODS

Patients and study design. This was a post hoc analysis of pooled data from the phase III, randomized, placebo-controlled, double-blind, 52-week TULIP-1 and TULIP-2 trials, for which patient disposition and study details have been previously published (13,14). In brief, eligible patients were ages 18–70 years, fulfilled the American College of Rheumatology revised classification criteria for SLE (15), and had seropositive moderate-to-severe SLE despite standard treatment. Patients with active severe lupus nephritis or neuropsychiatric SLE were excluded. Patients were randomized to receive intravenous infusions of placebo or anifrolumab every 4 weeks for 48 weeks in addition to standard treatment (TULIP-1: placebo, anifrolumab 150 mg, or anifrolumab 300 mg [2:1:2]; TULIP-2: placebo or anifrolumab 300 mg [1:1]). Primary end points were assessed at week 52. Other treatments were stable throughout the trial except for those resulting from protocol-determined intent-to-taper oral glucocorticoids. For patients receiving oral glucocorticoids ≥10 mg/day (prednisone or equivalent) at baseline, an attempt to taper oral glucocorticoids to ≤7.5 mg/day was required between weeks 8 and 40; tapering was also permitted for patients receiving oral glucocorticoids <10 mg/day at baseline. Stable oral glucocorticoid dosage was required between weeks 40 and 52.

Study end points and assessments. BICLA response was defined as all of the following: reduction of all baseline BILAG 2004 A and B domain scores to B/C/D and C/D domains, respectively, and no worsening in other BILAG 2004 organ systems, as defined by ≥1 new BILAG 2004 A domain scores or ≥2 new BILAG 2004 B domain scores; no increase in SLEDAI-2K score (from baseline); no increase in PhGA score (≥0.3 points from baseline); no discontinuation of investigational product; and no use of restricted medications beyond protocol-allowed thresholds (5). Pooled data were analyzed according to the TULIP-2 restricted medication analytical rules to classify responders/nonresponders (14).

Clinical outcome measures were compared between BICLA responders and nonresponders at week 52, regardless of treatment group assignment, and results are presented in a hierarchy of clinical relevance, agreed upon by consensus between the authors. Outcome measures include the percentage of patients with flares (defined as ≥1 new BILAG 2004 A domain score or ≥2 new BILAG 2004 B domain scores compared with the prior visit) through week 52, as well as an analysis by flare severity (percentage of patients with ≥1 new BILAG 2004 A domain score or ≥2 new BILAG 2004 B domain scores versus prior visit through week 52), annualized flare rates, percentage of patients achieving sustained oral glucocorticoid taper (defined as oral glucocorticoid dosage reduction to ≤7.5 mg/day prednisone or equivalent, achieved by week 40 and sustained through week 52, in patients receiving ≥10 mg/

day at baseline), and change in daily oral glucocorticoid dosage from baseline to week 52.

Changes in PROs were assessed from baseline to week 52, including responses in Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT–F) (defined as an improvement of >3 points) (16), responses in Short Form 36 health survey (SF-36) version 2 (acute) physical component summary (PCS) and mental component summary (MCS) (defined as an improvement of >3.4 in the PCS and >4.6 in the MCS) (17), and changes from baseline in patient global assessment of disease activity (PtGA).

Attainment of Lupus Low Disease Activity State (LLDAS) (18) was assessed as previously described (19). Medical resource utilization (health care visits, emergency department [ED] use, and hospital visits) was also assessed. Other indices compared between BICLA responders and nonresponders included changes from baseline to week 52 in SLEDAI-2K score, PhGA score, joint counts (active, swollen, tender), and the Cutaneous Lupus Erythematosus Disease Area and Severity Index Activity (CLASI) (20) responses (defined as a ≥50% reduction in CLASI activity score among patients with a CLASI activity score ≥10 at baseline).

Table 1. Patient demographic data and baseline clinical characteristics*

Baseline characteristic	BICLA responders at week 52 (n = 318)	BICLA nonresponders at week 52 (n = 501)
Age, mean ± SD years	41.5 ± 11.67	41.7 ± 12.13
Sex, female	294 (92.5)	466 (93.0)
Time from SLE diagnosis to randomization, median (range) months	85.5 (0–555)	84.0 (4–503)
IFN gene signature test-high at screening	261 (82.1)	415 (82.8)
BILAG 2004 score, mean ± SD	18.9 ± 5.20	19.2 ± 5.59
≥1 A item	171 (53.8)	222 (44.3)
No A and ≥2 B items	126 (39.6)	254 (50.7)
No A and <2 B items	21 (6.6)	25 (5.0)
SLEDAI-2K score, mean ± SD	10.8 ± 3.19	11.7 ± 4.01
<10	104 (32.7)	127 (25.3)
≥10	214 (67.3)	374 (74.7)
PhGA score, mean ± SD	1.76 ± 0.425	1.81 ± 0.396
Oral GC use†	263 (82.7)	410 (81.8)
<10 mg/day	98 (30.8)	152 (30.3)
≥10 mg/day	165 (51.9)	258 (51.5)
Antimalarial use	225 (70.8)	361 (72.1)
Immunosuppressant use	158 (49.7)	230 (45.9)
CLASI activity score, mean ± SD	8.5 ± 7.56	7.8 ± 7.18
<10	215 (67.6)	373 (74.5)
≥10	103 (32.4)	128 (25.5)
0	13 (4.1)	20 (4.0)
>0	305 (95.9)	481 (96.0)
SDI global score, mean ± SD	0.6 ± 1.08	0.5 ± 0.89
SJC, mean ± SD	6.5 ± 5.27	7.4 ± 6.08
TJC, mean ± SD	9.8 ± 6.94	11.1 ± 7.85
AJC, mean ± SD‡	6.1 ± 5.22	6.9 ± 5.97
Anti-dsDNA positive§	142 (44.7)	224 (44.7)
Anti-dsDNA level, median (range) IU/ml¶	48.2 (15–3,790)	57.0 (15–4,404)
Anti-dsDNA level, mean ± SD, IU/ml§	142.5 ± 401.84	220.4 ± 526.38
Abnormal C3 level#	123 (38.7)	178 (35.5)
C3 level, median (range) gm/liter§	0.729 (0.36–0.90)	0.715 (0.18–0.90)
C3 level, mean ± SD, gm/liter§	0.711 ± 0.1279	0.685 ± 0.1603
Abnormal C4 level**	72 (22.6)	118 (23.6)
C4 level, median (range) gm/liter§	0.080 (0.05–0.10)	0.072 (0.05–0.10)
C4 level, mean ± SD, gm/liter§	0.075 ± 0.0168	0.071 ± 0.0145

