Impact of belimumab on patient-reported outcomes in systemic lupus erythematosus: review of clinical studies

Elvira Bangert¹
Laura Wakani²
Mehveen Merchant³
Vibeke Strand⁴
Zahi Touma²

¹Division of Rheumatology, Department of Medicine, Queen’s University, Kingston, ON, Canada; ²University of Toronto Lupus Clinic, Toronto Western Hospital, Centre for Prognosis Studies in the Rheumatic Diseases, Toronto, ON, Canada; ³Division of Rheumatology, Nova Scotia Health Authority, Dalhousie University, Halifax, NS, Canada; ⁴Division of Immunology/Rheumatology, Stanford University, Palo Alto, CA, USA

Abstract: Systemic lupus erythematosus (SLE) is a chronic autoimmune, multisystem rheumatic disease with significant impact on health-related quality of life (HRQoL). Patient-reported outcomes (PROs) provide valuable data on patient perceptions across a variety of domains, such as HRQoL, pain, fatigue, and depression. The measurement and results of PROs with respect to HRQoL in randomized controlled trials (RCTs) on belimumab (B-lymphocyte stimulator inhibitor) in SLE are reviewed here, including BLISS-52 and BLISS-76, as well as publications related to belimumab trials that included HRQoL data. Other trials that evaluated belimumab did not include HRQoL data and were therefore not included in the analysis. The BLISS-52 and BLISS-76 RCTs met their primary endpoints and demonstrated improvements in PROs, measured by the 36-item Short Form Health Survey, EuroQol 5 Dimensions, and Functional Assessment of Chronic Illness Therapy-Fatigue Scale. Belimumab was shown overall to improve PROs in adult autoimmune-positive lupus patients.

Keywords: quality of life, fatigue, systemic lupus, Benlysta, trials

Introduction

Systemic lupus erythematosus (SLE) is a chronic autoimmune, multisystem rheumatic disease with significant impact on many dimensions of patients’ well-being, causing marked impairment in health-related quality of life (HRQoL).¹ Evidence suggests that the impact on HRQoL experienced by SLE patients is not adequately captured by assessments of physician-assessed disease activity and damage alone.² Patient-reported outcomes (PROs) reflect the patient’s perspectives of their illness, HRQoL, and well-being, and can be helpful in informing treatment decisions.³ Recognizing their value, PROs are identified as an independent core domain by the Outcome Measures in Rheumatology (OMERACT) and other regulatory agencies and health care decision makers. OMERACT recommended four domains to be assessed in SLE randomized controlled trials (RCTs) and longitudinal observational studies: disease activity, HRQoL, damage, and adverse events.⁴ The US Food and Drug Administration (FDA) requires that control group participants in SLE RCTs receive, at minimum, standard of care therapy.⁵ Detecting a difference in endpoints between the treatment and control groups may be challenging if a similar level of response to the experimental, add-on treatment is observed. To address this challenge, responder vs non-responder analysis is used to detect differences in outcomes. The metrics and results of PROs in the belimumab RCTs in SLE are reviewed here.
Belimumab in SLE

Similar to the clinical heterogeneity of SLE, its etiopathogenesis is also multifactorial, involving environmental (e.g., sex hormones, ultraviolet light, viral infections), genetic and epigenetic, hormonal, and immunologic factors. In genetically predisposed individuals, a normal immune response is impaired, leading to increased autoantibodies and activation of the innate and adaptive immune responses.

Patients with SLE have both abnormal activation pathways and activation of a large proportion of B and T cells causing autoimmunity and inflammation. B cells play a crucial role in the pathogenesis of SLE, including acting as precursors for plasma cells that secrete autoantibodies, presenting antigenic peptides to T cells, and secreting pro-inflammatory cytokines that amplify and downregulate immune responses.

