Received: 1 August 2022 | Revised: 12 October 2022 | Accepted: 27 October 2022

DOI: 10.1002/psp4.12888

ARTICLE

Disease trajectory of SLE clinical endpoints and covariates affecting disease severity and probability of response: Analysis of pooled patient-level placebo (Standard-of-Care) data to enable model-informed drug development

Kosalaram Goteti[1] | **Jonathan French**[2] | **Ramon Garcia**[2] | **Ying Li**[1] | **Florence Casset-Semanaz**[1] | **Aida Aydemir**[1] | **Robert Townsend**[1] | **Cristina Vazquez Mateo**[1] | **Matthew Studham**[1] | **Oliver Guenther**[3] | **Amy Kao**[1] | **Marc Gastonguay**[2] | **Pascal Girard**[4] | **Lisa Benincosa**[1] | **Karthik Venkatakrishnan**[1]

[1]EMD Serono Research and Development Institute, Inc (an affiliate of Merck KGaA, Darmstadt Germany), Billerica, Massachusetts, USA

[2]Metrum Research Group, Tariffville, Connecticut, USA

[3]Merck KGaA, Darmstadt, Germany

[4]Merck Institute of Pharmacometrics, Lausanne, Switzerland

Correspondence
Kosalaram Goteti, EMD Serono Research and Development Institute, Inc., 45A Middlesex Turnpike, Billerica, MA 01821, USA.
Email: kosalaram.goteti@emdserono.com

Abstract

Systemic lupus erythematosus (SLE) is an autoimmune disease affecting multiple organ systems. Many investigational agents have failed or shown only modest effects when added to standard of care (SoC) therapy in placebo-controlled trials, and only two therapies have been approved for SLE in the last 60 years. Clinical trial outcomes have shown discordance in drug effects between clinical endpoints. Herein, we characterized longitudinal disease activity in the SLE population and the sources of variability by developing a latent disease trajectory model for SLE component endpoints (Systemic Lupus Erythematosus Disease Activity Index [SLEDAI], Physician's Global Assessment [PGA], British Isles Lupus Assessment Group Index [BILAG]) and composite endpoints (Systemic Lupus Erythematosus Responder Index [SRI], BILAG-based Composite Lupus Assessment [BICLA], and Lupus Low Disease Activity State [LLDAS]) using patient-level historical SoC data from nine phase II and III studies. Across all endpoints, in predictions up to 52 weeks from the final disease trajectory model, the following baseline covariates were associated with a greater decrease in SLE disease activity and higher response to placebo + SoC: Hispanic ethnicity from Central/South America, absence of hypocomplementemia, recent SLE diagnosis, and high baseline disease activity score using SLEDAI and BILAG separately. No discernible differences were observed in the trajectory of response to placebo + SoC across different SoC medications (antimalarial and immunosuppressant such as mycophenolate, methotrexate, and azathioprine). Across all endpoints, disease trajectory showed no difference in Asian versus non-Asian patients, supporting Asia-inclusive global SLE drug development. These results

describe the first population approach to support a model-informed drug development framework in SLE.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Drug development in systemic lupus erythematosus (SLE) is challenged by low success rates, discordance in trial performance across primary endpoints, and the lack of reliable short-term outcome biomarkers of efficacy.

WHAT QUESTION DID THIS STUDY ADDRESS?

What are the clinical sources of variability in SLE disease trajectory that influence response to standard of care (SoC) treatment?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

This work describes the first population disease trajectory model of SLE. Patients from Central/South America with Hispanic ethnicity, baseline C3 not less than the lower limit of normal, or time since diagnosis <1 year, high baseline total Systemic Lupus Erythematosus Disease Activity Index (SLEDAI), or high baseline total number of British Isles Lupus Assessment Group Index (BILAG) organ systems A or B, have greater decrease in SLE disease activity during SoC treatment. Disease trajectory is similar between Asian and non-Asian patients.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

Patient enrollment strategies based on the identified covariates may enhance SLE proof-of-concept trial designs to ultimately maximize success rates. Consistency in disease trajectory in Asian versus non-Asian patients supports Asia-inclusive multiregional clinical trials, whereas the other identified covariates may inform appropriate stratification in pivotal SLE trials.

INTRODUCTION

Systemic lupus erythematosus (SLE) is an autoimmune disease that affects multiple organ systems, with fluctuating disease activity, including flares. Disease prevalence varies with race and ethnicity, and SLE is more common in women, with onset typically occurring during their childbearing years. Patients with moderate to severe SLE are chronically exposed to medications with significant side effects, such as corticosteroids and other immunosuppressive agents, and consider their health-related quality of life to be poor.

In clinical trials, SLE disease activity is assessed by several instruments like component or composite scores. Commonly used SLE instruments and component scores include:

- Systemic Lupus Erythematosus Disease Activity Index (SLEDAI), which measures presence/absence of 24 disease features, from which a weighted total score can be derived.[1]
- British Isles Lupus Assessment Group (BILAG) index, which evaluates eight individual organ systems;[1] the organ systems included in this analysis included: musculoskeletal, mucocutaneous, or renal.
- Physician's Global Assessment (PGA) on a 0 (none) to 3 (severe) visual analogue scale.
- Average prednisone-equivalent daily corticosteroid dose (aPEDD).

Composite SLE outcomes are based on component scores like the SLEDAI and/or BILAG, PGA, and non-initiation of protocol-prohibited treatments, and include:

- Systemic Lupus Erythematosus Responder Index (SRI).[2]
- BILAG-based Composite Lupus Assessment (BICLA).[3]
- Lupus Low Disease Activity State (LLDAS).[4]

Many therapeutic interventions have failed or shown only modest effect compared to standard of care (placebo + SoC) in clinical trials.[5,6] Furthermore, trial outcomes have been discordant between endpoints,[7,8] which may, in part, be related to the complexity of the composite endpoints. In addition, it has been shown that shortfalls of SRI and BICLA may be due to BICLA requiring only partial improvement but in all organs versus SRI requiring full

improvement in some manifestations and not necessarily in all organs.[9] Thus, the importance of understanding endpoint sensitivity and eventual clinical phenotypes[10,11] associated with it to describe the SLE disease are required and beyond the scope of this paper. Of note, there are efforts within the SLE community for newer or modified endpoints that describe the SLE disease and perhaps these will be available in the future.[12]

Therefore, there is a need to understand these composite endpoints via longitudinal models of historical placebo + SoC data. Of note, the importance of considering the time course of SLE endpoints in longitudinal analyses for evaluating sources of variability in treatment response have been emphasized previously.[13–15] In a recently convened workshop by the US Food and Drug Administration (FDA) the value of longitudinal disease progression modeling for enabling efficient drug development and informing patient selection and cross-population extrapolation in chronic diseases was extensively discussed.[16] Herein, we describe development of disease trajectory models for SLE endpoints to provide insight into the factors influencing SLE response following placebo + SoC and help identify study populations who are likely to have a low response to placebo + SoC and who could benefit from new treatment.