* Except where indicated otherwise, values are the number (%) of patients. BICLA = British Isles Lupus Assessment Group (BILAG)–based Composite Lupus Assessment; SLE = systemic lupus erythematosus; IFN = interferon; BILAG 2004 = 2004 Update of the BILAG; SLEDAI-2K = SLE Disease Activity Index 2000; PhGA = physician global assessment of disease activity; CLASI = Cutaneous Lupus Erythematosus Disease Area and Severity Index; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SJC = swollen joint count; TJC = tender joint count; AJC = active joint count.
† Oral glucocorticoids (GCs) include prednisone or equivalent.
‡ Defined as a joint with swelling and tenderness.
§ Positivity defined as >15 IU/ml.
¶ Only patients with anti–double-stranded DNA (anti-dsDNA) antibodies and abnormal complement levels at baseline are included in the summary statistics for the respective variables.
# Abnormal levels defined as <0.9 gm/liter.
** Abnormal levels defined as <0.1 gm/liter.

Serologies (anti–double-stranded DNA [anti-dsDNA] antibodies and complement C3 and C4) were evaluated; anti-dsDNA antibody levels were classified as "positive" (>15 IU/ml) or "negative" (≤15 IU/ml), and complement levels were classified as "abnormal" (C3, <0.9 gm/liter; C4, <0.1 gm/liter) or "normal" (C3, ≥0.9 gm/liter; C4, ≥0.1 gm/liter). Adverse events (AEs) were also assessed.

Statistical analysis. The similar designs of the TULIP-1 and TULIP-2 studies allowed for the results to be pooled. Sample sizes were selected for TULIP-1 and TULIP-2 based on powering of the primary and key secondary end points and to ensure an adequate safety database. In TULIP-1 and TULIP-2, 180 subjects per arm yielded >99% and 88% power, respectively, to reject the hypothesis (no difference in the primary end point) using a 2-sided alpha value of 0.05. Responder versus nonresponder rates were calculated using a stratified Cochran-Mantel-Haenszel approach (21), which included stratification factors of SLEDAI-2K score at screening (<10 or ≥10), baseline oral glucocorticoid dosage (<10 mg/day or ≥10 mg/day), type I IFN gene signature test status at screening (test-low or test-high); the study in which the patients participated (TULIP-1 or TULIP-2) was also included in the model. For all responder analyses, patients were considered nonresponders if they received restricted medications beyond the protocol-allowed thresholds or discontinued the investigational drug before the week 52 assessment. Comparison of estimated change from baseline to week 52 between BICLA responders and nonresponders was assessed using a mixed repeated-measures model with fixed effects for baseline value, group, visit, study, and the stratification factors used at screening; a group-by-visit interaction term was used, and visit was a repeated variable in the model. Missing data were imputed using the last observation carried forward for the first visit with missing data; subsequent visits with missing data were not imputed. For responder analyses, if any component of the variable could not be derived owing to missing data, the patient was classified as a nonresponder for that visit.

Study approval. The TULIP-1 and TULIP-2 trials were conducted in accordance with the principles of the Declaration of Helsinki and followed the International Conference on Harmonisation Guidelines for Good Clinical Practice, and all patients provided written informed consent in accordance with local requirements (13,14). As this was a post hoc analysis of anonymized data, no ethics committee or institutional review board approvals were required—all such approvals were obtained in the original trials (13,14).

RESULTS

Trial populations. Data were pooled for 457 patients in TULIP-1 and 362 patients in TULIP-2 (n = 819). Across both trials, 360 patients received anifrolumab 300 mg, 93 patients received anifrolumab 150 mg (in TULIP-1 only), and 366 patients

received placebo. Regardless of treatment group assignment, at week 52 there were 318 BICLA responders (anifrolumab 300 mg, n = 171; anifrolumab 150 mg, n = 35; placebo, n = 112) and 501 BICLA nonresponders (anifrolumab 300 mg, n = 189; anifrolumab 150 mg, n = 58; placebo, n = 254).

Patient demographics and baseline clinical characteristics were generally balanced across BICLA responders and nonresponders (Table 1 and Supplementary Table 1, available on the *Arthritis & Rheumatology* website at http://onlinelibrary.wiley.com/doi/10.1002/art.41778/abstract). The majority of patients were female (92.5% of responders, 93.0% of nonresponders), and the mean ± SD age was 41.5 ± 11.67 years for responders and 41.7 ± 12.13 years for nonresponders. Similar proportions of BICLA responders and nonresponders were white (67.0% versus 65.9%), Black/African American (14.2% versus 12.6%), or Asian (9.1% versus 11.0%).

Flares. More BICLA responders than nonresponders were flare-free over the 52-week treatment period (76.1% versus 52.2%), meaning that fewer BICLA responders than nonresponders experienced ≥1 flare (23.9% versus 47.8%; difference −23.9% [95% confidence interval (95% CI) −30.4, −17.5]; nominal $P < 0.001$) (Figure 1A and Supplementary Table 2, http://onlinelibrary.wiley.com/doi/10.1002/art.41778/abstract). Compared with their prior visits, fewer BICLA responders than nonresponders experienced ≥1 new BILAG 2004 A domain score or ≥2 new BILAG 2004 B domain scores. Fewer patients experienced 1, 2, or ≥3 flares, and the annualized flare rate was lower for BICLA responders than nonresponders (rate ratio [RR] 0.36 [95% CI 0.29, 0.47]; nominal $P < 0.001$) (Supplementary Table 2).