Not surprisingly, given the central role for B cells in the pathogenesis of SLE, B cells emerged as a potential therapeutic target. Recent therapeutic interventions focus on strategies such as inhibiting B-cell activation, decreasing B-cell survival, or deleting specific lineages of B cells. B-lymphocyte stimulator (BLyS), also known as B-cell activating factor (BAFF), and a proliferation-inducing ligand, are cytokines belonging to the tumor necrosis factor ligand superfamily; they activate B-cell stimulation and maturation of B cells through the binding of several receptors on B cells. Studies have demonstrated that BLyS levels are correlated with anti-dsDNA antibody levels and SLE disease activity, therefore representing a potentially important therapeutic target in SLE. Belimumab (Benlysta) is a human IgG1k monoclonal antibody that binds BLyS/BAFF, thus preventing it from binding to its receptors on B cells. Consequently, belimumab depletes activated and naive B cells but not directly memory B cells. In 2011, intravenous (IV) belimumab was approved for the treatment of SLE, the first time for a new therapy in >50 years. In 2017, the subcutaneous belimumab formulation was approved in the USA, Japan, and the European Union.

RCTs of belimumab in SLE

The efficacy of belimumab in addition to standard of care in adult patients with autoantibody-positive SLE has been demonstrated in four phase III, multicenter, prospective RCTs. (Table 1 summarizes the trials.) Inclusion criteria for the four trials were very similar and differences are highlighted in Table 1: age ≥18 years, classification of SLE according to the American College of Rheumatology criteria, seropositivity for autoantibodies, and a stable treatment regimen with glucocorticosteroids, antimalarials, or immunosuppressives for at least 30 days before the first study dose. In concordance with OMERACT recommendations and regulatory agencies, the belimumab trials assessed the robustness of responses across four domains: disease activity, damage, adverse effects, and HRQoL. The primary outcome measure in all four RCTs was the Systemic Lupus Erythematosus Responder Index-4 (SRI-4) at week 52. The RCTs demonstrated significant improvement in the SRI-4 with 10 mg/kg IV belimumab compared with placebo. Improvements in multiple domains were reported vs placebo and the 10 mg/kg dose received formal approval for treatment of SLE by the US FDA, Health Canada, and European Medicines

Table 1: Summary of the trials

<table>
<thead>
<tr>
<th>RCT</th>
<th>BLISS-52 Ref.</th>
<th>BLISS-76 Ref.</th>
<th>Northeast Asia Trial Ref.</th>
<th>BLISS-SC Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration (weeks)</td>
<td>52</td>
<td>76</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>Number of patients</td>
<td>852</td>
<td>819</td>
<td>677</td>
<td>836</td>
</tr>
<tr>
<td>Location</td>
<td>Eastern Europe, Asia-Pacific, and Latin America</td>
<td>North and Central America, and Europe</td>
<td>China, Japan, and South Korea</td>
<td>North, Central, and South America, Eastern and Western Europe, Australia, and Asia</td>
</tr>
<tr>
<td>Belimumab form of administration</td>
<td>Intravenous</td>
<td>Intravenous</td>
<td>Intravenous</td>
<td>Subcutaneous</td>
</tr>
<tr>
<td>Belimumab dose groups</td>
<td>1 mg/kg or 10 mg/kg days 0, 14, and 28, and then every 28 days until 48 weeks ≥6</td>
<td>1 mg/kg or 10 mg/kg days 0, 14, and 28, and every 28 days through week 72 ≥6</td>
<td>10 mg/kg on days 0, 14, and 28, and then every 28 days up to week 48 ≥8</td>
<td>200 mg weekly</td>
</tr>
<tr>
<td>SELENA-SLEDAI scores for inclusion</td>
<td>≥6</td>
<td>≥6</td>
<td>≥8</td>
<td>≥8</td>
</tr>
</tbody>
</table>

Abbreviations: BLISS-52, Study of Belimumab in Subjects with SLE 52-week trial; BLISS-76, Study of Belimumab in Subjects with SLE 76-week trial; BLISS-SC, Study of Belimumab Administered Subcutaneously in Subjects with SLE; RCT, randomized controlled trial; SELENA-SLEDAI, Safety of Estrogens in Lupus Erythematosus National Assessment–SLE Disease Activity Index.
Agency. In the pooled data analyses of the BLISS-52 and BLISS-76, responses occurred in 50.6% of patients assigned to belimumab 10 mg/kg vs 38.6% placebo ($P<0.001$). Across the pooled data of the BLISS-52 and BLISS-76 trials, belimumab 10 mg/kg was associated with improvements in multiple domains, including: disease activity, demonstrated by reductions from baseline in Safety of Estrogens in Lupus Erythematosus National Assessment—SLE Disease Activity Index scores at week 52; reductions in the incidence and severity of flares; and decreased glucocorticosteroid doses. The BLISS-SC trial demonstrated similar treatment effects to the 10 mg/kg IV doses in BLISS-52 and BLISS-76. Importantly, patients with severe active lupus nephritis or central nervous system lupus were excluded from all four trials. During post-marketing experience, in an open-label extension over a 7-year period (1,746 patient-years), adverse events rates due to belimumab either remained stable or substantially decreased.

