



# Medicine Based Evidence for Individualized Decision Making: Case Study of Systemic Lupus Erythematosus

Ashley E. Wivel, MD, Kate Lapane, PhD,<sup>a</sup> Christi Kleoudis, MS,<sup>b</sup> Burton H. Singer, PhD,<sup>c</sup> Ralph I. Horwitz, MD<sup>d</sup>

<sup>a</sup>Lewis Katz School of Medicine, Temple University, Philadelphia, Pa; <sup>b</sup>University of Massachusetts Medical School, Amherst; <sup>c</sup>PAREXEL, International, Waltham, Mass; <sup>d</sup>University of Florida, Gainesville.

## ABSTRACT

**BACKGROUND:** To guide management decisions for an index patient, evidence is required from comparisons between approximate matches to the profile of the index case, where some matches contain responses to treatment and others act as controls.

**METHODS:** We describe a method for constructing clinically relevant histories/profiles using data collected but unreported from 2 recent phase 3 randomized controlled trials assessing belimumab in subjects with clinically active and serologically positive systemic lupus erythematosus. Outcome was the Systemic lupus erythematosus Responder Index (SRI) measured at 52 weeks.

**RESULTS:** Among 1175 subjects, we constructed an algorithm utilizing 11 trajectory variables including 4 biological, 2 clinical, and 5 social/behavioral. Across all biological and social/behavioral variables, the proportion of responders based on the SRI whose value indicated clinical worsening or no improvement ranged from 27.5% to 42.3%. Kappa values suggested poor agreement, indicating that each biological and patient-reported outcome provides different information than gleaned from the SRI.

**CONCLUSION:** The richly detailed patient profiles needed to guide decision-making in clinical practice are sharply at odds with the limited information utilized in conventional randomized controlled trial analyses. © 2017 Elsevier Inc. All rights reserved. • *The American Journal of Medicine* (2017) 130, 1290–1297

**KEYWORDS:** Clinical decision-making; Evidence based medicine; Medicine based evidence; Systemic lupus erythematosus

SEE RELATED ARTICLE, p. 1246.

## INTRODUCTION

The starting point of an evidential base to manage an individual patient is a longitudinal patient record, the patient history. At multiple points, a clinician may consider starting, stopping, or modifying particular treatments. At these times, the evidence that guides the clinician are the follow-up results of the contemplated intervention to patients whose histories—prior to the intervention—closely match the history

of the patient at hand. In the practice of clinical medicine, this is the relevant comparison group. Within the archive of matched histories, the clinician also wants to know what happened to patients who received different interventions, or none at all at a comparable time in the clinical course.

As the profile of the patient evolves further, the clinician may consider new, or revised, interventions. At each point where an intervention is contemplated, evidence to guide the decision can be acquired by examining the performance of the intervention on a patient set that approximately matches the patient at hand. The comparison histories will likely be from different sets of patients as time passes, because histories that match those of the given patient up to a first decision point may not be good matches when you consider longer histories up to a second, third, or later decision point. We will refer to the above general strategy as Medicine Based Evidence, emphasizing the point that in clinical practice, it is

**Funding:** None.

**Conflicts of Interest:** AEW, CK, and RIH were employed by GlaxoSmithKline (GSK); KL and BHS were consultants to GSK.

**Authorship:** All authors had access to the data and contributed to writing the manuscript.

Requests for reprints should be addressed to Ralph I. Horwitz, MD, Lewis Katz School of Medicine, Temple University, 3500 N. Broad Street, Suite 1111E, Philadelphia, PA 19140.

E-mail address: [ralph.horwitz@temple.edu](mailto:ralph.horwitz@temple.edu)

information about a single patient that drives the assembly of closely comparable patient histories with their responses to various interventions used as guides for decisions about treatment for the patient at hand.

The first step toward implementing Medicine Based Evidence is constructing an archive of patient histories, tuned to the information needs of clinical decision-making. We illustrate the process in systemic lupus erythematosus, using information collected in 2 randomized controlled trials (RCTs) of the drug belimumab. That the raw information comes from RCTs is beside the point. It could just as well derive from patient records in clinical practices or observational studies of systemic lupus erythematosus patients. The challenge is specification of the multidimensional histories themselves.

The purposes of this paper are to: 1) illustrate a method for constructing a taxonomy of multidimensional patient histories where systemic lupus erythematosus is the disease of interest and 2) describe how this more nuanced information guides management for an individual patient. In the Appendix (available online), we further clarify the distinction between information routinely reported in RCTs as part of evidence based medicine and information tuned to clinical practice (Medicine Based Evidence) that is often available, but not used, in RCTs.

## RESEARCH METHODS

### Systemic Lupus Erythematosus and Patient Histories

Systemic lupus erythematosus is a chronic autoimmune disease characterized by unpredictable flares, irreversible organ damage, adverse impacts on health-related quality of life, and increased mortality.<sup>1-4</sup> The waxing and waning clinical course of systemic lupus erythematosus along with the breadth of organ systems involved make it challenging to reliably measure clinical outcomes and demonstrate the effectiveness and safety of new and established treatments.

In this study, we combine data from 2 phase 3, multi-center, randomized, double-blind, placebo-controlled trials evaluating the efficacy and safety of belimumab in subjects with serologically positive systemic lupus erythematosus (Belimumab in Subjects with Systemic Lupus Erythematosus [BLISS]-52 and BLISS-76). Further details regarding BLISS-52 and BLISS-76 are provided below.<sup>5,6</sup> Briefly, dosing with the study agent occurred on Days 0, 14, and 28, then every 28 days through 48 weeks for BLISS-52 and 72 weeks for BLISS-76. In both studies, the primary endpoint was measured at week 52 using the Systemic lupus erythematosus

Responder Index (SRI), where a response means that there is a reduction of at least 4 points in the Safety of Estrogens in Lupus Erythematosus National Assessment-Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI) score; no new British Isles Lupus Assessment Group (BILAG) A or not more than one new BILAG B organ scores, and no worsening on the Physician Global Assessment score (increase <0.3).

The phase III studies demonstrated the efficacy of belimumab in this seropositive population of systemic lupus erythematosus patients based on the SRI endpoint. In BLISS-52, SRI response rate was statistically significantly higher in the belimumab 1-mg/kg (51.4%) and belimumab 10-mg/kg (57.6%) groups when compared with the placebo group (43.6%). In BLISS-76, the SRI response rate was 33.5% in the placebo group, 40.6% in the belimumab 1-mg/kg group, and 43.2% in the belimumab 10-mg/kg group, with the higher belimumab dose meeting prespecified levels of statistical significance.