Oral glucocorticoid use and steroid sparing. At baseline, similar percentages of BICLA responders and nonresponders were receiving oral glucocorticoids at any dosage and at ≥10 mg/day (Table 1). Compared with BICLA nonresponders, BICLA responders had greater reductions in daily oral glucocorticoid dosage from baseline to week 52 (least squares mean [LSM] difference −4.29 mg/day [95% CI −5.37, −3.20]; nominal $P < 0.001$) (Figure 1B). The key secondary end point of sustained oral glucocorticoid dosage reduction to ≤7.5 mg/day among patients who were receiving oral glucocorticoids ≥10 mg/day at baseline was achieved by more BICLA responders than nonresponders (79.2% versus 19.1%; difference 60.1% [95% CI 52.1%, 68.1%]; nominal $P < 0.001$) (Figure 1C). The mean ± SD cumulative oral glucocorticoid dose through week 52 was 31.3% lower in BICLA responders compared with nonresponders (mean ± SD 2,159.20 ± 1,661.39 mg versus 3,140.81 ± 3,081.19 mg) (Figure 1D).

PROs. FACIT–F, SF-36 MCS, and SF-36 PCS scores were similar among BICLA responders and nonresponders at baseline (Supplementary Table 3, http://onlinelibrary.wiley.com/doi/10.1002/art.41778/abstract). Improvement in FACIT–F scores
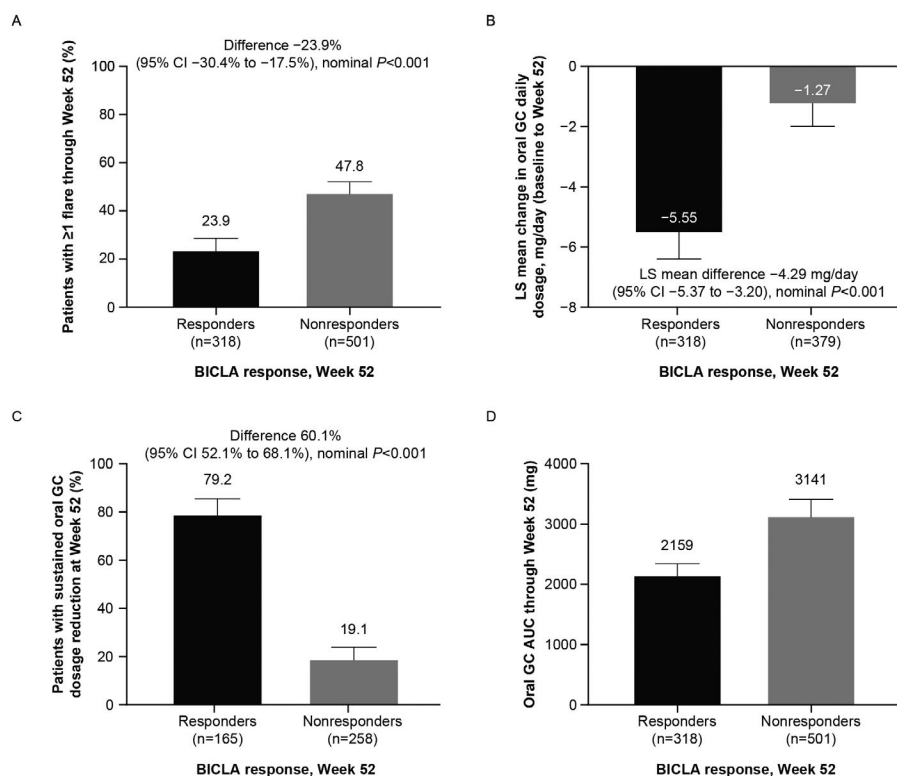
Figure 1. Flares and oral glucocorticoid (GC) use in British Isles Lupus Assessment Group (BILAG)–based Composite Lupus Assessment (BICLA) responders compared with nonresponders. **A**, Patients with ≥1 BILAG 2004 flare through week 52. Bars show the mean and 95% confidence interval (95% CI). **B**, Least squares mean (LSM) change in oral glucocorticoid daily dosage from baseline to week 52 in all patients regardless of baseline oral glucocorticoid dosage. Bars show the LSM change and 95% CI. **C**, Patients achieving sustained oral glucocorticoid dosage reduction to ≤7.5 mg/day among patients receiving oral glucocorticoids ≥10 mg/day at baseline. Sustained oral glucocorticoid dosage reduction is defined as oral glucocorticoid dosage of ≤7.5 mg/day sustained from weeks 40 to 52. Bars show the mean and 95% CI. **D**, Oral glucocorticoids area under the curve (AUC) through week 52 for all patients regardless of baseline oral glucocorticoid dosage. Bars show the mean ± SD. Rate difference, 95% CIs, and nominal *P* values were calculated using a stratified Cochran-Mantel-Haenszel approach.

was reported in more BICLA responders than nonresponders (55.6% versus 15.7%; difference 40.0% [95% CI 33.6%, 46.3%]; nominal *P* < 0.001) (Figure 2A). Similarly, more BICLA responders than nonresponders had improvement above the predefined threshold in SF-36 PCS scores (57.9% versus 12.8%; difference 45.1% [95% CI 38.9%, 51.3%]; nominal *P* < 0.001) and SF-36 MCS scores (42.6% versus 12.3%; difference 30.3% [95% CI 24.1%, 36.5%]; nominal *P* < 0.001) (Figures 2B–C). PtGA scores were similar for BICLA responders and nonresponders at baseline (Supplementary Table 2). Greater improvements in PtGA scores from baseline to week 52 were reported for BICLA responders compared with nonresponders (LSM difference −11.1 [95% CI −14.9, −7.3]; nominal *P* < 0.001) (Figure 2D).