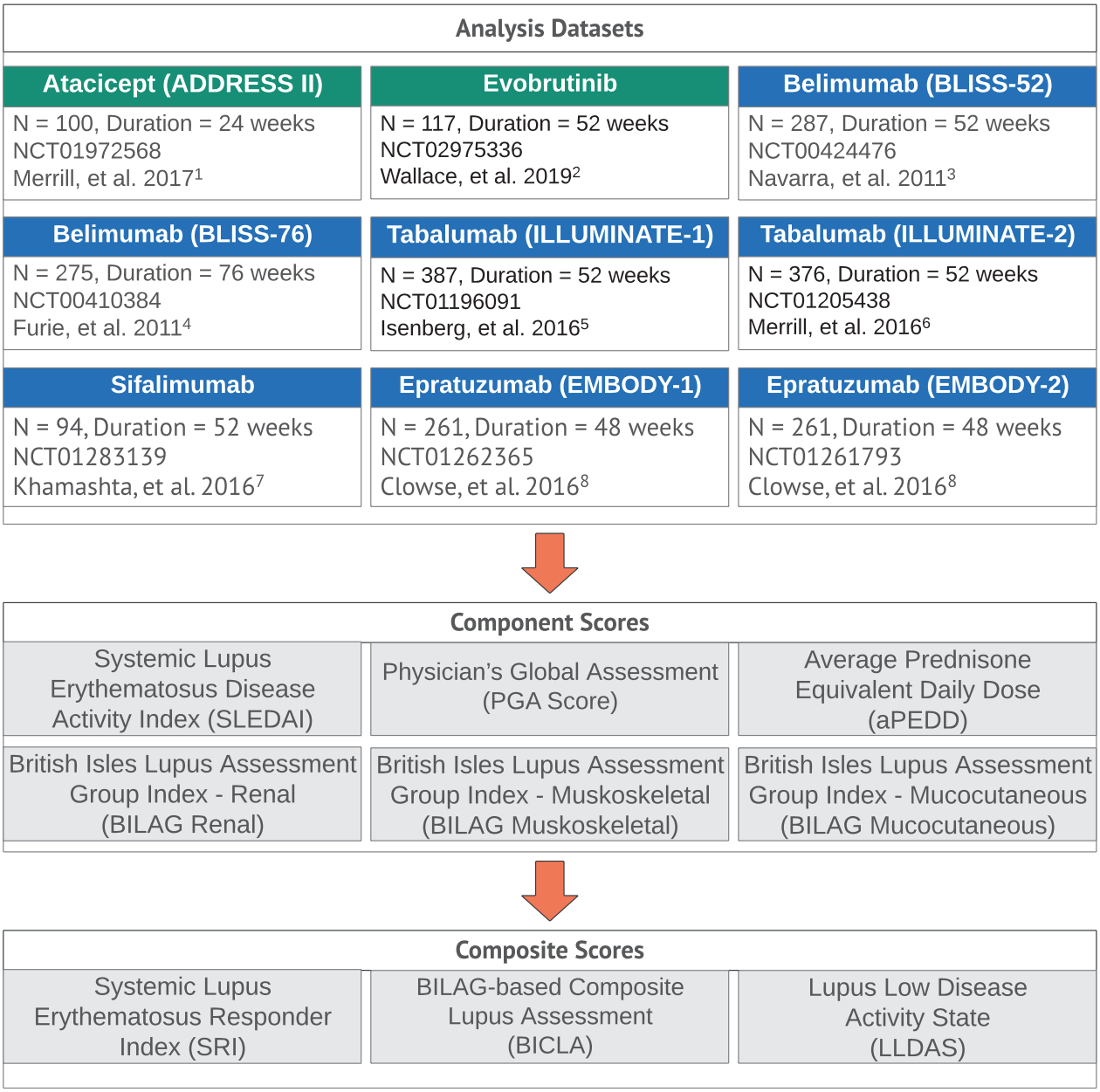
The objective of this analysis was to quantitatively characterize the time course of SLE disease trajectory and identify associated demographic and clinical sources of variability in patients with SLE in the placebo + SoC arms of randomized controlled trials.

METHODS

Data came from nine randomized, placebo controlled clinical trials in patients with SLE who received the SoC treatment (placebo + SoC) of each trial. Two trials were conducted by the sponsor EMD Serono Inc. (a business of Merck KGaA), and placebo + SoC data from seven trials were obtained through the TransCelerate BioPharma's Historical Trial Data Sharing Initiative.[17] Figure 1 summarizes the studies, including their respective clinicaltrials.gov identifiers, and primary publications. The individual components for the composite endpoints were collected on the same patient at the same visit over a specified study duration described in the protocol as per the inclusion and exclusion criteria of the respective trials. Details of study design and duration are as noted in Figure 1 and section 7 in Supporting information and the associated references. Data integration, visualization, modeling, and simulations were conducted using R version 3.6,[18] and estimation and inference were carried out using Bayesian methods implemented in Stan version 2.4[19] via RStan package version 2.21.2.[20] Because the number of compounds evaluated are minimal, only an empirical method using a latent variable model similar to the ones in literature was used[21,22] and this analysis does not tease out the mechanism of action for each compound as only placebo + SoC data were modeled and patient-level data from treatment arms were not available from the TransCelerate BioPharma's Historical Trial Data Sharing Initiative. It should be noted that latent variable framework described in the paper reflects the unobserved disease trajectory. Using such an approach, the main purpose was to distinguish between the time course of endpoints from disease trajectory. The baseline continuous covariates by study and population are provided in Supporting information (Section 7).

Model development

Joint (multiple endpoint) longitudinal models were developed separately for the component and composite SLE outcomes. As shown in Figure 1, the component outcomes consisted of six endpoints: SLEDAI total score; scores for the BILAG musculoskeletal, mucocutaneous, and renal organ system domains; PGA; and aPEDD. Composite SLE outcomes consisted of three endpoints: SRI, BICLA, and LLDAS. Although some scales may present collinearity, the assumption in these analyses is that the outcomes from different scales are independent conditional in the latent SLE disease. Having an underlying latent SLE disease induces correlation on the same outcome of a patient and among different scale outcomes of the same patient. Consequently, the use of latent variable modeling to express outcomes as a function of latent disease shows an approach that can be used to model various component or composite psychometric instruments simultaneously while taking into account correlation between them. Both component and composite SLE outcome models assumed that there was a unidimensional and structurally similar latent continuous measure of SLE disease activity that varies as a continuous function of time. Each longitudinal outcome measure was modeled as a function of the latent disease activity. The component and composite outcomes were analyzed separately but used the same structural form for the latent disease trajectory in each analysis. The adherence to corticosteroid and dropout were modeled in the latent disease trajectory framework of the component and composite endpoints (discussed further in Section 4.3 in Supporting information). This was important to be included in the modeling plan to mitigate the impact of potential non-random dropout, especially in longer studies. The structural models for latent disease activity, observed component SLE outcomes, and observed composite SLE outcomes are described in the Supporting information (Sections 1, 2, and 3, respectively).