**PROs in the belimumab trials**

Across all trials, there was an overall consensus that belimumab was superior to placebo, reducing the burden of disease with a safe medication profile.

Besides physician-centered indices, the BLISS-52 and BLISS-76 trials (but not the Northeast Asia Trial) evaluated PROs relevant to HRQoL in SLE, including the generic Medical Outcomes Study 36-item Short Form Health Survey (SF-36 v2); secondary analyses were also performed using other generic questionnaires: the Functional Assessment of Chronic Illness Therapy–Fatigue Scale (FACIT-F) and the EuroQol 5 Dimensions (EQ-5D). Although measuring multiple aspects of HRQoL, these instruments together capture two of the most frequently reported symptoms in SLE: fatigue and pain. More specifically, all three—SF-36, FACIT-F, and EQ-5D—capture fatigue, and SF-36 and EQ-5D capture pain as well. As secondary endpoints, these two trials included mean changes in the SF-36 Physical Component Summary (PCS) and Mental Component Summary (MCS) scores (at weeks 24 and 52 for BLISS-52, and weeks 24, 52, and 76 for BLISS-76); FACIT-F, and EQ-5D scores at various time points (weeks 12, 24, 52, and 76) (BLISS-76 only). Analyses of PRO data from both the BLISS-52 and BLISS-76 trials were performed according to SRI-4 responders vs non-responders across all doses by SF-36 summary and domain scores, FACIT-F and EQ-5D (at weeks 12, 24, and 52, and, for BLISS-76 data, also at 76 weeks). In the BLISS-SC trial, FACIT-F was administered at weeks 4, 8, 12, 24, 36, and 52 and analyzed with the corresponding visits. The PROs used to measure HRQoL and fatigue, and their results from the BLISS trials, will be reviewed here.

**36-Item Short Form Health Survey**

The SF-36 was first developed in 1988 and was followed by the Standard SF-36 Health Survey in 1992, containing 36 questions, and is one of the most widely used and studied generic PRO questionnaires in SLE. It measures various domains from the WHO quality of life list of domains. The questionnaire is scored from 0 (worse health) to 100 (better health), capturing eight domains: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health. From these domains, the PCS and MCS are calculated. Since the instrument has been introduced in the late 1980s, its psychometric measurement properties across cultures and languages have been well established and it continues to be one of the most common questionnaires included in clinical research studies and RCTs. The reliability (internal consistency and test-retest reliability), validity, and responsiveness of SF-36 have been extensively studied.

It is important to note that PRO measures, such as the SF-36, are to be interpreted in the context of minimal clinically important differences (MCIDs). MCIDs are an important concept used to determine whether a medical intervention improves outcomes perceived by patients to be clinically meaningful. A clinical intervention is considered clinically meaningful if that PRO score change at follow-up meets or exceeds the MCID for that instrument in that disease. MCIDs for improvement and worsening in SF-36 have been established in SLE, as well as MCIDs for FACIT: ≥2.5 point increases and ≥−0.8 decreases in PCS and MCS scores for improvement and worsening, respectively, ≥5.0 point increases and ≥−2.5 decreases in SF-36 individual domain scores for improvement and worsening, respectively, and ≥4.0 MCID for FACIT improvement; and were applied to the analyses of SF-36 in the BLISS trials. Mean SF-36 domain scores in patients at baseline and week 52 were compared with age- and gender-matched healthy US subjects. The well-established MCID definition of improvements of ≥10 points on a 0–100 VAS was used for EQ-5D.