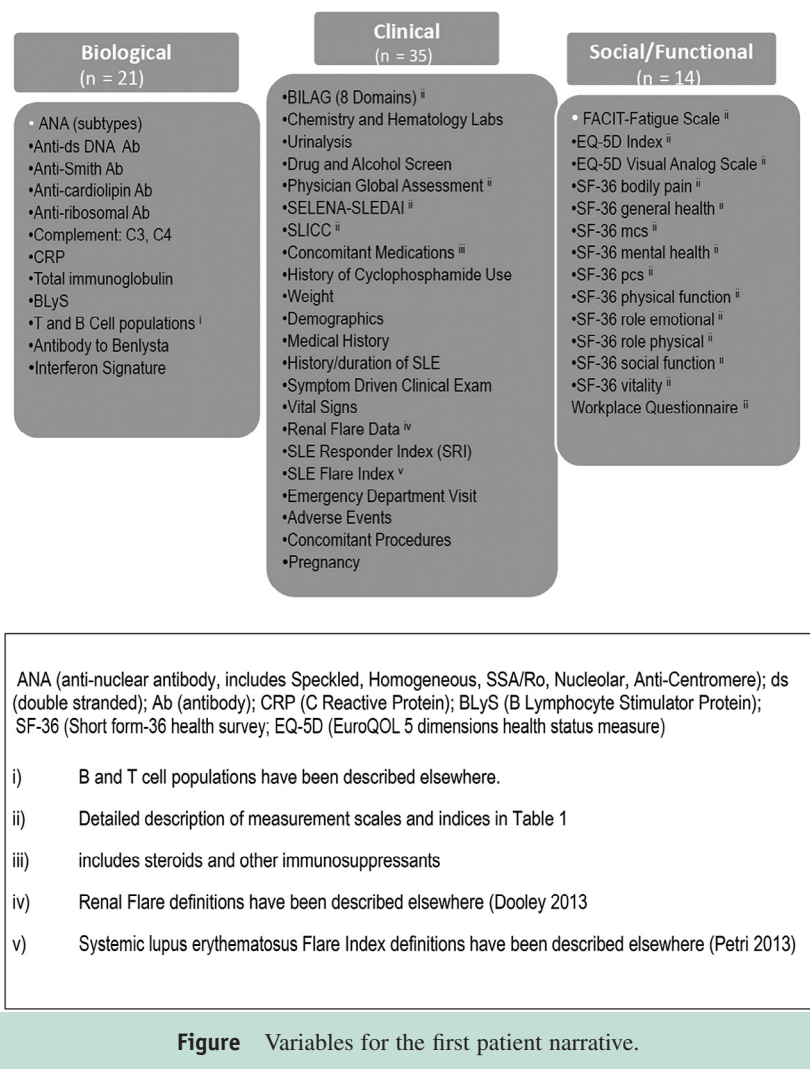
This conventional summary of the results masks the rich longitudinal data on patient experience in BLISS-52 and BLISS-76. This point is elaborated in detail in the Appendix (available online). BLISS-52 and BLISS-76 included 14 and 20 longitudinal assessments, respectively, during the on-treatment periods of the trials. They occur at baseline, 14 days, 28 days, and then every 28 days thereafter. The variables assessed at follow-up visits include: vital signs, weight, symptom-driven physical examination, hematology and chemistry assessments, pregnancy test, urinalysis, complement levels (eg, C3 and C4), anti-double-stranded DNA (anti-dsDNA), disease activity scales (eg, SLEDAI, systemic lupus erythematosus Flare Index, Physicians Global Assessment, BILAG), Short Form-36 health survey (SF-36), Functional Assessment of Chronic Illness Therapy (FACIT) Fatigue Scale, EuroQOL 5 dimensions health status measure (EQ-5D), workplace productivity questionnaire, emergency room visit questionnaire, adverse events, and concurrent medications. Autoantibodies (eg, anti-Smith, antiribosomal, antinuclear antibodies) and immunoglobulin levels were also measured, but less frequently.

### Constructing Representations of Patient Histories

Our 2 data sets contain individual responses on variables described above at each of up to 14 assessment times in BLISS-52 and 20 in BLISS-76. For inclusion in our analyses, we require that a patient be assessed at least 5 times, including an assessment close to the end date of the trial. Construction of a taxonomy of patient histories, our primary objective, requires a priori specification of organizing principles and a minimal information set to characterize the disease (in this

## CLINICAL SIGNIFICANCE

- Evidence to guide treatment decisions for a given patient is a multidimensional longitudinal profile incorporating biological, clinical, psychological, and social environmental information idiosyncratic to the particular patient.
- We developed a taxonomy of systemic lupus erythematosus patient histories in the context of response to a new therapy that provides an example of how far more nuanced information than that customarily made available from randomized controlled trials can be tuned to the needs of clinical practice.



case, systemic lupus erythematosus) dynamics. Three broad categories of variables were specified: biological, clinical, and social/functional. To obtain a succinct description of an individual patient, the first step is writing a narrative description of a single patient's record. It was decided to start with a narrative because it more closely resembles the narrative approach clinicians use to integrate the whole patient's experience.

**The First Patient.** A female subject on belimumab who completed BLISS-76 and failed to meet the primary endpoint was randomly selected from the dataset. As the first narrative was written, the clinician eliminated many variables based on her experience, as well as by review of literature and consultation with other clinicians about information they found useful for guiding management of individual systemic lupus erythematosus patients.

Seventy variables from biological, clinical, social, and functional domains were included in the analysis for the first patient (Figure and Table 1). Extensive discussion led to consensus for which variables in the original data set to discard as a *first proposal* for a clinically meaningful data set to guide management of systemic lupus erythematosus patients.

**An Iterative Process Using Increasing Numbers of Patients.** Construction of the second set of narratives was based on experience with the first subject. Twenty-one subjects who completed BLISS protocols were reviewed first and then the methodology was applied to 7 subjects who had not completed the study. To reduce bias based on preconceived ideas about treatment impact, study treatment was abstracted after all other data had been reviewed and summarized.

Variables were classified into 3 categories: baseline variables to describe patient characteristics and disease activity; excluded variables not considered in future analyses; and trajectory algorithm variables for future quantitative analyses (see Table 2).

Many of the clinical variables of interest (eg, BILAG, SELENA-SLEDAI, SF-36 domain scores) were incorporated into the baseline variable category. The Systemic Lupus International Collaborating Clinics score was designated a baseline variable because there were only 1-2 postbaseline measurements. Other baseline variables were collected only at entry into the study (eg, duration of systemic lupus erythematosus, demographics, history of cyclophosphamide use).