Medical resource utilization. During the 52-week trials, fewer BICLA responders than nonresponders had health care visits (62.5% versus 70.7%; difference −8.3 [95% CI −14.9%, −1.6%]; nominal *P* = 0.015) (Table 2). Fewer BICLA responders required ED visits compared with nonresponders (11.9% versus

21.8%; difference −9.9% [95% CI −15.2%, −4.5%]; nominal *P* < 0.001), and fewer ED visits were related to increased SLE activity (2.6% versus 24.0%; difference −21.4% [95% CI −35.3%, −7.5%]; nominal *P* = 0.003). Similarly, fewer BICLA responders than nonresponders had hospital visits (4.5% versus 14.4%; difference −10.0% [95% CI −14.3%, −5.7%]; nominal *P* < 0.001), and no hospital visits were related to increased SLE activity among BICLA responders, compared with 38.5% among BICLA nonresponders (difference −38.5% [95% CI −58.8%, −18.2%]; nominal *P* < 0.001).

SLEDAI-2K, PhGA, and LLDAS. The mean ± SD of SLEDAI-2K and PhGA scores were similar among responders and nonresponders at baseline (Table 1). From baseline to week 52, greater improvements were observed for BICLA responders compared with nonresponders in total SLEDAI-2K scores (LSM difference −3.5 [95% CI −4.1, −3.0]; nominal *P* < 0.001) (Figure 3A) and PhGA scores (LSM difference −0.59 [95% CI −0.67, −0.51]; nominal *P* < 0.001) (Figure 3B).
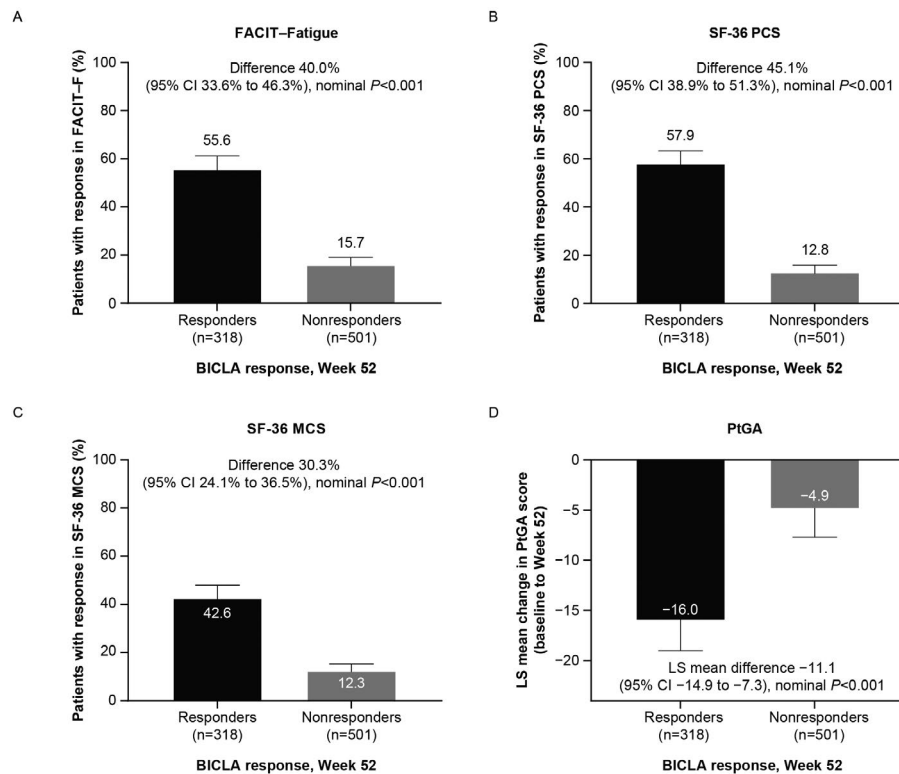
Figure 2. Patient-reported outcomes at week 52 in BICLA responders compared with nonresponders. **A–C**, Patients with a response according to following assessments: the Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT–F), defined as an improvement of >3 points from baseline to week 52 (**A**), Short Form 36 health survey (SF-36) physical component summary (PCS), defined as an increase of >3.4 in the PCS domain from baseline to week 52 (**B**), and SF-36 mental component summary (MCS), defined as an increase of >4.6 in the MCS domain from baseline to week 52 (**C**). Bars show the mean and 95% CI. Response rates, 95% CIs, and nominal *P* values were calculated using a stratified Cochran-Mantel-Haenszel approach. **D**, LSM change in patient global assessment (PtGA) score from baseline to week 52. Bars show the LSM change and 95% CI. LSM difference, 95% CIs, and nominal *P* values were calculated using mixed-model repeated measures. See Figure 1 for other definitions.

At week 52, 32.0% of BICLA responders versus 2.3% of BICLA nonresponders had attained LLDAS (difference 29.7% [95% CI 24.3%, 35.1%]; nominal *P* < 0.001) (Supplementary Table 4, http://onlinelibrary.wiley.com/doi/10.1002/art.41778/abstract). Accordingly, 89.5% of patients who attained LLDAS were BICLA responders.