A combined analysis across both RCTs revealed that at week 52 mean improvements in SF-36 PCS scores were greater in SRI-4 responders vs non-responders (4.9 vs 2.6; $P<0.001$) and exceeded MCID. Similarly, at week 52, mean improvements in SF-36 MCS were greater in SRI-4
responders vs non-responders (4.4 vs 1.7; \( P<0.001 \)), and also exceeded MCID. In addition, all SF-36 domain scores showed statistically significant improvements in responders vs non-responders \( (P<0.001) \).

Further post-hoc analysis was performed. At 24 weeks, the first major prespecified endpoint of both BLISS trials demonstrated that mean improvements from baseline in SF-36 PCS did not differ significantly between treatment arms with belimumab 10 mg/kg and placebo (+3.34 vs +3.26) in BLISS-52 and (+3.21 vs +3.35) in BLISS-76. However, by week 52, significant improvements were reported in BLISS-52 PCS scores (+4.19 vs +2.84; \( P=0.0247 \)) but not in BLISS-76 (+3.44 vs +2.85). In BLISS-52, improvements were noted with the 1 mg/kg and 10 mg/kg belimumab doses in the following SF-36 domains at 52 weeks: physical functioning, bodily pain, and role-emotional. With the 1 mg/kg dose, there were improvements in the social functioning and general health domains as well. In BLISS-76, statistically significant improvements were reported with the 1 mg/kg dose in the following domains: role-physical, bodily pain, general health, and vitality at week 52.

### Functional Assessment of Chronic Illness Therapy-Fatigue Scale

The FACIT-F is a 13-item fatigue questionnaire that assesses self-reported aspects of physical and mental fatigue and impact on function and daily living. Items are scored using a four-point Likert scale. The score ranges from 0 to 52 and lower scores represent more fatigue. The FACIT-F was originally developed in 1997 for the assessment of fatigue in patients with anemia, and validated for use in rheumatic diseases. The content validity of the FACIT-F was confirmed by Kosinski et al in lupus patients. Other psychometric properties of FACIT-F, reliability, validity, and responsiveness have been demonstrated in several trials. In the SLE population, the instrument’s ability to detect change over time has been consistently demonstrated.

Similar to the outcomes of the SF-36 in the BLISS trials, at week 52, SRI-4 responders reported higher mean improvements in FACIT-F scores (5.2) compared to non-responders (3.0), but only values reported by SRI-4 responders exceeded the MCID. Both belimumab treatment groups demonstrated significant improvements by week 52 compared to placebo, with greater improvements by week 8 in responders, which were sustained through week 52. In post-hoc subgroup analyses of FACIT-F in Hispanic or Latino patients in the BLISS-SC trial, improvements were reported over time in both treatment groups and mean changes from baseline were significantly greater in the belimumab compared with placebo groups at weeks 8, 36, and 52, but not weeks 4, 12, and 24. At week 52, more patients who received belimumab reported improvements that equaled or exceeded the MCID in FACIT-F scores ≥4 compared with placebo (44.4% vs 36.1%; OR 1.42, 95% CI 1.05–1.94; \( P=0.0245 \)).

### EuroQol 5 Dimensions

The EQ-5D is composed of a six-item questionnaire and a VAS, allowing patients to rate their overall health. The VAS is scored from 0 to 100, reflecting “worst imaginable health” to “best imaginable health”. The EQ-5D VAS score permits an economic evaluation, through the calculation of quality-adjusted life-years. The questionnaire portion is a three-item response scale, reporting results as a summary score from 0 to 1, reflecting “death or health worse than death” to “best imaginable health”. The reliability and validity of the EQ-5D have been established in the SLE population and it is commonly used in RCTs. The EQ-5D was evaluated in the BLISS-52 and BLISS-76 RCTs, with changes in the pain/discomfort domain noted.

### Discussion

In the phase III BLISS RCTs, the primary endpoints were met and a combined analysis of SRI-4 responders vs non-responders across all treatment groups demonstrated statistically significant and clinically meaningful improvements in PROs. These findings supported the approval of belimumab for treatment of patients with adult autoantibody-positive SLE.