**Table 1** Detailed Description of Measurement Scales and Indices

Variable	Description	Scale
Physician Global Disease Assessment (PGA)	10-cm visual analog scale to capture physician assessment of a patient's current disease activity with anchors at 0 (none), 1 (mild), 2 (moderate), and 3 (severe).	0-3 (higher scores indicate higher disease activity)
British Isles Lupus Assessment Group (BILAG) disease activity index	Disease activity scale developed as a transitional index. Focuses on identification of disease activity that may require treatment. Includes 86 items that must be attributable to SLE and related to the last 4 weeks. Includes 8 organ-based domains: general (eg, constitutional symptoms such as pyrexia, fatigue, anorexia); mucocutaneous; neurological; musculoskeletal; cardiovascular and respiratory; vasculitis; renal; and hematology.	A: Disease sufficiently active to require disease-modifying treatment (prednisolone >20 mg or immunosuppressants). B: Disease less active than in "A"; mild reversible problems requiring only symptomatic therapy such as antimalarials, nonsteroidal anti-inflammatory drugs, or prednisolone <20 mg/day C: Stable mild disease D: System previously affected but currently inactive E: System never involved
Safety of Estrogens in Lupus Erythematosus National Assessment- Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI)	Includes 24 variables and measures disease activity within the last 10 days.	0-105 (higher scores indicate higher disease activity)
Systemic Lupus International Collaborating Clinics (SLICC) damage index	Instrument developed for assessment of accumulated permanent damage not related to active inflammation with a requirement that an item be present for at least 6 months to be scored. Includes 39 items across 12 organ systems, including: ocular, neuropsychiatric, renal, pulmonary, cardiovascular, peripheral vascular, gastrointestinal, musculoskeletal, skin, premature gonadal failure, diabetes, malignancy (except dysplasia).	0-46 (higher scores indicate higher damage levels)
SLE Responder Index (SRI)	Composite endpoint developed for the belimumab phase III studies, which includes: 1) a reduction of at least 4 points in the SELENA-SLEDAI disease activity score 2) no new BILAG A or no more than 1 new BILAG B organ scores 3) no worsening on the Physician Global Assessment (PGA) score (increase <0.3)	N/A
Functional Assessment of Chronic Illness Therapy (FACIT)-fatigue scale	Includes 13 items covering the last 7 days.	0-52 (higher scores indicate less fatigue)
EuroQOL 5 dimensions (EQ-5D) health status measure	5-item index and one visual analog scale (VAS), related to the date of administration of the scale. The 10-cm VAS is anchored at 0 (worst imaginable health state) and 100 (best imaginable health state). The index includes 5 questions related to mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.	VAS: 0-100 Index: 0-1 (higher scores indicate lower disease activity)
SF-36	Validated questionnaire that evaluates general health-related quality of life across 8 domains, including bodily pain, general health, mental health, physical function, role-emotional, role-physical, social function, and vitality. It also includes a mental composite score (MCS) and physical composite score (PCS).	Scoring based on norms derived from the 1998 Survey of Functional Health Status in the US with transformed scores of 50 representing the mean with a standard deviation of 10. (higher scores indicate better health in the domain or composite score).
Workplace Productivity Questionnaire	3 questions asking about employment in the past 4 weeks. Score represents the percentage of planned work days missed due to illness.	0-1 (higher score reflects less work missed due to illness)

BILAG = British Isles Lupus Assessment Group; SF-36 = Short Form-36 health survey; SLE = systemic lupus erythematosus.

**Table 2** Final Categorization of Variables

Variable Domain	Baseline Variables (n = 33)	Excluded Variables (n = 26)	Trajectory Analysis Variables (n = 11)
Biological	<ul style="list-style-type: none"> <li>■ ANA</li> <li>■ Anti-Smith antibody</li> <li>■ Anti-ribosomal antibody</li> <li>■ Anti-cardiolipin immunoglobulin (Ig)A, IgG, IgM</li> <li>■ BLYS protein</li> </ul>	<ul style="list-style-type: none"> <li>■ Total IgA, IgG, IgM</li> <li>■ T- and B-cell populations</li> <li>■ Antibody to benlysta</li> <li>■ Interferon signature</li> </ul>	<ul style="list-style-type: none"> <li>■ Anti-dsDNA antibody</li> <li>■ C3</li> <li>■ C4</li> <li>■ CRP</li> </ul>
Clinical	<ul style="list-style-type: none"> <li>■ BILAG (8 domains)</li> <li>■ Physician Global Assessment</li> <li>■ SELENA-SLEDAI</li> <li>■ SLICC</li> <li>■ History of SLE, including duration of disease</li> <li>■ Demographics</li> <li>■ History of cyclophosphamide use</li> </ul>	<ul style="list-style-type: none"> <li>■ Laboratory data: hematology, chemistry, urinalysis, drug and alcohol screen</li> <li>■ Vital signs</li> <li>■ Weight</li> <li>■ Symptom-driven clinical examination</li> <li>■ Medical history</li> <li>■ Renal flare data</li> <li>■ SLE Flare Index</li> <li>■ Emergency department visit</li> <li>■ Adverse events</li> <li>■ Concomitant procedures</li> <li>■ Pregnancy</li> </ul>	<ul style="list-style-type: none"> <li>■ SRI</li> <li>■ SLE medications, including steroids</li> </ul>
Social/functional	<ul style="list-style-type: none"> <li>■ SF-36 bodily pain</li> <li>■ SF-36 general health</li> <li>■ SF-36 mental health</li> <li>■ SF-36 physical function</li> <li>■ SF-36 role emotional</li> <li>■ SF-36 role physical</li> <li>■ SF-36 social function</li> <li>■ SF-36 vitality</li> </ul>	<ul style="list-style-type: none"> <li>■ Workplace questionnaire</li> </ul>	<ul style="list-style-type: none"> <li>■ FACIT-Fatigue Scale</li> <li>■ EQ-5D Index</li> <li>■ EQ-5D VAS</li> <li>■ SF-36 PCS</li> <li>■ SF-36 MCS</li> </ul>

ANA = antinuclear antibodies; BILAG = British Isles Lupus Assessment Group; BLYS = B-lymphocyte stimulator; C3 = complement C3; C4 = complement C4; CRP = C-reactive protein; EQ-5D = EuroQOL 5 dimensions health status measure; FACIT = Functional Assessment of Chronic Illness Therapy; MCS = Mental Component Summary; PCS = Physical Component Summary; SELENA-SLEDAI = Safety of Estrogens in Lupus Erythematosus National Assessment–Systemic Lupus Erythematosus Disease Activity Index; SF-36 = Short Form-36 health survey; SLE = systemic lupus erythematosus; SLICC = Systemic Lupus International Collaborating Clinics damage index; SRI = Systemic lupus erythematosus Responder Index; VAS = visual analog scale.

Variables were excluded from analysis for various reasons, including pragmatic considerations such as whether the variable would be utilized during normal clinical practice (eg, B- and T-cell subpopulations, interferon signature) or because there were no postbaseline measurements (eg, drug and alcohol screen). Other variables were excluded because it was impossible to determine whether the variable was specifically related to systemic lupus erythematosus (eg, adverse events, weight changes, workplace questionnaire). Other variables were eliminated due to redundancy.

Eleven trajectory algorithm variables were identified, including 4 biological, 2 clinical, and 5 social/functional. These are listed in **Table 2**. We included SRI as a clinical variable because it captured clinically significant improvement in the SELENA-SLEDAI and a lack of clinically significant worsening in both the BILAG and the Physician Global Assessment, reducing the dimensionality of patient histories. The biological variables we selected were likely to be measured in routine practice. We retained a number of patient-reported social and functional outcome variables because they provided direct in-

formation about how patients were functioning while participating in the studies. The FACIT-fatigue score was chosen because fatigue is a key symptom of systemic lupus erythematosus. The EQ-5D is a visual analog scale (VAS) score providing a global assessment by the patient of their health status as well as an index score summarizing health status across 5 domains. We chose the SF-36 mental and physical composite scores because they aggregate the 8 domain scores and allowed simplification of the number of variables.