[1] Merrill JT, et al. Efficacy and safety of atacicept in patients with systemic lupus erythematosus. Arthritis Rheumatol. 2018;70(2):266-276.
[2] Wallace DJ, et al. 212 Phase 2, randomized, double-blind, placebo-controlled, dose-finding study, evaluating the Bruton's tyrosine kinase inhibitor evobrutinib in patients with systemic lupus erythematosus: study design. Lupus Science & Medicine. 2019;6(Suppl 1).
[3] Navarra SV, et al. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. Lancet. 2011;377(9767):721-731.
[4] Furie R, et al. A phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. Arthritis Rheum. 2011;63(12):3918-3930.
[5] Isenberg DA, Petri M, Kalunian K, et al. Efficacy and safety of subcutaneous tabalumab in patients with systemic lupus erythematosus: results from ILLUMINATE-1, a 52-week, phase III, multicentre, randomised, double-blind, placebo-controlled study. Ann Rheum Dis. 2016;75(2):323-331.
[6] Merrill JT, et al. Efficacy and safety of subcutaneous tabalumab, a monoclonal antibody to B-cell activating factor, in patients with systemic lupus erythematosus: results from ILLUMINATE-2, a 52-week, phase III, multicentre, randomised, double-blind, placebo-controlled study. Ann Rheum Dis. 2016;75(2):332-340.
[7] Khamashta M, Merrill JT, Werth VP, et al. Sifalimumab, an anti-interferon-α monoclonal antibody, in moderate to severe systemic lupus erythematosus: a randomised, double-blind, placebo-controlled study. Ann Rheum Dis. 2016;75(11):1909-1916.
[8] Clowse MEB, et al. Efficacy and safety of epratuzumab in moderately to severely active systemic lupus erythematosus. Arthritis Rheumatol. 2017;69(2):362-375.

FIGURE 1 Summary of datasets and analysis workflow

Identification of covariates

To identify covariates, machine learning approaches, including random forest and least absolute shrinkage and selection operator (LASSO) models, were applied to the empirical Bayes estimates in the latent disease activity time course. Multiple machine learning methods were used so that the choice of one particular analytical technique does not inadvertently pick up a covariate that is not likely to explain the disease trajectory in SLE. Relative covariate importance was assessed from the random forest model using the Boruta algorithm and Shapley values.[23–25] With the latent disease model, it is not possible to directly assess clinical relevance of covariate effects; hence, covariates were assessed for clinical relevance through simulations.

Selection of covariates was dependent on which baseline data were consistently available across all studies. Baseline covariates considered for identification as predictive in the latent disease model were:

- Demographic: age (years), body weight (kg), sex (female vs. male), race (White vs. Asian vs. other vs. missing), and ethnicity (non-Hispanic vs. Hispanic – North America vs. Hispanic – Central/South America vs. missing).
- Laboratory: complement C3 below the lower limit of normal (no vs. yes), complement C4 below the lower limit of normal (no vs. yes), anti-double stranded DNA positive status (no vs. yes), baseline renal function (estimated glomerular filtration rate [eGFR]), renal involvement (normal: eGFR >90 ml/min/1.73m^2 vs. mild: eGFR 60–90 ml/min/1.73m^2, and moderate: eGFR <60 mL/min/1.73m^2).
- Disease status: time since diagnosis (≤1 year vs. >1 year), baseline SLEDAI total score, baseline total number of organ systems with scores of A or B (0 vs. 1, 2, and 3+).
- Previous or concomitant medication (yes vs. no): antimalarial use, immunosuppressant use, mycophenolate use, methotrexate use, and azathioprine use. Baseline prednisone equivalent daily corticosteroid dose (mg).

Although the baseline SLEDAI total score was included in the structural model for the SRI outcome for the composite analysis, it was included in the covariate selection to assess its ability to predict the LLDAS and BICLA composite outcomes. Missing covariates were either imputed using the median or the mode for continuous and categorical covariates, respectively. If a study did not provide any data for a particular covariate, missing data were imputed at the overall median or mode across all studies. For categorical variables with substantial missing data, a "missing" category was defined.

Model evaluation and validation

Modeling was performed on a subset of the full data using a train/evaluate/test paradigm,[26] which consisted of 60%, 20%, and 20% of the total number of patients, respectively. Hereafter, these data are referred to as training, evaluation, and test datasets. Although the approach of using subsets of the full data for model development has its limitations, it should be noted that appropriate stratification and cross-validation like leave-one-study-out cross-validation procedures were used to ensure robustness of the analyses. The partitioning into training, evaluation, and test datasets was stratified by sex, total number of organ systems with BILAG index scores of A or B at baseline, and baseline anti-double stranded DNA positive status. Model development was performed using the training dataset, and model evaluation was performed using the evaluation dataset. The test dataset was used at the end of the modeling process as a one-time external model evaluation. Additionally, the models were evaluated using a leave-one-study-out cross-validation to assess the model's ability to predict future placebo + SoC response in a study up to 52 weeks. The primary method for evaluation was visual predictive checks (VPCs), the widely applicable information criterion (WAIC), and log posterior density (LPD) criterion, whereby smaller values for the WAIC and LPD criterions reflect a more favorable model.[27,28] Models were initially estimated using the training dataset with no covariates. The WAIC and LPD were computed using the training and evaluation datasets, respectively, and compared across models. Somers' D criterion was used to quantify the predictive performance in the leave-one-study-out cross-validations.[29,30] For continuous variables, Somers' D corresponds to a measure of rank correlation; for binary variables, it is a linear transformation of the area under the concentration time curve (AUC) under the receiver operating characteristic curve (D = 2 AUC-1). The final model was estimated using the entirety of the data.

RESULTS

Source data

Data were pooled from 2158 patients across nine phase II and III studies that ranged from 24 to 76 weeks in duration and from 94 to 387 patients in size (Figure 1). The data were partitioned randomly stratified across trials into training, evaluation, and test datasets with sample sizes of 1309, 432, and 417 patients, respectively. Patient covariates are summarized across all studies in Table 1. The median baseline SLEDAI score was 10.0 (range: 0 to 34), and the median number of BILAG organ systems