Table 2. Medical resource utilization for BICLA responders and nonresponders*

Medical resource utilization	BICLA responders at week 52 (n = 318)	BICLA nonresponders at week 52 (n = 501)	Difference between groups (95% CI)	Nominal *P*
Health care visits (specialist and primary care)				
Patients with ≥1 health care visit†	198 (62.5)	348 (70.7)	−8.3 (−14.9, −1.6)	0.015
ED visits				
Patients with ≥1 ED visit†	38 (11.9)	107 (21.8)	−9.9 (−15.2, −4.5)	<0.001
Visit related to increase in SLE activity†	1 (2.6)	25 (24.0)	−21.4 (−35.3, −7.5)	0.003
No. of ED visits per patient, mean ± SD‡	1.4 ± 0.86	1.7 ± 1.48	–	–
Hospitalizations				
Patients with ≥1 hospital visit†	14 (4.5)	71 (14.4)	−10.0 (−14.3, −5.7)	<0.001
Visits related to increase in SLE activity†	0	25 (38.5)	−38.5 (−58.8, −18.2)	<0.001
No. of hospital visits per patient, mean ± SD‡	1.6 ± 2.13	1.4 ± 0.72	–	–
Length of hospital stay, mean ± SD days‡	5.7 (2.64)	7.4 (8.02)	–	–
Total days in ICU‡	0	5	–	–
No. of days in ICU, mean ± SD	–	1.8 (0.45)	–	–

* Except where indicated otherwise, values are the number (%) of patients. ICU = intensive care unit (see Table 1 for other definitions).
† Percentages, differences, 95% confidence intervals (95% CIs), and nominal *P* values were calculated using a stratified Cochran-Mantel-Haenszel approach.
‡ Data on hospital visits and emergency department (ED) visits were missing for 8 patients in the BICLA nonresponders group.
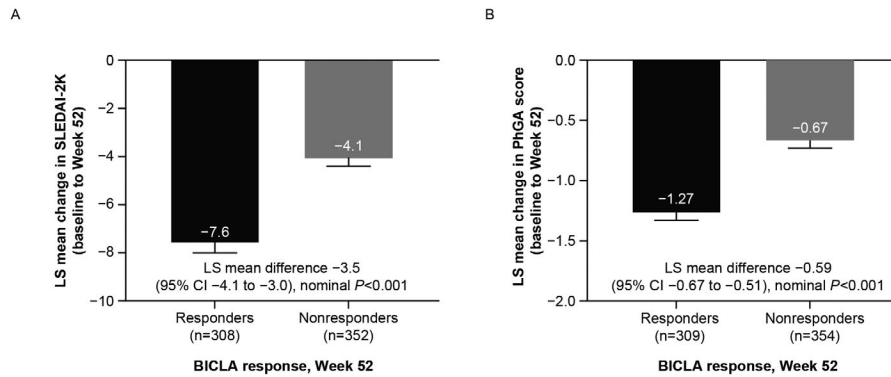
Figure 3. Change in Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) (**A**) and physician global assessment (PhGA) scores (**B**) from baseline to week 52 in BICLA responders compared with nonresponders. Bars show the mean and 95% CI. LSM difference, 95% CIs, and nominal *P* values were calculated using mixed-model repeated measures. See Figure 1 for other definitions.

Skin disease. Overall, 32.4% of BICLA responders and 25.5% of nonresponders had a baseline CLASI activity score ≥10 (Table 1). Among these patients, more BICLA responders achieved a ≥50% reduction in CLASI activity score at week 52 compared with nonresponders (92.0% versus 23.2%; difference 68.8% [95% CI 59.2%, 78.3%]; nominal *P* < 0.001) (Figure 4A).

Arthritis. The mean ± SD active joint counts (defined as joints with both swelling and tenderness) were 6.1 ± 5.22 (BICLA responders) and 6.9 ± 5.97 (nonresponders), the mean ± SD swollen joint counts were 6.5 ± 5.27 (BICLA responders) and 7.4 ± 6.08 (nonresponders), and the mean ± SD tender joint counts were 9.8 ± 6.94 (BICLA responders) and 11.1 ± 7.85 (nonresponders) (Table 1). From baseline to week 52, joint counts improved more for
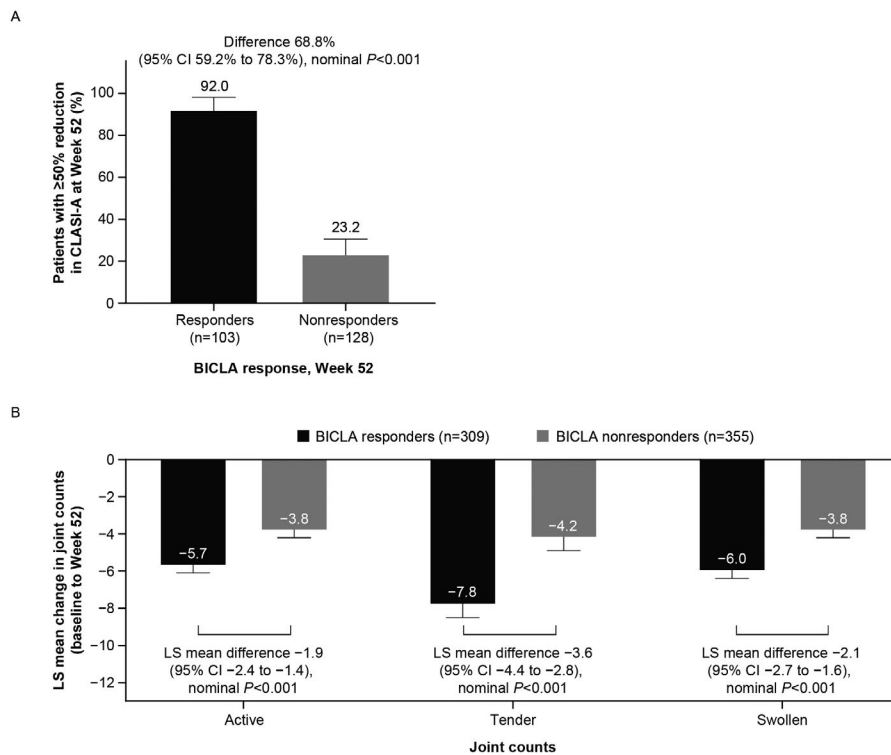
Figure 4. Cutaneous Lupus Erythematosus Disease Area and Severity Index (CLASI) response and joint counts in BICLA responders compared with nonresponders. **A**, Patients with a CLASI response at week 52 (defined as ≥50% reduction from baseline to week 52) among patients with a CLASI activity score of ≥10 at baseline. Bars show the mean and 95% CI. Response rates, 95% CIs, and nominal *P* values were calculated using a stratified Cochran-Mantel-Haenszel approach. **B**, Change in LSM joint counts from baseline to week 52 for active joints (defined as a joint with swelling and tenderness), tender joints, and swollen joints. Bars show the mean and 95% CI. LSM difference, 95% CIs, and nominal *P* values were calculated using mixed-model repeated measures. See Figure 1 for other definitions.