**Defining Patient Experience Trajectories.** Graphs of responses on each variable over time suggest that changes in magnitude between one response value and the next (in time) was less important for characterizing a patient history than the percentage of time over a year that a patient spent, with response defined as improved, worsened, or stable compared with their baseline status. Here, “improved” or “worsened” means that a response at a particular time is greater in magnitude (either + or –) than a minimal clinically important difference for the variable. Then we defined a trajectory as



**Table 3** Baseline Disease Activity and Minimal Clinically Important Difference Definitions

Variable	High Disease Activity	Moderate Disease Activity	Low Disease Activity	MCID
<b>Biological</b>				
Anti-dsDNA	≥30 IU/mL	N/A	<30 IU/mL (WNL)	WNL ↔ ≥30 IU/mL or ± 10% of BL if in the abnormal range at BL*
C3	<900 mg/L	N/A	≥900 mg/L (WNL)	WNL ↔ <900 mg/L
C4	<16 mg/dL	N/A	≥16 mg/dL (WNL)	WNL ↔ <16 mg/dL
CRP	>3 mg/L	N/A	≤ 3 mg/L (WNL)	WNL ↔ >3 mg/L
<b>Clinical</b>				
SLEDAI†	≥10	≥6 to <10	<6	≥4
Prednisone equivalents	>7.5 mg/day	N/A	≤ 7.5 mg/day	>7.5 mg/day ↔ ≤7.5 mg/day
<b>Social/behavioral</b>				
FACIT‡	<30	≥30 to <42	≥42 to 52	≥4 point change
SF-36 Composite (MCS, PCS)§	<30	≥30 to <40	≥40	≥+2.5(improved) <sup>9</sup> ≥-0.8 (worsened) <sup>9</sup>
EQ-5D VAS	<65	≥65 to <80	≥80 to 100	≥7 point change
EQ-5D Index	<0.65	≥0.65 to <0.8	≥0.8 to 1	≥0.1 point change

Anti-dsDNA = anti-double-stranded DNA; BL = baseline; C3 = complement C3; C4 = complement C4; CRP = C-reactive protein; EQ-5D = EuroQOL 5 dimensions health status measure; FACIT = Functional Assessment of Chronic Illness Therapy; MCID = minimal clinically important difference; MCS = Mental Component Summary; PCS = Physical Component Summary; SF-36 = Short Form-36 health survey; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; VAS = visual analog scale; WNL = within normal limits.

↔ = Transition between WNL and abnormal levels.

Normal limits for laboratory values are based on definitions in the original Belimumab in Subjects with Systemic Lupus Erythematosus (BLISS)-52 and BLISS-76 studies.

\*MCID of ± 10% change from baseline selected after review of published analysis of abetimus sodium (which specifically reduces anti-dsDNA) that identified an inverse relationship between levels of anti-dsDNA antibodies (at both ≥10 % and ≥20% reduction from baseline) and SF-36 scores.

†In BLISS studies, subjects were required to have “active” disease defined as a screening SLEDAI score of ≥6 and at randomization subjects were stratified by their screening SLEDAI score (6-9 vs ≥10). MCID for SLEDAI score aligns to SRI endpoint requirement of a reduction ≥4 for classification as a responder.

‡Based on literature from a general population as well as oncology, rheumatoid arthritis (RA), and SLE patients.

§SF-36: normed mean = 50, standard deviation (SD) = 10. Low, moderate, and high disease activity were defined to be within 1 SD below the mean, between 1 and 2 SDs and >2 SDs below the mean, respectively.

||EQ-5D: Disease activity not available for systemic lupus erythematosus and was adapted from literature in other chronically ill populations. MCID is based on data in rheumatoid arthritis, psoriasis, and oncology patients.

improving or worsening if at least two-thirds of postbaseline measurements were improved (or worsened) by more than one minimal clinically important difference relative to the baseline measurement. A trajectory was *stable* if at least two-thirds of postbaseline measurements are within one minimal clinically important difference of the baseline measurement. Finally, we classified an important category of response histories as *variable* trajectories.

**Table 3** shows the minimal clinically important difference definitions. Most are based on published systemic lupus erythematosus data. If no specific literature was identified for a variable, definitions were derived from information from both healthy populations and populations with significant health-related quality-of-life impact, including other rheumatological conditions.

Based on the variables described above, the coarsest level of an information set we propose as part of an archive of systemic lupus erythematosus patient histories consists of a description of 11 trajectories for each patient, augmented by baseline information on biological (anti-dsDNA, C3, C4, C-reactive protein), clinical (SLEDAI), and social/functional (FACIT-fatigue; EQ-5D index and VAS; SF-36 Mental Component Summary and Physical Component Summary [PCS]) variables. The trajectory representations

for the 21 patients for whom narratives were constructed are displayed in **Table 4**. In addition, **Table 4** includes characterization of responder status based on a single time point for comparison with the trajectory data.

Biological and self-reported health measures trajectories can develop in opposite directions. For example, patient #16 is improving on anti-dsDNA and worsening on SF-36 PCS. Patient #20 is worsening on anti-dsDNA and improving on EQ-5D index, SF-36 Mental Component Summary, and SF-36 PCS. Note, also, that RCT outcome classification of a patient as a responder is not necessarily the same as the patient's SRI trajectory classification. A more nuanced version of SRI trajectory information can be conveyed to the clinician by listing the patient's SRI response category (Yes or No) at each of her assessment times. This listing can be viewed as one of 11 graphical presentations for each cell in the trajectory columns where actual scores vs time are plotted. This added level of detail provides a basis from which a clinician can prescribe conditions for a given patient that would be the basis for identifying approximate matches to that patient from an archive containing records analogous to the content of **Table 4**, augmented by the graphical displays.

**Table 4** Summary of Variable Trajectories for 21 Subjects

Subject ID	Biological Variables				Social and Functional Variables					Clinical Variable*	Responder Status†
	Anti-dsDNA	C3	C4	CRP	FACIT-Fatigue	EQ-5D VAS	EQ-5D Index	SF-36 MCS	SF-36 PCS	SRI Trajectory	SRI at Single Time Point
1	SH	∇	V	V	V	∇	V	∇	©©	Y	Y
2	SL	SL	SL	SL	V	SL	SL	V	∇	N	N
3	SH	SH	∇	SH	SM	∇	SM	SM	∇	N	Y
4	V	SL	SL	SH	SH	SL	∇	V	SM	Y	Y
5	∇	SL	SH	SH	SM	SL	SM	∇	V	Y	Y
6	SH	V	SH	©	∇	∇	∇	∇	©	N	N
7	∇	SL	SL	SL	SL	SM	SL	V	V	N	Y
8	∇	SL	V	SH	V	©	V	V	©	N	N
9	SL	SL	SH	SH	SH	SM	V	SH	V	Y	Y
10	SH	SH	SH	SL	∇	∇	V	∇	∇	Y	Y
11	SH	©	SH	SL	∇	©	©	∇	V	N	N
12	SH	SH	SH	SL	V	SH	V	©	∇	Y	Y
13	SH	SH	∇	SL	∇	SL	SL	∇	∇	Y	N
14	SH	∇	SH	SH	∇	∇	SM	∇	V	N	Y
15	∇	∇	V	SL	V	V	V	V	V	Y	Y
16	∇	SL	©	V	V	V	SL	V	©	N	N
17	V	SL	SH	SL	©	SH	SH	©	∇	N	N
18	SH	SL	SL	SL	V	V	SL	©	V	Y	Y
19	SL	SL	SL	SH	SM	SL	SH	∇	SL	Y	Y
20	©	SH	SH	SH	∇	V	∇	∇	∇	N	N
21	V	SH	SH	SH	∇	∇	V	V	V	N	N