TABLE 1 Patient demographics and baseline characteristics

Statistic	All data $N = 2158$
Baseline age, years	
Mean (SD)	39.4 (12.2)
Minimum/maximum	17.5/80.0
Missing	522
Baseline body weight, kg	
Mean (SD)	69.6 (18.0)
Minimum/maximum	34.7/177
Missing	26
Time since diagnosis (at screening), years	
Mean (SD)	7.93 (7.52)
Minimum/maximum	0.00/57.0
Missing	772
Baseline SLEDAI total score	
Mean (SD)	10.2 (3.74)
Minimum/maximum	0.00/34.0
Missing	12
N organ systems w scores of A or B at BL	
Mean (SD)	1.76 (0.848)
Minimum/maximum	0.00/7.00
Missing	12
BL prednisone-equivalent daily corticosteroid, mg	
Mean (SD)	13.0 (32.4)
Minimum/maximum	0.0500/1250
Missing	566
Baseline renal function, ml/min/1.73 m^2	
Mean (SD)	99.8 (27.7)
Minimum/maximum	23.6/291
Missing	542
Sex	
Female	2017 (93.5)
Male	141 (6.5)
Race	
Asian	280 (13.0)
White	919 (42.6)
Other	429 (19.9)
Missing	530 (24.6)
Hispanic vs. Non-Hispanic	
Non-Hispanic	1008 (46.7)
Hispanic – Central/South America	449 (20.8)
Hispanic – North America	109 (5.1)
Missing	592 (27.4)
eGFR categorization	
Normal	1020 (47.3)
Mild	503 (23.3)
Moderate	93 (4.3)
Missing	542 (25.1)
BL complement C3 below lower limit of normal	
No	1308 (60.6)
Yes	815 (37.8)
Missing	35 (1.6)
BL complement C4 below lower limit of normal	
No	1376 (63.8)
Yes	718 (33.3)
Missing	64 (3.0)
Time since missing diagnosis	
Less than a year	163 (7.6)
Greater than a year	1223 (56.7)
Missing	772 (35.8)

Note: Categorical summary is count (percent).

Abbreviations: BL, baseline; eGFR, estimated glomerular filtration rate; *N*, number of records summarized; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index.

with scores of A or B was 2.0 (range: 0 to 7). The percentage of missing baseline covariate data was imbalanced across studies with certain studies providing no information on some of the covariates due to data anonymization or other reasons from the sponsor companies. The ILLUMINATE-1 and ILLUMINATE-2 studies had completely missing time since diagnosis data (35.3%); the EMBODY1 and EMBODY2 studies had completely missing eGFR, ethnicity, and race data (24.2% missing for these 3 variables), and the Sifalimumab study had missing data on prior and concomitant medication use (4.36%).

Component score SLE modeling

Latent variable model development

The WAIC and LPD were computed and compared across models (Table S1), showing that the log linear model was preferable; however, the VPCs indicated that the mono-exponential models provided a better fit to the observed data (Figure 2). The log linear model also did not provide monotone decreasing predictions of some endpoints. The WAIC and LPD were similar for the maximum effect and mono-exponential models, but, ultimately, the mono-exponential model values were lower (indicating a more favorable model). Of the two mono-exponential models,

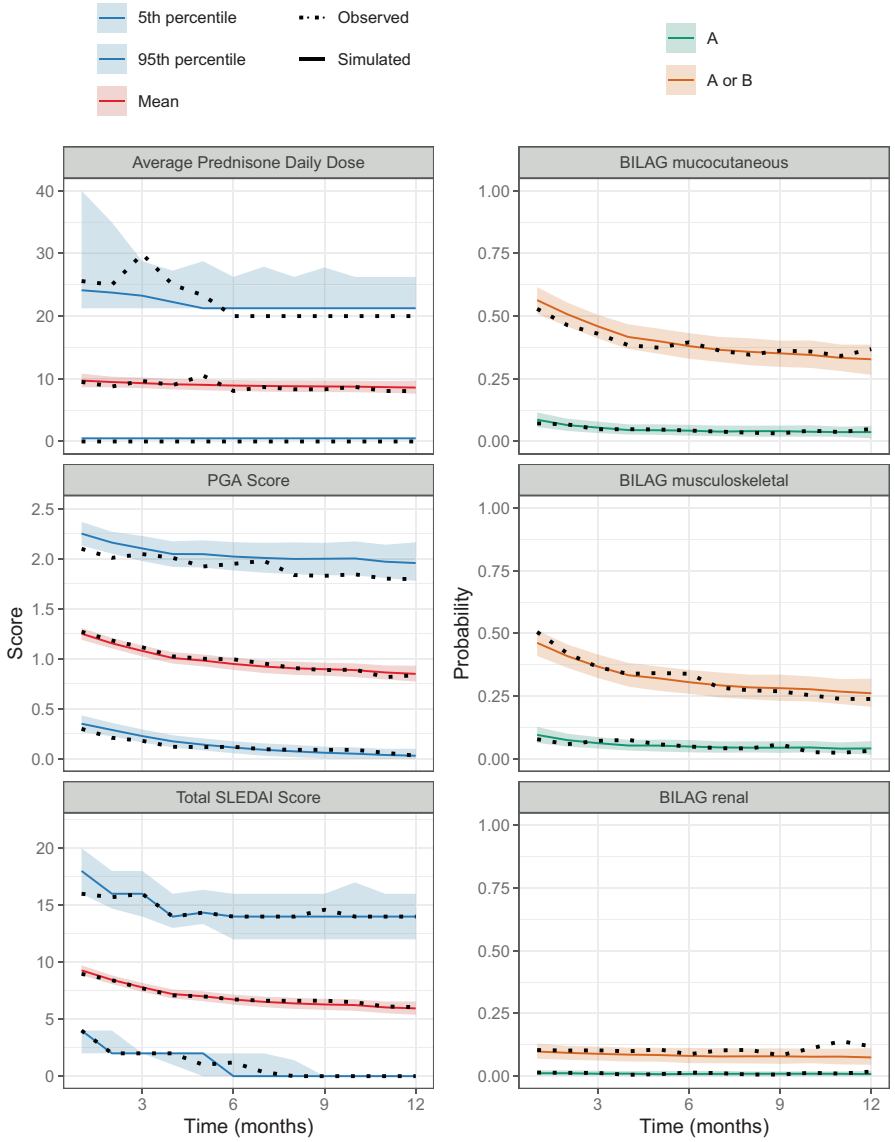


FIGURE 2 Visual predictive check of final model for component score analysis. BILAG, British Isles Lupus Assessment Group Index; PGA, Physician's Global Assessment; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index

the one with random effects on baseline disease activity and long-term disease activity (defined as ≥1 year) was most parsimonious and was selected as the final latent variable model.

Identification of covariates

The covariates identified for baseline disease activity and long-term disease activity differed depending on the machine learning method or rule implemented (Table S2). All covariates identified were included in the final model. To ensure missing information about previous and concomitant medication use in the Sifalimumab study was not biasing the results, covariate selection was performed with and without patients from the Sifalimumab study, and no additional covariates were identified.

Final model

The final latent SLE disease trajectory for the component score outcomes were best described using mono-exponential models with random effects on baseline and long-term disease activity and covariate effects (see Table S5 and Section 6 in Supporting information for the formulaic expressions). The covariates for baseline disease activity included ethnicity, the number of organ systems with BILAG index scores of A or B, age, and total SLEDAI score (Table S2). The covariates for long-term disease activity included race, ethnicity, complement C4 below the lower limit of normal (LLN); the number of organ systems with BILAG index scores of A or B, age, body weight, time since diagnosis, aPEDD, total SLEDAI score, and renal function. The Somers' D criterion indicated the model did an adequate job of predicting disease trajectory in each study (Figure S1).