BICLA responders compared with nonresponders, for the active joint count (LSM difference −1.9 [95% CI −2.4, −1.4]; nominal $P < 0.001$), tender joint count (LS mean difference −3.6 [95% CI −4.4, −2.8]; nominal $P < 0.001$), and swollen joint count (LS mean difference −2.1 [95% CI −2.7, −1.6]; nominal $P < 0.001$) (Figure 4B).

Serology. Equal percentages of BICLA responders and nonresponders were positive for anti-dsDNA antibodies at baseline (Table 1). Improvement from positive to negative anti-dsDNA antibody status was observed in similar percentages of BICLA responders and nonresponders (5.0% versus 4.4%) (Supplementary Table 5, http://onlinelibrary.wiley.com/doi/10.1002/art.41778/abstract).

Similar proportions of BICLA responders and nonresponders had abnormal C3 and C4 levels at baseline (Table 1). Percentage changes from baseline to week 52 in complement levels did not differ between BICLA responders and nonresponders for C3 (LSM difference 2.82 [95% CI −4.185, 9.819]; nominal $P = 0.429$) or C4 (LSM difference −9.63 [95% CI −25.174, 5.910]; nominal $P = 0.223$) (Supplementary Table 5, http://onlinelibrary.wiley.com/doi/10.1002/art.41778/abstract). More BICLA responders than nonresponders showed improvement from abnormal to normal C3 levels (10.4% versus 7.0%) and C4 levels (7.5% versus 4.8%).

Safety. AE frequencies were similar between BICLA responders and nonresponders (83.6% versus 85.2%) (Supplementary Table 6, http://onlinelibrary.wiley.com/doi/10.1002/art.41778/abstract). Mild and moderate AEs were reported in similar percentages of BICLA responders and nonresponders, whereas fewer BICLA responders than nonresponders experienced severe AEs (3.8% versus 9.4%). There were no AEs leading to discontinuation in BICLA responders compared with 8.2% in nonresponders. Fewer BICLA responders than nonresponders experienced serious AEs (SAEs) (5.0% versus 19.0%), including SAEs due to lupus nephritis or SLE (0.3% versus 3.0%). Fewer BICLA responders than nonresponders had non-opportunistic serious infections (2.2% versus 6.8%). The percentage of patients with herpes zoster was similar in BICLA responders and nonresponders (4.7% versus 3.6%), as was the percentage of patients with influenza (1.9% versus 2.0%) or malignancy (0.6% versus 1.0%).

DISCUSSION

BICLA response is a dichotomous SLE trial end point that classifies a patient as a responder or nonresponder based on changes in organ domain activity (5,6). As BICLA response is primarily used in the clinical trial setting, the aim of this study was to assess the meaningfulness of BICLA response in terms of outcomes that are relevant to patients and physicians. In this post hoc analysis of pooled data acquired from 819 patients enrolled in the TULIP-1 and TULIP-2 trials, BICLA response was significantly associated with improved clinical outcomes across a range of SLE assessments, key PROs, and medical resource utilization measures.

Flares, with or without an increase in glucocorticoid dosage, pose significant risks to patients with SLE. In the long term, both disease flares and oral glucocorticoid use have been linked to organ damage, which itself increases mortality (22–26). Flares are also associated with reduced health-related quality of life, and flare severity together with oral glucocorticoid use correlate with health care costs (27–31). A key SLE treatment goal therefore is to prevent flares while minimizing oral glucocorticoid exposure, which in turn is expected to reduce medical resource utilization (27–30,32,33). BICLA responders had fewer disease flares, and there were more BICLA responders who were flare-free, consistent with the BICLA response definition, which requires improvement in all BILAG 2004 organs affected at baseline and no new activity in the remaining organ systems. BICLA responders also had a lower daily oral glucocorticoid dosage, indicating that the observed responses were not due to greater oral glucocorticoid use. A greater percentage of BICLA responders achieved sustained oral glucocorticoid reduction to target dosage, and they also had fewer hospitalizations and ED visits than did nonresponders, including those related to increased SLE activity.

In addition, greater improvements in global and organ-specific disease activity were observed in BICLA responders compared with nonresponders, as measured by PhGA, SLEDAI-2K, LLDAS attainment, skin disease as measured by CLASI activity score, and joint counts. Since improved disease activity, reduced oral glucocorticoid exposure, and LLDAS attainment have been shown to be associated with reduced health care costs (29,30), BICLA responders may incur lower health care costs than nonresponders.

We also assessed AEs in BICLA responders and nonresponders occurring during the TULIP trials. There were fewer SAEs in BICLA responders compared with nonresponders, potentially because of greater disease control and/or reduced oral glucocorticoid dosage in the BICLA responder group. There were also fewer SAEs related to lupus nephritis or SLE disease worsening in BICLA responders, consistent with the efficacy findings of lower flare rates and fewer SLE-related ED visits and hospitalizations associated with BICLA response. BICLA nonresponders had a greater propensity to discontinue treatment due to an AE than BICLA responders, consistent with the BICLA response definition, where patients who discontinued investigational product were classified as nonresponders.