∇ = Improved; anti-dsDNA = anti-double-stranded DNA; C3 = complement C3; C4 = complement C4; CRP = C-reactive protein; EQ-5D = EuroQOL 5 dimensions health status measure; FACIT = Functional Assessment of Chronic Illness Therapy; MCS = Mental Component Summary; PCS = Physical Component Summary; SF-36 = Short Form-36 health survey; SH = stable high; SM = stable moderate; SL = stable low; SRI = Systemic lupus erythematosus Responder Index; V = variable; © = Worse; Y = Yes; N = No.

\*Yes = SRI criteria met at ≥ 2/3 of follow-up assessments over either 52 or 76 weeks for Belimumab in Subjects with Systemic Lupus Erythematosus (BLISS)-52 and BLISS-76, respectively. No = SRI criteria met for <2/3 followup assessments over either 52 or 76 weeks for BLISS-52 and BLISS-76, respectively.

†Yes = SRI criteria met at either 52 or 76 weeks for BLISS-52 and BLISS-76, respectively. No = SRI criteria not met at either 52 or 76 weeks for BLISS-52 and BLISS-76, respectively.

In decision-making for an individual patient, think of this set of trajectories as a small prototype archive of cases for comparison purposes. Assume, for illustrative purposes, that the patient at hand is in the *improving* category. In addition, assume that a decision about further management of the patient is to be made 6 months into the observation period. Thus, only the first 6 months of a particular full trajectory is known to the clinician. Now, for comparison purposes, the clinician identifies other *improving* cases that match her patient, as well as some variable cases that are improving in the first 6 months. If some of these archived cases received a treatment the clinician contemplates for her patient while others did not receive it, then the information to guide management of the patient is the follow-up results in the comparison cases. If the clinician again wants to make comparisons after 9 months, some of the previously used trajectories will not enter her comparison group. They are cases of variable trajectories that could not be seen as such after only 6 months.

## DISCUSSION

We presented an approach to construction of longitudinal patient profiles that can serve as primary evidence to guide the practice of individualized medicine. Our detailed development of a taxonomy of systemic lupus erythematosus histories in response to a new therapy provides a concrete example of how far more nuanced information than that customarily made available from RCTs can be the basis for enhanced clinical decision-making.

Constructing multidimensional patient histories started with a labor-intensive process. Considerable expert judgment and qualitative assessment was required at multiple points in narrative writing, variable selection, and trajectory specification, and we acknowledge that a different group of researchers might have arrived at alternate definitions. Once explicit criteria are at hand for defining key baseline characteristics and subsequent trajectories, algorithms for automatically building up an archive of such patient experience records can be produced and routinely implemented.

To enable decision-making about a given patient, the clinician interrogates a large archive of patient histories, keying in criteria for an approximate match to her patient at a particular follow-up point. At multiple other points, it should be expected that different sets of comparative histories would be drawn from the archive as the patient experience progresses.

It should be possible to automate the initial phase of developing patient narratives by curating the clinical, biological,

and social/functional measurements and relying on an expert system to produce patient experience histories. Success in this area would provide a first step toward resolving a vexing dilemma posed by John Tukey 40 years ago.<sup>7</sup>

“It is a difficult task to drive the nearly incompatible two-horse team: on the one hand knowledge of a most carefully evaluated kind (the RCT), where, in particular, questions of multiplicity (multiple comparisons) are faced up to; and, on the other hand, informed professional opinion, where impressions gained from statistically inadequate numbers of cases often, and so far as we can see, often should, control the treatment of individual patients. The same physician or surgeon must be concerned with what is his knowledge (gleaned from rigorous population studies) and what is his informed professional opinion, often as part of treating a single patient. I wish I understood better how to help in this essentially ambivalent task.”

## References

1. Yazdany J, Yelin E. Health related quality of life and employment among persons with systemic lupus erythematosus. *Rheum Disease Clin North Am.* 2010;36(1):15-32.
2. McElhone K, Abbott J, Gray J, Williams A, Teh L-S. Patient perspective on systemic lupus erythematosus in relation to health-related quality of life concepts: a qualitative study. *Lupus.* 2010;19(14):1640-1647.
3. Gallop K, Nixon A, Swinburn P, Sterling KL, Naegeli AN, Silk MET. Development of a conceptual model of health-related quality of life for systemic lupus erythematosus from the patient's perspective. *Lupus.* 2012;21(9):934-943.
4. Rahman A, Isenberg DA. Systemic lupus erythematosus. *N Engl J Med.* 2008;358(9):929-939.
5. Navarra SV, Guzman RM, Gallacher AE, et al. RF for the BLISS-52 study group. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomized, placebo-controlled, phase 3 trial. *Lancet.* 2011;377(9767):721-731.
6. Furie R, Petri M, Zamani O, et al. RF for the BLISS-76 study group. 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.
7. Tukey JW. Some thoughts on clinical trials, especially problems of multiplicity. *Science.* 1977;198:679-684.

## SUPPLEMENTARY DATA

Supplementary Material accompanying this article can be found in the online version at [doi:10.1016/j.amjmed.2017.06.013](https://doi.org/10.1016/j.amjmed.2017.06.013).



## APPENDIX

### Comparison of Information in Standard RCT Reporting with that in Patient Histories

In this Appendix, we show where the information used in our patient history constructions is qualitatively different and where it is similar to what is used in a conventional randomized controlled trial (RCT) analysis.

**Study Sample.** First, we evaluated the extent to which the Belimumab in Subjects with Systemic Lupus Erythematosus (BLISS)-52 and BLISS-76 study participants ( $n = 1684$ ) had missing baseline data for each variable considered for the analysis. We required a baseline value for each feature to determine clinically relevant differences over time. From this process, 97 subjects were deemed ineligible for the current analysis.

Next, we examined the distribution of postbaseline measurements available in the pooled analytic dataset of the 2 trials. Based on this review, we included participants who had at least 4 postbaseline assessments for features of interest: complement C3 (C3), C4, anti-double-stranded DNA; C-reactive protein (CRP), Functional Assessment of Chronic Illness Therapy (FACIT), Short-Form 36 (SF-36) (Mental Component Summary [MCS] and Physical Component Summary [PCS]), EuroQOL 5 dimensions (EQ-5D) health status measure visual analog scale (VAS) and Index, and Systemic lupus erythematosus Responder Index (SRI;  $n = 1326$ ). We excluded 129 subjects who lacked data for week 52 assessment. We required at least 2 assessments in the first 6 months of the study and at least 2 assessments in the second 6 months. Twenty-two subjects did not meet these criteria and were also excluded. The remaining 1175 participants comprised the study sample, described in [Supplementary Table 1](#).