After identifying the final model structure and covariates, the training and evaluation datasets were combined, and the model was re-fit. As an external validation of model fit, this second iteration of the model was used to predict into the test dataset. VPCs were generated to assess model performance by comparing each model-predicted outcome to those observed in the test dataset and did not suggest model deficiencies across component endpoints or studies (Figure 2). Stratifying the results by different baseline covariates also did not suggest model deficiencies across component endpoints or studies (not shown). Finally, the training, evaluation, and test datasets were recombined, and the model was re-fit using the full dataset. These final parameter estimates were used to assess magnitude and directionality of the covariates (Table S3) and to simulate the change from baseline in clinical endpoints, over 1 year of treatment, for covariates (Figures 3, 4, and 5).

Parameter estimates of covariate effects can be found in the Supporting information. The model-predicted change in SLE component endpoints, for a given change in baseline covariate, differed in size and directionality across endpoints (Table S3). Higher baseline total SLEDAI score was associated with larger reductions in total SLEDAI, PGA, BILAG mucocutaneous and BILAG musculoskeletal endpoints at all study timepoints; the association between baseline total SLEDAI score and aPEDD and BILAG renal endpoints was weaker. Across all endpoints, predictions up to 52 weeks generally showed that patients from Central/South America with Hispanic ethnicity or baseline C4 not below the LLN or time since diagnosis of less than a year or high baseline total SLEDAI or high baseline total number of BILAG organ systems A or B had greater decrease in SLE disease activity compared with other patients.

Predictions of outcomes indicated that Hispanic ethnicity from Central/South America, baseline total number

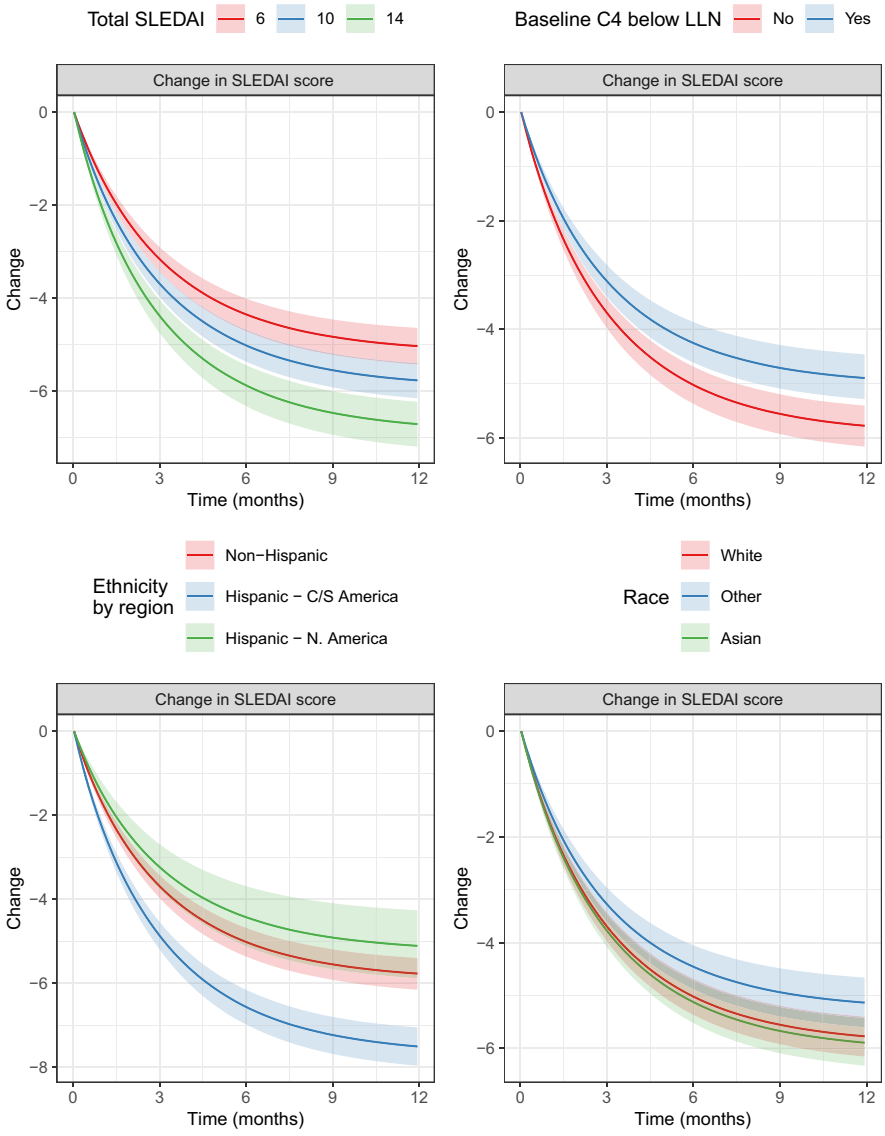


FIGURE 3 Impact of baseline total SLEDAI, C4 below LLN, ethnicity by region or race on simulated change in SLEDAI score. C/S America, Central/South America; LLN, lower limit of normal; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index