PROs have been incorporated into nearly all SLE clinical trials. However, analyses have often yielded discordance between clinical outcomes and PROs, as patient perceptions of disease activity and illness are heavily impacted by fatigue and quality of life, which are not captured by measures of disease activity (34–38). Furthermore, fatigue and other PROs do not always associate with disease activity measures in clinical practice (39–41). In the TULIP trials, BICLA responders had improvements in

validated PROs, including the physical and mental components of the SF-36 health survey and the FACIT assessment of fatigue. Fatigue, a common symptom in patients with SLE, interferes with daily life (42), and more than half of the patients in the TULIP trials with BICLA responses experienced improvement in fatigue. PtGA and PhGA scores both showed greater degrees of improvement among BICLA responders compared with nonresponders. Our results suggest that BICLA response translates to general improvements in the physical and mental well-being of patients with SLE.

Investigation of the correlations of SRI-4 response to clinical outcomes in pooled data from 2 phase IIb trials (sifalimumab and anifrolumab), as well as 2 phase III trials of belimumab, demonstrated improved general clinical outcomes in SRI-4 responders compared with nonresponders, including improved fatigue (36,43). While changes in serologic outcomes were not significantly different between BICLA responders and nonresponders in the TULIP trials, SRI-4 response was associated with significant improvements in anti-dsDNA antibody and complement C3 levels (but not C4 levels) in the belimumab phase III trials (43). This discordance may reflect the different mechanisms of action of the 2 evaluated drugs, and/or because the BILAG 2004, on which the BICLA is anchored, does not include serology in its scoring system (4,44).

The limitations of this study include the post hoc nature of the analyses, as groups were compared based on postrandomization characteristics, which may have resulted in imbalances of baseline characteristics and stratification factors. Furthermore, these results were assessed with a single intervention and used predetermined inclusion and exclusion criteria. As such, results may not translate to other interventions or patient populations. For example, the TULIP trials excluded patients with severe active renal disease and central nervous system lupus. Therefore, future work is needed to assess the relationship between BICLA response and clinical outcomes across the full spectrum of SLE disease. In addition, because we assessed SF-36 score improvement using meaningful change thresholds established in the general population (17), future work will be needed to measure improvements in PROs in BICLA responders using SLE-specific meaningful change thresholds.

Our data confirm the value of BICLA as a clinical trial end point and that a BICLA response is associated with improvements in a range of important outcomes that resonate with the priorities of both clinicians and patients in everyday practice.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Tummala had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study conception and design. Furie, Morand, Bruce, Isenberg, Abreu, Pineda, Tummala.
Acquisition of data. Furie, Bruce, Pineda, Tummala.
Analysis and interpretation of data. Furie, Morand, Bruce, Isenberg, van Vollenhoven, Abreu, Tummala.

ROLE OF THE STUDY SPONSOR

REFERENCES

1. US Food and Drug Administration. Systemic lupus erythematosus–developing medical products for treatment. June 2010. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/systemic-lupus-erythematosus-developing-medical-products-treatment.

2. Furie RA, Petri MA, Wallace DJ, Ginzler EM, Merrill JT, Stohl W, et al. Novel evidence-based systemic lupus erythematosus responder index. Arthritis Rheum 2009;61:1143–51.

3. Gladman DD, Ibanez D, Urowitz MB. Systemic Lupus Erythematosus Disease Activity Index 2000. J Rheumatol 2002;29:288–91.

4. Isenberg DA, Rahman A, Allen E, Farewell V, Akil M, Bruce IN, et al. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology (Oxford) 2005;44:902–6.

5. Wallace DJ, Kalunian K, Petri MA, Strand V, Houssiau FA, Pike M, et al. Efficacy and safety of epratuzumab in patients with moderate/severe active systemic lupus erythematosus: results from EMBLEM, a phase IIb, randomised, double-blind, placebo-controlled, multicentre study. Ann Rheum Dis 2014;73:183–90.

6. Wallace D, Strand V, Furie R, Petri M, Kalunian K, Pike M, et al. Evaluation of treatment success in systemic lupus erythematosus clinical trials: development of the British Isles Lupus Assessment Group-based Composite Lupus Assessment Endpoint [poster]. Presented at the Annual Meeting of the American College of Rheumatology; 2011 November 5–9; Chicago, Illinois.

7. 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.

8. Yee CS, Farewell V, Isenberg DA, Griffiths B, Teh LS, Bruce IN, et al. The BILAG-2004 index is sensitive to change for assessment of SLE disease activity. Rheumatology (Oxford) 2009;48:691–5.

9. Wallace D, Popa S, Spindler A, Eimon A, González-Rivera T, Utset T, et al. Improvement of disease activity and reduction of severe flares following subcutaneous administration of an IL-6 monoclonal antibody (mAb) in subjects with active generalized systemic lupus erythematosus (SLE) [abstract]. Arthritis Rheumatol 2014;66:3531.

10. ClinicalTrials.gov. National Library of Medicine: NIH clinical trials database. URL: https://clinicaltrials.gov/.

11. Clowse ME, Wallace DJ, Furie RA, Petri MA, Pike MC, Leszczyński P, et al. Efficacy and safety of epratuzumab in moderately to severely active systemic lupus erythematosus: results from two phase III randomized, double-blind, placebo-controlled trials. Arthritis Rheumatol 2017;69:362–75.

12. Furie R, Khamashta M, Merrill JT, Werth VP, Kalunian K, Brohawn P, et al. Anifrolumab, an anti–interferon-α receptor monoclonal antibody, in moderate-to-severe systemic lupus erythematosus. Arthritis Rheumatol 2017;69:376–86.

13. Furie R, Morand E, Bruce I, Manzi S, Kalunian K, Vital E, et al. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. Lancet Rheumatol 2019;1:E208–19.

14. Morand EF, Furie R, Tanaka Y, Bruce IN, Askanase AD, Richez C, et al. Trial of anifrolumab in active systemic lupus erythematosus. N Engl J Med 2020;382:211–21.

15. Hochberg MC for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725.

16. 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;38:672–9.

17. Maruish ME. User's manual for the SF-36v2 Health Survey. 3rd ed. Lincoln, (Rhode Island): QualityMetric Incorporated; 2011.