As per OMERACT recommendations, both generic and disease-specific instruments should be used to gauge HRQoL; disease-specific PRO instruments can complement generic ones such as SF-36. While generic instruments allow comparison to other diseases, disease-specific ones may contribute additional information. The ultimate goal with these PROs is to identify areas that could potentially be targeted with therapeutic interventions.

Generic PRO questionnaires frequently evaluated in SLE include the Medical Outcomes Study SF-36, EQ-5D, and the FACIT-F. Several HRQoL domains of importance to SLE patients, such as appearance, confidence, and body image, are not represented in the generic PRO tools such as SF-36. SLE-specific PRO instruments have been developed to capture these domains. These SLE-specific tools include the LupusQoL, SLE Symptom Checklist (SSC), SLE Quality of Life (SLEQoL), Lupus Quality of Life (L-QoL), Lupus Patient-Reported Outcome tool (LupusPRO), Lupus Impact
Tracker (LIT), Simple Measure of Impact of Lupus Erythematosus in Youngsters (SMILEY), and Body Image in Lupus Scale (BILS).12

In SLE, pain, fatigue, and physical function are frequently assessed both by generic PROs, such as the SF-36 and EQ-5D, and by the SLE-specific LupusQoL and LupusPRO questionnaires.2 The review by Holloway et al in 2014 showed that the SF-36 v2, the LupusQoL, and the FACIT-F demonstrated content and face validity and overall strong psychometric properties in an SLE population.25 More recently, in 2018, Izadi et al reviewed the literature to assess the responsiveness of different PROs to interventions in RCTs. This review confirmed that SF-36 and LupusQoL have strong evidence for adequate psychometric measurement properties.26 In addition, the LIT and Patient-Reported Outcomes Measurement Information System (PROMIS) showed strong evidence for measurement properties; however, PROMIS responsiveness has not been studied in lupus, nor have LupusPRO or LIT been assessed in RCTs.

LupusQoL is an SLE-specific HRQoL capturing the following eight HRQoL domains: physical health, emotional health, body image, pain, planning, fatigue, intimate relationships, and burden to others. The LupusQoL has been validated in multiple language cohorts and countries.2 Despite its use in multiple longitudinal observational studies and two RCTs,7 the LupusQoL was not evaluated in the belimumab trials. Inclusion of SLE-specific PRO measurements in clinical trials, such as the LupusQoL, can potentially represent patients’ perspectives of effects of disease and/or treatment on pertinent HRQoL domains not captured by generic metrics such as SF-36. LupusQoL includes additional pertinent domains to patients with SLE: sleep, body image, and sexual health. Two phase III RCTs, assessing the efficacy of epratuzumab in lupus, have included LupusQoL.59,60

To conclude, PRO outcomes in BLISS trials have been assessed as secondary endpoints in a responder–nonresponder analysis, which confirmed that clinical improvements measured by the SRI-4 were associated with significant and clinically meaningful reported improvements in HRQoL and fatigue. Thus, it is very important to include PROs, including generic and disease-specific measures, such as SF-36, LupusQoL, and FACIT-F, as secondary endpoints in trials assessing the efficacy of new drugs in lupus. RCTs in SLE have mainly utilized PROs to assess HRQoL, fatigue, and pain; however, as the field rapidly evolves and more trials develop, other PROs can be included to address cognitive function, anxiety, and depression, among other domains.

Disclosure
Dr Z Touma holds a Young Investigator Salary and Operating Awards provided by the Arthritis Society and the Physicians’ Services Incorporated Foundation. Dr Z Touma’s research is supported by a CRA (CIORA)-Arthritis Society Clinician Investigator Award. The University of Toronto Lupus Clinic is supported the Lupus Foundation of Ontario, the Lou Rocca Family, and the Kaiser’s Family. Dr Z Touma has consulted for GlaxoSmithKline, Merck Serono, Pfizer, and Janssen Pharmaceuticals. Dr V Strand serves as a consultant to Human Genome Sciences, now GSK, as well as Abbvie Amgen Corporation, AstraZeneca, BMS, Boehringer Ingelheim, EMD Serono, Genentech/Roche, Glenmark, Janssen, Kypha, Lilly, Novartis, and UCB. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

References