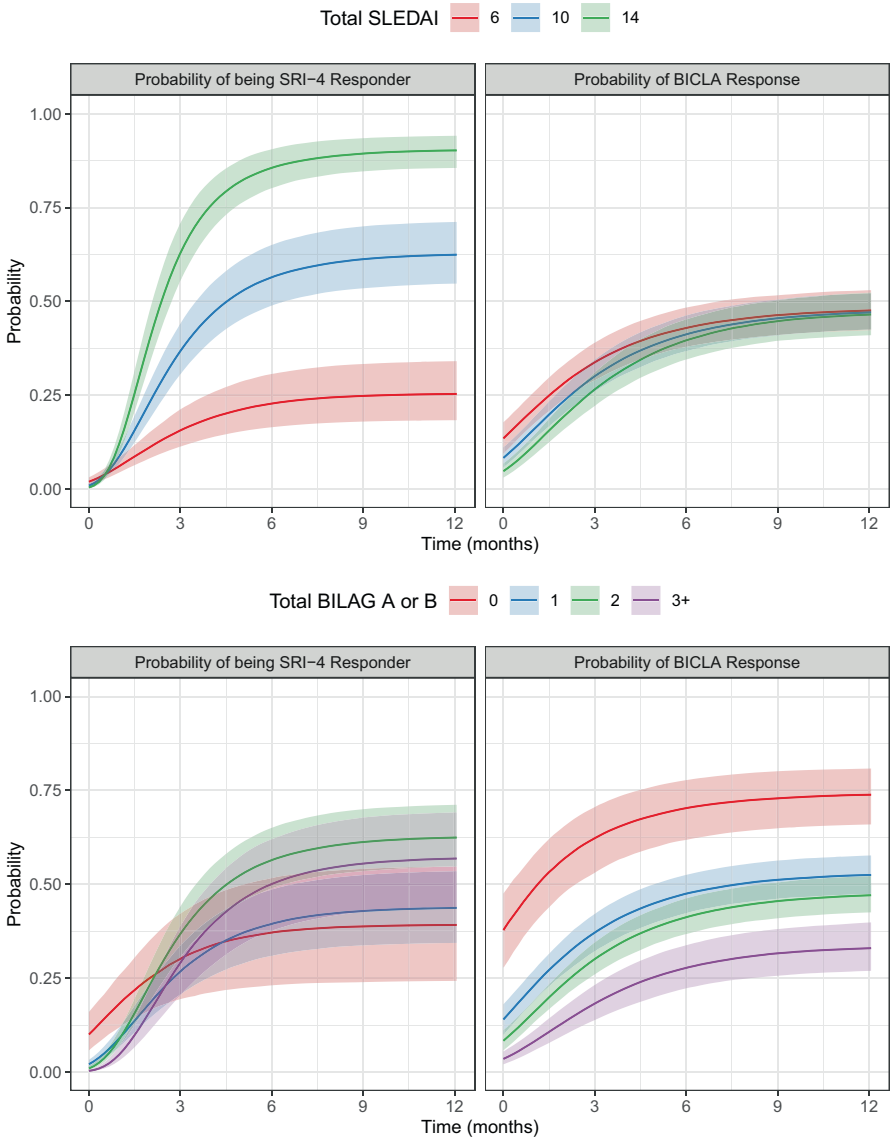


FIGURE 4 Impact of baseline total SLEDAI or BILAG scores on simulated probabilities of being an SRI-4 or BICLA responder. BICLA, British Isles Lupus Assessment Group Index-based Composite Lupus Assessment; BILAG, British Isles Lupus Assessment Group Index; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index

of BILAG scores A and B, time since diagnosis, C4 below LLN, and baseline total SLEDAI had an impact on model-predicted future SLE disease activity (Figures 3, 4, and 5).

Composite score SLE modeling

Latent variable model development

The mono-exponential model that best fit the component outcomes was used to fit the composite outcomes, although variations on the latent variable structure were explored (Table S1). The models were initially estimated using the training dataset only and no covariate predictors. The WAIC and LPD indicated models with and without endpoint specific interindividual variability (IIV) fit the training and evaluation datasets better than the time-scaled and shifted disease trajectory model. The WAIC was identical for models with and without endpoint specific IIV, indicating that they both fit the training data equally well. However, the model with endpoint specific IIV better fit the evaluation dataset (per the LPD), and it was selected as the final latent variable model for the composite score SLE modeling.

Identification of covariates

The identified covariates again differed depending on the machine learning method or rule implemented (Table S2). All covariates identified were included in the final model. Similar to the component score modeling, covariate selection was performed with and without patients from the Sifalimumab study, and no additional covariates were identified when Sifalimumab study patients were removed.

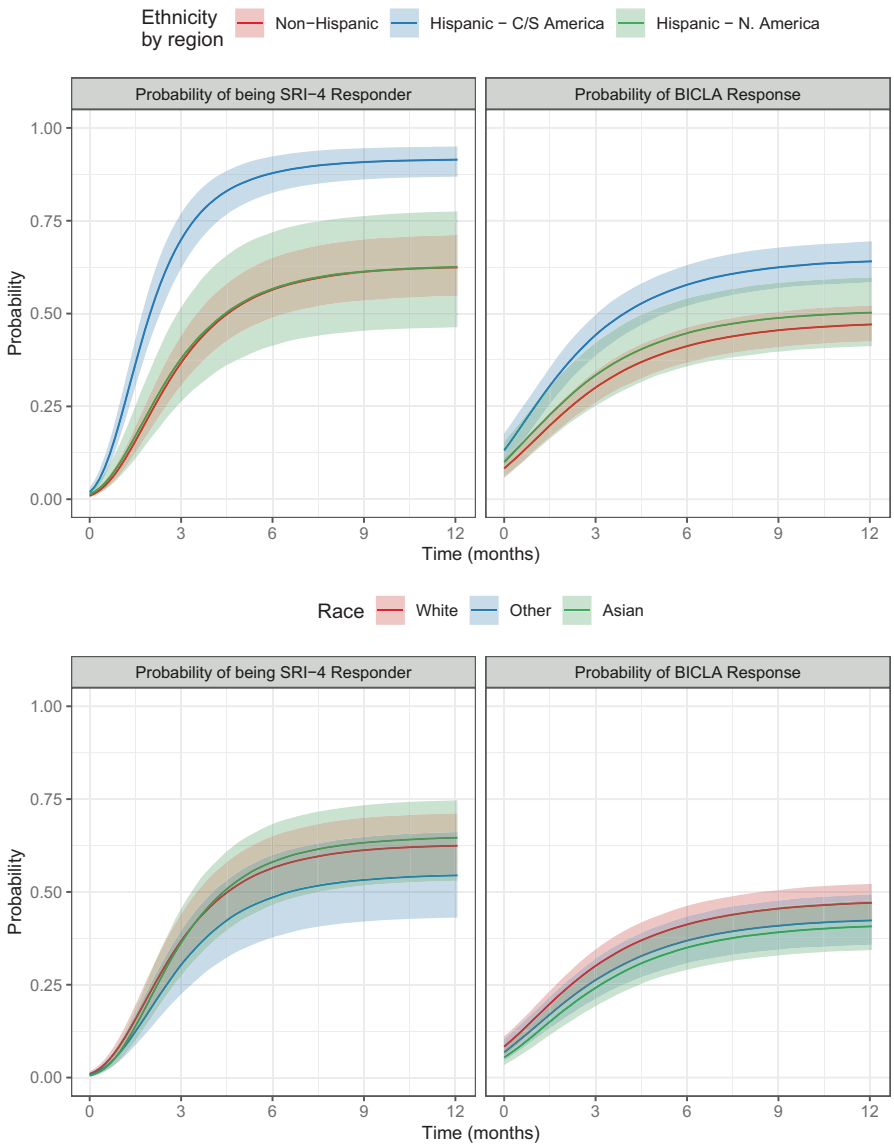


FIGURE 5 Impact of ethnicity by region or race on simulated probabilities of being an SRI-4 or BICLA responder. BICLA, British Isles Lupus Assessment Group Index-based Composite Lupus Assessment

Final model

The final latent SLE disease trajectory for the composite score outcomes were best described using a mono-exponential model with endpoint specific IIV, random effects on baseline and long-term disease activity, and covariate effects (see Table S5 for covariate effects and Section 6 in Supporting information for the formulaic expressions). The covariates for the baseline disease activity included race, ethnicity by region, total number of organ systems with BILAG scores of A or B, body weight, time since diagnosis, aPEDD, and total SLEDAI (Table S2). The covariates for long-term disease activity included race, ethnicity, complement C3 below the LLN, total number of organ systems with BILAG scores of A or B, age, body weight, aPEDD, total SLEDAI, and renal function. The Somers' D criterion indicated the model did an adequate job of predicting disease trajectory in all studies (Figure S1). The extent to which the IIV is explained by the covariates differs between the two models. In the component model, covariates explain 49% and 16% of the variability in baseline and long-term disease activity, respectively. In the composite endpoint model, covariates explain 20% and 58% of the corresponding variability.