We compared our study sample to the clinical trial “completers” who were defined in the Belimumab in Subjects with Systemic Lupus Erythematosus (BLISS) study protocols as those who completed at least 48 weeks of treatment and had a Week 52 efficacy evaluation.

Seventeen subjects who met our minimum eligibility criteria for the current analysis were considered trial noncompleters. These subjects either did not receive treatment at 48 weeks or withdrew late within the 1-year study period before a designated 52-week assessment was completed.

**Analytical Approach.** We evaluated the discordance between SRI response at 52 weeks and each of the biological and social/functional/behavioral variables at 52 weeks. Within each SRI response category, we calculated the proportion who improved or remained stable at low levels of disease activity based on each biological or social/functional/behavioral variable and the proportion who appeared to have worsened or remained stable at moderate to high disease activity. The percent agreement corrected for chance was estimated using kappa statistics with 95% confidence intervals provided.

For each biological and social/functional/behavioral variable, we estimated the proportion in each trajectory category

stratified by SRI response. We also estimated the overall proportion of those who were classified as clinically variable. For these comparisons, estimation of kappas was not possible, as the resulting tables of comparison were not square (took the form of  $2 \times 3$  tables). The comparison of SRI response at 52 weeks and SRI trajectory algorithm resulted in  $2 \times 2$  tables, and kappas (and 95% confidence intervals) were calculated. Lastly, we evaluated agreement between each biological variable with each social/functional/behavioral variable by estimating the kappa statistics and 95% confidence intervals. Kappa statistics indicating very poor or less than chance agreement would favor our hypothesis that social/functional/behavioral variables provide different information about the patient experience than biological variables.

**Concordance with SRI at 52 Weeks.** In [Supplementary Table 2](#), the distribution of the SRI index at week 52 is compared with each biological and social/behavioral variable measured at 52 weeks.

Using CRP as an example, 40.5% of patients classified as responders based on the 52-week SRI had CRP values that worsened or remained moderate or high at 52 weeks. Across all the biological and social/behavioral variables, the proportion of responders based on the SRI at 52 weeks whose value indicated clinically relevant worsening or whose levels remained stable and moderate to high at 52 weeks ranged from 27.5% for the EQ-5D VAS to 42.3% for the EQ-5D Index. Similarly, among those who were considered nonresponders based on the SRI at 52 weeks, there was lack of agreement in the individual biological and social/behavioral variables at 52 weeks. For the C4 variable that had the greatest observed agreement, 47.5% of nonresponders had C4 values consistent with clinically relevant improvements or remaining stable with low C4 values. The greatest extent of discordance among the nonresponders was observed for the SF-36 PCS measure (64.5%). The kappa values ranged from 0.03 for CRP (95% confidence interval [CI],  $-0.02$ - $0.08$ ) to  $-0.18$  for C4 (95% CI,  $-0.24$  to  $-0.13$ ), suggesting poor chance corrected agreement and indicating that each biological and patient-reported outcome provides different information than gleaned from the SRI when assessed at the same 52-week time point.

[Supplementary Figure](#) shows the distribution of the outcome based on the trajectory over 52 weeks of each biological and social/behavioral variable for patients who were considered responders based on the SRI index at week 52 (Panel A) and patients considered nonresponders based on the SRI index at week 52 (Panel B).

In Panel A, regardless of the biological or social/behavioral variable, only half of the patients were considered to have improved or remained stable with low disease activity based on their trajectories (range: 47.9% for SF-36 MCS to 62.7% for C3). The proportion of patients whose trajectories indicated worsening based on each biological or social/behavioral variable ranged from 19.4% on the FACIT to 35.8% on the EQ-5D Index. Patients whose trajectories were classified as clinically variable ranged from 7.5% for C4 to 29.9% on the

SF-36 PCS. Among patients who did not respond based on the SRI at 52 weeks (Panel B), discordance between the 52-week SRI measure and each trajectory distribution was observed. For those classified as nonresponders using the SRI measure at 52 weeks, many patients appeared to have improvements in the biologic measures based on the trajectory algorithms (40.9% for C4 to 61.7% for CRP). Similar patterns were observed for the social/behavioral variables. The discordance between the SRI index at 52 weeks and the trajectory method was apparent among nonresponders as well, with a range of 46.7% for EQ-5D VAS to 49.7% for FACIT appearing to improve or remain stable with low disease activity throughout the year.

The comparison of the SRI index at 52 weeks to the trajectory method for the SRI is shown in [Supplementary Table 3](#).

In responders, discordance between the 52-week SRI measure and the SRI trajectory distribution was observed. Among SRI responders at 52 weeks, 26.2% were nonresponders using the SRI trajectory method. Discordance among the nonresponders at 52 weeks was less (8.4%). The overall kappa was 0.63 (95% CI, 0.59-0.67).

The kappa values and corresponding 95% CIs provide a summary measure of agreement for each biologic variable and the social/behavioral variables ([Supplementary Table 4](#)).

Comparisons were made based on variables measured at 52 weeks and based on the experience over the year using the trajectory method. Regardless of the time period in which the comparisons were made, the biological variables do not tend to track along with the social/behavioral variables. For comparisons made on variables measured at 52 weeks, kappas ranged from  $-0.09$  (95% CI,  $-0.14$  to  $-0.03$ ) for C3 relative to the EQ-5D Index to  $0.13$  (95% CI,  $0.07$ - $0.19$ ) for CRP relative to SF-36 PCS, consistent with poor chance corrected agreement. Using the trajectory method, poor agreement was observed between biological and social/behavioral variables, with kappas ranging from  $-0.07$  (95% CI,  $-0.12$ - $0.03$ ) for C3 relative to the EQ-5D Index to  $0.03$  (95% CI,  $-0.01$ - $0.07$ ) for CRP relative to FACIT and EQ-5D VAS.

**Discussion.** Our comparisons of the information gleaned from conventional RCT analyses vs what has been learned from patient histories constructed using the rich longitudinal follow-up information from the RCT data sets emphasizes the sharp distinction between information needed for regulatory purposes and what is important to consider in clinical decision-making.