18. Franklyn K, Lau CS, Navarra SV, Louthrenoo W, Lateef A, Hamijoyo L, et al. Definition and initial validation of a Lupus Low Disease Activity State (LLDAS). Ann Rheum Dis 2016;75:1615–21.

19. Morand EF, Trasieva T, Berglind A, Illei GG, Tummala R. Lupus Low Disease Activity State (LLDAS) attainment discriminates responders in a systemic lupus erythematosus trial: post-hoc analysis of the Phase IIb MUSE trial of anifrolumab. Ann Rheum Dis 2018;77:706–13.

20. Albrecht J, Taylor L, Berlin JA, Dulay S, Ang G, Fakharzadeh S, et al. The CLASI (Cutaneous Lupus Erythematosus Disease Area and Severity Index): an outcome instrument for cutaneous lupus erythematosus. J Invest Dermatol 2005;125:889–94.

21. Stokes ME, Davis CS, Koch GG. Categorical data analysis using SAS. 3rd ed. Carey (North Carolina): SAS Institute; 2012.

22. Bruce IN, O'Keeffe AG, Farewell V, Hanly JG, Manzi S, Su L, et al. Factors associated with damage accrual in patients with systemic lupus erythematosus: results from the Systemic Lupus International Collaborating Clinics (SLICC) inception cohort. Ann Rheum Dis 2015;74:1706–13.

23. Apostolopoulos D, Kandane-Rathnayake R, Louthrenoo W, Luo SF, Wu Y, Lateef A, et al. Factors associated with damage accrual in patients with systemic lupus erythematosus with no clinical or serological disease activity: a multicentre cohort study. Lancet Rheumatol 2020;2:e24–30.

24. 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:S427–32.

25. Tsang AS, Bultink IE, Heslinga M, Voskuyl AE. Both prolonged remission and Lupus Low Disease Activity State are associated with reduced damage accrual in systemic lupus erythematosus. Rheumatology (Oxford) 2017;56:121–8.

26. Murimi-Worstell IB, Lin DH, Nab H, Kan HJ, Onasanya O, Tierce JC, et al. Association between organ damage and mortality in systemic lupus erythematosus: a systematic review and meta-analysis. BMJ Open 2020;10:e031850.

27. Katz P, Wan GJ, Daly P, Topf L, Connolly-Strong E, Bostic R, et al. Patient-reported flare frequency is associated with diminished quality of life and family role functioning in systemic lupus erythematosus. Qual Life Res 2020;29:3251–61.

28. Katz P, Nelson WW, Daly RP, Topf L, Connolly-Strong E, Reed ML. Patient-reported lupus flare symptoms are associated with worsened patient outcomes and increased economic burden. J Manag Care Spec Pharm 2020;26:275–83.

29. Yeo AL, Koelmeyer R, Kandane-Rathnayake R, Golder V, Hoi A, Huq M, et al. Lupus low disease activity state and reduced direct health care costs in patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2020;72:1289–95.

30. Kan HJ, Song X, Johnson BH, Bechtel B, O'Sullivan D, Molta CT. Healthcare utilization and costs of systemic lupus erythematosus in Medicaid. Biomed Res Int 2013;2013:808391.

31. Zhu TY, Tam LS, Lee VW, Lee KK, Li EK. The impact of flare on disease costs of patients with systemic lupus erythematosus. Arthritis Rheum 2009;61:1159–67.

32. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Ann Rheum Dis 2019;78:736–45.

33. Van Vollenhoven RF, Mosca M, Bertsias G, Isenberg D, Kuhn A, Lerstrøm K, et al. Treat-to-target in systemic lupus erythematosus: recommendations from an international task force. Ann Rheum Dis 2014;73:958–67.

34. Alarcón GS, McGwin G Jr, Brooks K, Roseman JM, Fessler BJ, Sanchez ML, et al. Systemic lupus erythematosus in three ethnic groups. XI. Sources of discrepancy in perception of disease activity: a comparison of physician and patient visual analog scale scores. Arthritis Rheum 2002;47:408–13.

35. Elera-Fitzcarrald C, Vega K, Gamboa-Cárdenas RV, Zúñiga K, Medina M, Pimentel-Quiroz V, et al. Discrepant perception of lupus disease activity: a comparison between patients' and physicians' disease activity scores. J Clin Rheumatol 2020;26:S165–9.

36. 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;27:955–62.

37. Mahieu M, Yount S, Ramsey-Goldman R. Patient-reported outcomes in systemic lupus erythematosus. Rheum Dis Clin North Am 2016;42:253–63.

38. 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;71:829–38.

39. Wang B, Gladman DD, Urowitz MB. Fatigue in lupus is not correlated with disease activity. J Rheumatol 1998;25:892–5.

40. Bruce IN, Mak VC, Hallett DC, Gladman DD, Urowitz MB. Factors associated with fatigue in patients with systemic lupus erythematosus. Ann Rheum Dis 1999;58:379–81.

41. Nantes SG, Strand V, Su J, Touma Z. Comparison of the sensitivity to change of the 36-item Short Form Health Survey and the lupus quality of life measure using various definitions of minimum clinically important differences in patients with active systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2018;70:125–33.

42. Morgan C, Bland AR, Maker C, Dunnage J, Bruce IN. Individuals living with lupus: findings from the LUPUS UK Members Survey 2014. Lupus 2018;27:681–7.

43. Van Vollenhoven RF, Stohl W, Furie RA, Fox NL, Groark JG, Bass D, et al. Clinical response beyond the Systemic Lupus Erythematosus Responder Index: post-hoc analysis of the BLISS-SC study. Lupus Sci Med 2018;5:e000288.

44. Ceccarelli F, Perricone C, Massaro L, Cipriano E, Alessandri C, Spinelli FR, et al. Assessment of disease activity in systemic lupus erythematosus: lights and shadows [review]. Autoimmun Rev 2015;14:601–8.