As with the component score modeling, the training and evaluation datasets were combined, and the final model was re-fit. VPCs did not suggest model deficiencies across composite endpoints or studies (Figure 6). Results were comparable when stratifying by different baseline covariates and for patients with and without previous and concomitant medication use (i.e., no model deficiencies were identified across component endpoints or studies; Figure S2). Finally, the model was re-fit using the full, recombined dataset. These final parameter estimates were used to assess magnitude and directionality of the covariates (Table S4) and to simulate the change from baseline in clinical endpoints, over 1 year

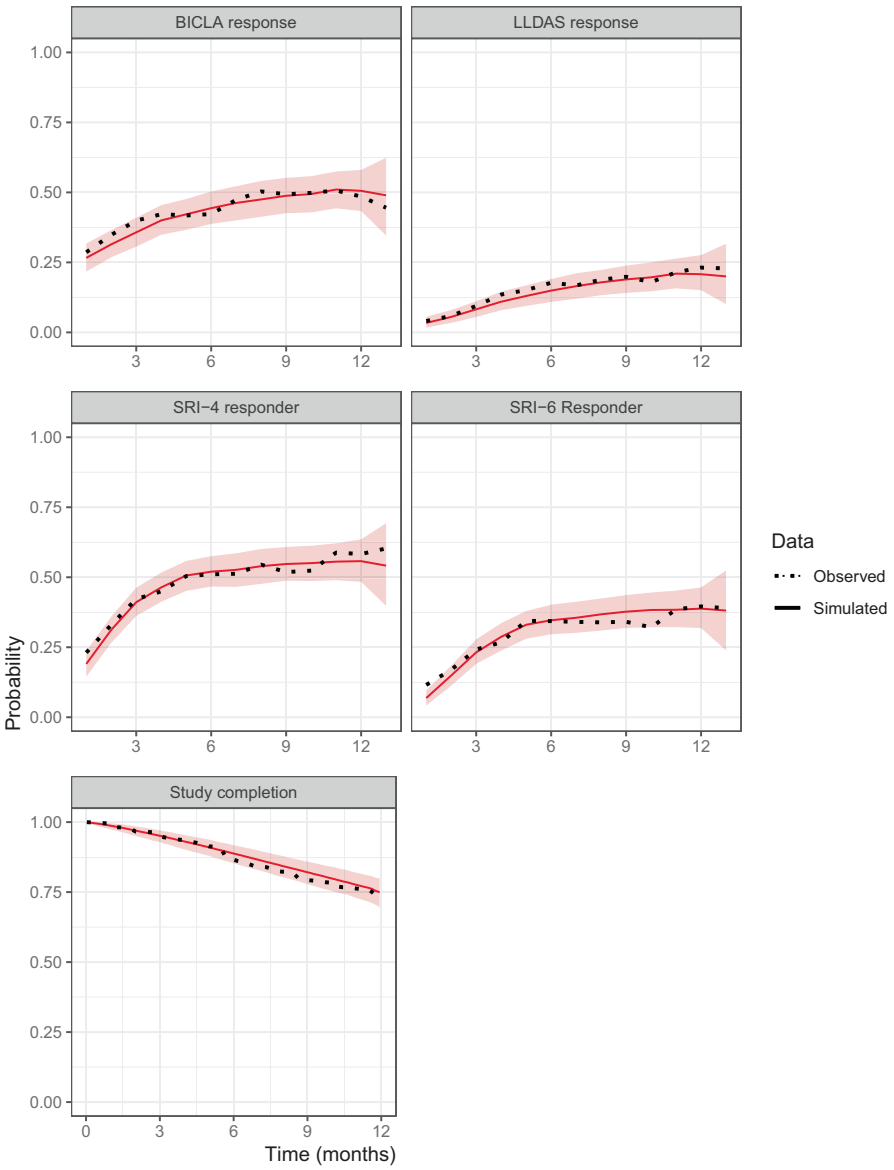


FIGURE 6 Visual predictive check of final model for composite score analysis. BICLA, British Isles Lupus Assessment Group Index-based Composite Lupus Assessment; LLDAS, Lupus Low Disease Activity State

of treatment, for all covariates (Figures 3, 4, and 5). Model performance stratified by covariates of interest in global drug development (race and ethnicity) are shown in the Supporting information (Section 8: Figure S3 and Figure S4). VPCs stratified by other covariates also yielded similar plots (data not shown). Further, sensitivity analyses of the missing data were performed and did not result in any model instability as predicted in the test and validation data sets using the leave-one-study-out cross-validation technique.

The model predicted change in SLE composite endpoints for a given change in baseline covariate differed in size and directionality across endpoints (Table S4). Higher baseline total SLEDAI score was associated with higher proportion of responders for SRI-4 and SRI-6 response (where SRI-4 and -6 are defined as a 4- and 6-point reduction in SLEDAI score); the association between baseline total SLEDAI score and BICLA and LLDAS endpoints was weaker over time. Higher baseline total number of BILAG organ systems A or B was associated with larger increases in BICLA response and LLDAS attainment but was associated with decreases with the SRI-4 and SRI-6 response. Across all the endpoints, predictions up to 52 weeks generally showed that patients from Central/South America with Hispanic ethnicity or C3 not below the LLN or time since diagnosis of less than a year or high baseline total SLEDAI or high baseline total number of BILAG organ systems A or B had greater decrease in SLE disease activity compared with other patients.

Predictions of outcomes indicated that ethnicity, baseline total number of BILAG scores A and B, time since diagnosis, C3 below the LLN, and baseline total SLEDAI had an impact on model-predicted future SLE disease activity (Figures 4 and 5).

DISCUSSION

SLE has high unmet medical needs along with complex composite endpoints making the readout of the trials cumbersome, therefore there is a need for model-informed drug development. Population models were developed to describe the time course of the component and composite endpoints used in phase II and III SLE clinical trials in a latent disease model framework. In the current analysis, placebo effect (without SoC treatment) on the time course of clinical endpoints was not evaluated because no data were available on patients receiving only the placebo without SoC. Consequently, no information was available on pure placebo effect. It should be noted that whereas endpoints may overlap in characterizing the endpoint time course, both the component and composite scores were treated as un-nested in this analysis. Furthermore, individual component or composite scores (e.g., BICLA and SRI) were treated as uncorrelated, conditional on the latent disease severity. Covariate analyses were enabled by various machine learning approaches with rules for identifying covariates using multiple measures of variable importance. These methods and rules produced different results, with LASSO typically identifying the least number of covariates and the Shapley importance measure identifying the most covariates, although the identified covariates between the component and composite score analyses were similar (Table S2). The covariates identified by the Shapley variable importance measure were consistently important, but this was less so with the other variable importance measures. Therefore, it was necessary to find the most common covariates that influence the SLE disease trajectory by running multiple algorithms.

Multiple rules were used to ensure covariate identification properly. Covariates identified were not necessarily strongly associated with latent disease trajectory or SLE outcomes. For this reason, simulations using the final estimated model were primarily used to determine whether a covariate was an influential predictor of the SLE outcomes. For example, age was identified as covariate on baseline and long-term disease activity, but predictions of change in SLEDAI score were similar across ages in the patient population. This may be due to the large sample size, which can detect small covariate effects that may not necessarily be strongly associated with an outcome.