We first explored the alignment between biological and social/functional variables assessed at 52 weeks and responder status based on the SRI endpoint at 52 weeks. This was an important context to understanding whether any differences we noted when comparing the trajectory approach to the single time-point approach were due entirely to the strategy being applied or whether it also potentially reflected our hypothesis that each individual variable might

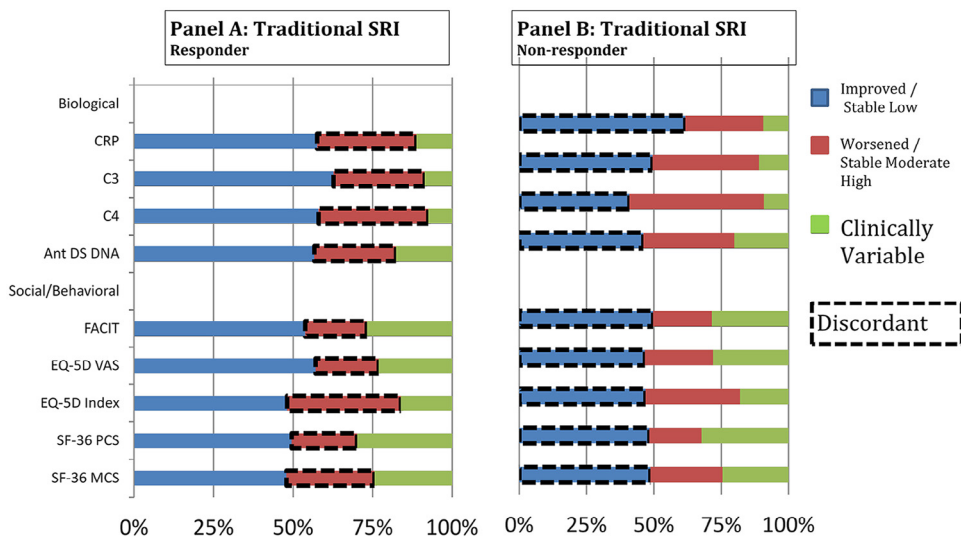
contribute different information than the SRI at 52 weeks. The kappa statistics for these comparisons were very poor (range:  $-0.18$  to  $0.03$ ), indicating less than chance agreement, strengthening the inference that the information content is distinctive.

Because both the British Isles Lupus Assessment Group and Systemic Lupus Erythematosus Disease Activity Index, components of the SRI endpoint, include anti-dsDNA, C3, and C4, one might have anticipated that improvement/stable low activity or worsening/stable moderate-high activity of the biological variables would be associated with SRI response or nonresponse, respectively. However, systemic lupus erythematosus is a complex disease involving multiple organ systems for which no single biomarker has been identified that can reliably predict clinical outcomes, and no biomarkers have been validated as a surrogate endpoint according to the US Food and Drug Administration guidance on systemic lupus erythematosus.<sup>1</sup> Our results are consistent with these prior observations, and we agree with Liu et al<sup>2</sup> in their recent summary of the challenges of identifying biomarkers for systemic lupus erythematosus that future efforts will likely lead to “composite panels” of biomarkers. Based on our results assessing the relationship between the biological and social/functional variables ([Supplementary Table 4](#)), we propose that panels of biomarkers alone will not be sufficient to understand and predict outcomes because the patient-reported outcome variables we evaluated did not uniformly trend in the same direction as the biological variables we incorporated. It is a reminder that patients may feel and function differently than their clinicians may assess and their laboratory data may indicate.

Both the moderate kappa,  $0.63$  (95% CI,  $0.59$ - $0.67$ ), for the comparison of the SRI assessed at 52 weeks vs the SRI trajectory approach and the visual presentation of the trajectory approach for the other variables ([Supplementary Figure](#)), suggest that incorporation of longitudinal follow-up information provides different information on a patient’s experience across 52 weeks than is provided by the snapshot of assessment at 52 weeks. The simple algorithm employed to classify patients allows one to contextualize the experience of a patient over the year. In a disease such as systemic lupus erythematosus, where flares are common, the trajectory approach allows one to think about how a patient is doing over time in a way that provides more detail than simply counting the number of flares.

## Supplementary References

1. U.S. Department of Health and Human Services, Food and Drug Administration. Guidance for Industry Systemic Lupus Erythematosus – Developing Medical Products for Treatment. Silver Spring, MD: U.S. Food and Drug Administration; June 2010.
2. Liu CC, Kao AH, Manzi S, Ahearn JM. Biomarkers in systemic lupus erythematosus: challenges and prospects for the future. *Ther Adv Musculoskelet Dis*. 2013;5(4):210-233.



**Supplementary Figure** The figure shows the distribution of the outcome based on the trajectory of each biological and social/behavioral variable for patients who were considered responders based on the Systemic lupus erythematosus Responder Index (SRI) at week 52 (Panel A) and patients considered nonresponders based on the SRI index at week 52 (Panel B). In Panel A, regardless of the biological or social/behavioral variable, only half of the patients were considered to have improved or remained stable or low based on their trajectories (range: 47.9% for Short Form [SF]-36 Mental Component Summary [MCS] to 62.7% for complement C3 [C3]). The proportion of patients whose trajectories indicated worsening based on each biological or social/behavioral variable ranged from 19.4% on the Functional Assessment of Chronic Illness Therapy (FACIT) to 35.8% on the EuroQOL 5 dimensions health status measure (EQ-5D) Index. Patients whose trajectories were clinically variable based on each variable ranged from 7.5% for C4 to 29.9% on the SF-36 Physical Component Summary (PCS). Among patients who did not respond based on the SRI at 52 weeks (Panel B), discordance between the 52-week SRI measure and each trajectory distribution was observed. For those classified as nonresponders using the SRI measure at 52 weeks, many patients appeared to have improvements in the biologic measures based on the trajectory algorithms (40.9% for C4 to 61.7% for C-reactive protein [CRP]). Similar patterns were observed for the social/behavioral variables. The discordance between the SRI index at 52 weeks and the trajectory method was apparent among nonresponders as well, with a range of 46.7% for EQ-5D VAS to 49.7% for FACIT appearing to improve or remain stable with low disease activity throughout the year.

**Supplementary Table 1** Baseline Sociodemographic, Clinical, and Social/Behavioral Factors by SRI Response at 52 Weeks (n = 1175)

	SRI Responder at 52 Weeks (n = 676)	SRI Nonresponder at 52 Weeks (n = 499)	Total (n = 1175)
	Percentage		
<b>Demographics</b>			
Age >45 years	24.9	27.9	26.1
Women	94.8	93.4	94.2
Non-White	53.9	48.3	51.5
Non-USA	78.6	69.1	74.6
<b>Study treatment</b>			
Belimumab 10 mg/kg	36.8	29.1	33.5
Belimumab 1 mg/kg	33.7	35.5	34.5
Placebo	29.4	35.5	32.0
<b>Clinical variables</b>			
PGA score >1.5 (0-3 VAS)	44.5	33.7	39.9
SELENA-SLEDAI ≥10	63.3	30.3	49.3
<b>Prednisone equivalent at baseline</b>			
>7.5 mg	55.8	48.1	52.5
≤7.5 mg	31.2	36.5	33.5
None	13.0	15.4	14.0
<b>Biological variables</b>			
Anti-dsDNA ≥30 IU/mL	68.1	69.9	68.9
C3 <900 mg/L	40.5	46.9	43.2
C4 <16 mg/dL	52.1	59.1	55.1
CRP ≥3 mg/L	40.7	36.7	39.0
<b>Social and functional variables</b>			
<b>FACIT-fatigue &lt;30</b>			
30-<42	42.8	43.7	43.2
42	37.9	37.1	37.5
42	19.4	19.2	19.3
<b>EQ-5D VAS &lt;65</b>			
65-<80	45.4	45.9	45.6
80	27.4	26.3	26.9
≥80	27.2	27.9	27.5
<b>EQ-5D Index &lt;0.65</b>			
0.65-<0.80	20.9	18.6	19.9
0.65-<0.80	41.1	35.5	38.7
≥0.80	38.0	45.9	41.4
<b>SF-36 PCS &lt;30</b>			
30-<40	17.5	20.0	18.6
40	34.6	30.7	32.9
≥40	47.9	49.3	48.5
<b>SF-36 MCS &lt;30</b>			
30-<40	16.7	14.6	15.8
40	32.3	28.3	30.6
≥40	51.0	57.1	53.6

ANA = antinuclear antibodies; anti-dsDNA = anti-double-stranded DNA; BILAG = British Isles Lupus Assessment Group disease activity index; BlyS = B-lymphocyte stimulator; C3 = complement C3; C4 = complement C4; CRP = C-reactive protein; EQ-5D = EuroQOL 5 dimensions health status measure; FACIT = Functional Assessment of Chronic Illness Therapy; MCS = Mental Component Summary; PCS = Physical Component Summary; PGA = Physician Global Assessment; SELENA-SLEDAI = Safety of Estrogens in Lupus Erythematosus National Assessment-Systemic Lupus Erythematosus Disease Activity Index; SF-36 = Short Form-36 health survey; SLICC = Systemic Lupus International Collaborating Clinics damage index; VAS = visual analog scale.

Missing data on Anti-ribosomal P antibody (n = 16), BlyS protein level (n = 16), Anti-Smith antibody (n = 2), IgA (n = 4), IgG (n = 2), IgM (n = 2), and SLICC (n = 2). Totals may not equal 100% due to rounding.

**Supplementary Table 2** Evaluation of Biological and Social/Functional/Behavioral Variables at 52 Weeks vs SRI Response at 52 Weeks (n = 1175)

	Improved/Stable (Low)	Worsened/Stable (Moderate/High)	Kappa (95% Confidence Interval)
<b>Biological variables</b>			
CRP			
SRI responder	402 (59.5%)	274 (40.5%)	0.03 (−0.02-0.08)
Nonresponder	313 (62.7%)	186 (37.3%)	
C3			
SRI responder	452 (66.9%)	224 (33.1%)	−0.15 (−0.21 to −0.10)
Nonresponder	253 (50.7%)	246 (49.3%)	
C4			
SRI responder	450 (66.6%)	226 (33.4%)	−0.18 (−0.24 to −0.13)
Nonresponder	237 (47.5%)	262 (52.5%)	
Anti-dsDNA			
SRI responder	462 (68.3%)	214 (31.7%)	−0.09 (−0.14 to −0.03)
Nonresponder	295 (59.1%)	204 (40.9%)	
<b>Social/functional/behavioral variables</b>			
FACIT			
SRI responder	418 (61.8%)	258 (38.2%)	−0.02 (−0.08-0.03)
Nonresponder	297 (59.5%)	202 (40.5%)	
EQ-5D VAS			
SRI responder	490 (72.5%)	186 (27.5%)	−0.11 (−0.16 to −0.06)
Nonresponder	302 (60.5%)	197 (39.5%)	
EQ-5D index			
SRI responder	390 (57.7%)	286 (42.3%)	−0.03 (−0.08-0.03)
Nonresponder	274 (54.9%)	225 (45.1%)	
SF-36 PCS			
SRI responder	466 (68.9%)	210 (31.1%)	−0.04 (−0.09-0.01)
Nonresponder	322 (64.5%)	177 (35.5%)	
SF-36 MCS			
SRI responder	437 (64.6%)	239 (35.4%)	−0.06 (−0.12 to −0.01)
Nonresponder	289 (57.9%)	210 (42.1%)	

Anti-dsDNA = anti-double-stranded DNA; C3 = complement C3; C4 = complement C4; CRP = C-reactive protein; EQ-5D = EuroQOL 5 dimensions health status measure; FACIT = Functional Assessment of Chronic Illness Therapy; MCS = Mental Component Summary; PCS = Physical Component Summary; SF-36 = Short Form-36 health survey; SRI = Systemic lupus erythematosus Responder Index; VAS = visual analog scale.

**Supplementary Table 3** Evaluation of SRI Trajectory vs SRI at 52 Weeks (n = 1175)

SRI at 52 Weeks	SRI Trajectory		Kappa (95% Confidence Interval)
	Responder	Nonresponder	
SRI responder	499 (73.8%)	177 (26.2%)	0.63 (0.59-0.67)
Nonresponder	42 (8.4%)	457 (91.6%)	

SRI = Systemic lupus erythematosus Responder Index.



**Supplementary Table 4** Comparison of Biological Variables Compared with Social/Functional/Behavioral Variables at Week 52 and Over the Year Using the Trajectory Method

	FACIT	EQ-5D VAS	EQ-5D Index	SF-36 PCS	SF-36 MCS
Measured at 52 weeks					
kappa (95% confidence interval)					
C3	-0.01 (-0.07-0.04)	-0.02 (-0.07-0.04)	-0.09 (-0.14 to -0.03)	-0.07 (-0.12 to -0.01)	0.00 (-0.06-0.06)
C4	-0.02 (-0.08-0.04)	-0.04 (-0.09-0.02)	-0.09 (-0.15 to -0.03)	-0.09 (-0.14 to -0.03)	-0.02 (0.08-0.03)
CRP	0.05 (0.00-0.11)	0.08 (0.02-0.13)	0.11 (0.06-0.17)	0.13 (0.07-0.19)	0.03 (-0.03-0.08)
Anti-dsDNA	0.02 (-0.03-0.08)	0.01 (-0.05-0.06)	-0.02 (-0.08-0.04)	-0.02 (-0.08-0.04)	0.00 (-0.06-0.06)
Trajectory kappa (95% confidence interval)					
C3	-0.04 (-0.09-0.00)	-0.03 (-0.07-0.02)	-0.07 (-0.12-0.03)	0.00 (-0.04-0.04)	-0.06 (-0.10 to -0.02)
C4	-0.02 (-0.06-0.02)	-0.01 (-0.05-0.03)	-0.03 (-0.08-0.01)	0.01 (-0.03-0.05)	-0.01 (-0.06-0.03)
CRP	0.03 (-0.01-0.07)	0.03 (-0.01-0.07)	0.02 (-0.02-0.06)	0.03 (-0.01-0.07)	0.00 (-0.04-0.04)
Anti-dsDNA	0.01 (-0.03-0.06)	-0.01 (-0.06-0.04)	0.00 (-0.05-0.04)	-0.02 (-0.07-0.02)	-0.03 (-0.07-0.02)

Anti-dsDNA = anti-double-stranded DNA; C3 = complement C3; C4 = complement C4; CRP = C-reactive protein; EQ-5D = EuroQOL 5 dimensions health status measure; FACIT = Functional Assessment of Chronic Illness Therapy; MCS = Mental Component Summary; PCS = Physical Component Summary.