The component and composite outcomes were analyzed separately due to the distinct nature of the outcomes. The limitations of composite and general disease status scores are not restricted to SLE, but a feature of most clinical scales with comparable operating characteristics. Previous attempts to explore the sensitivity of clinical scales for the evaluation of antidepressant drugs have been challenging.[31,32] Often, scale items or dimensions are too rudimentary or insensitive to the specific pharmacological effects of drugs, which in many cases do not provide disease modifying properties, but rather symptomatic improvement. Additionally, while understanding the trajectory of endpoints is important, there are other challenges in establishing efficacy in clinical trials, such as the heterogeneity of disease pathogenesis, patient population, and SoC therapy. The challenges associated with evaluation of Belimumab have been well-documented due to the nature of SLE endpoints[33] and certainly a lot of work is actively ongoing in this space to derive modified or new disease endpoints that perhaps can better describe SLE as a disease. We envision that continued refinement and extension of the model may be needed for any newer endpoints that would emerge in this disease area in the upcoming years that could support shorter trials or the possibility to have a lower sample size if a population with lower placebo response is identified. The analyses were linked by using the same mono-exponential latent disease structural model developed in the component analysis for the composite outcome analysis. Similar covariates were identified as predictive, and influential in simulations, for both component and composite analyses, except that C4 below the LLN was predictive for the component analysis, whereas C3 below the LLN was predictive for the composite analysis. For the composite score outcomes, additional parameters, such as endpoint specific IIV were necessary to adequately explain the data. The impacts of influential predictors on the composite scores were different for the various outcomes, as reflected in both the simulations and data. For example, increases in the baseline total SLEDAI had a minimal impact on the BICLA and LLDAS endpoints over time, implying a minimal change in future SLE disease activity. In contrast, increases in the baseline total SLEDAI increased the SRI-4 and SRI-6 endpoints, implying a decrease in future SLE disease activity (Figure 4 and Table S4). These findings are consistent with previous results that showed discordant trial outcomes between SRI and BICLA (e.g., as noted in the two phase III trials of anifrolumab).[7,8] In future SLE clinical trials, patient stratification based on baseline SLEDAI or baseline BILAG organ A or B scores may warrant consideration for the primary endpoint of interest (SRI vs. BICLA).

The VPCs showed the model sufficiently explained the data for the overall placebo + SoC patient population and under different covariate classifications. The leave-one-study-out cross-validations demonstrated that the model adequately explained outcomes across studies (Figure S1). Among component endpoints, aPEDD and BILAG renal outcomes showed poorer predictive performance than the other component endpoints, which may be due to the flat trajectories of these outcomes (i.e., the model is unable

to predict change over time when aPEDD and BILAG remain fairly constant). Hence, these may be less meaningful in finding differences in SLE clinical trials.

Some studies were completely missing data due to anonymization or other reasons. The ILLUMINATE-1 and ILLUMINATE-2 studies had completely missing time since diagnosis data, and the EMBODY1 and EMBODY2 studies had completely missing ethnicity, race, and eGFR data. If the missing data of these covariates is acquired, a re-analysis may result in different results and conclusions. When substantial data were missing, such as with race and ethnicity, a "missing" category was included, and its effect estimated. Because these missing data were mostly attributable to the EMBODY studies, this covariate effect is analogous to a study effect rather than actual ethnicity or race effect. At the time of analysis, the selection of trials contributing to the individual-level dataset was based on what was accessible to the authors from the stated external source (TransCelerate BioPharma's Historical Trial Data Sharing Initiative). Importantly, to address the generalizability of the model, we have conducted leave-one-study-out cross-validation analyses and presented the results in the Supporting information along with Somer's D criteria to assess model performance across datasets. As more data would be available in the future, perhaps validation with external data can be performed with this current model.

Patient medication use was included in the covariate selection of the latent variable model. Medication use was distinguished by type (anti-malarial, immunosuppressant, mycophenolate, methotrexate, and azathioprine) and timing (either prior to study enrollment or concomitantly). No medications were identified as predictive of latent SLE disease during covariate selection, implying that SLE disease trajectory is consistent across current SoC regimens. This was supported by the fact that the model-based predictions were equally good for patients with and without previous and concomitant medication use (Figure S2). Although SoC drugs with distinct mechanisms of action were assessed in our disease trajectory model, it was not possible to discern mechanism- or class-related differences in their effects on the underlying disease dynamics in SLE based on this analysis. Application of the developed modeling framework is anticipated to be useful for future analyses of the effects of investigational agents with novel mechanisms of action. The treatment effect can be incorporated in the future within the latent disease framework of an investigational agent, where the new therapeutic may affect either the baseline latent disease activity and/or the long-term latent disease activity. In addition, the model can be used to support Bayesian analyses of future clinical trials through the use of a weakly or moderately informative prior distribution for a placebo + SoC arm. Such efforts on active treatments (beyond the placebo + SoC data utilized in the present analysis) should enhance our understanding of the effects of different pharmacologic mechanisms on the underlying dynamics of disease trajectory and disease pathophysiology, as has been described in other therapeutic areas.[34–36]

One goal of this analysis was to identify study populations likely to have a low response to placebo + SoC to inform potential patient enrollment hypotheses. This can be achieved by considering clinical trial enrollment criteria on the influential covariates (baseline total number of BILAG scores A and B, time since diagnosis, C3/C4 below the LLN, and baseline total SLEDAI) such that they minimize future change in outcomes. Ethnicity, defined as (i) non-Hispanic, (ii) Hispanic – North America, (iii) Hispanic – Central/South America, and (iv) missing, was also an influential predictor of patient outcome. Both non-Hispanic patients and Hispanic patients from North America were found to have similar disease trajectories, and both appeared to have substantially different trajectories than those of Hispanic patients from Central/South America (not shown).

Simulations from the final component and composite models indicate that disease trajectory across the endpoints is conserved in Asian versus non-Asian patients identified in our dataset (i.e., White and other). These findings may have important considerations for the design of multiregional clinical trials (MRCTs) as insight into the contribution of demographic covariates to interindividual differences in disease trajectory allow sponsors to better plan and refine sample size for MRCTs. A principal assumption underlying MRCT design is that drug and disease-related intrinsic and extrinsic factors are reasonably conserved across the overall population.[37,38] If conserved and if there is similarity in disease trajectory across Asian and non-Asian patients, conducting Asian-inclusive global pivotal trials would be supported. Given the similar trajectory of SLE clinical endpoints in Asian and non-Asian patients demonstrated in our analysis, we posit that Asia-inclusive development strategies should be considered in SLE drug development to minimize access lag via inclusive development strategies that leverage *totality of evidence* principles.[39,40] Of course, this will require timely completion of necessary drug-specific ethnic sensitivity evaluations to support the MRCT design. Of note, the observed differences in disease trajectory between Hispanic patients from Central/South America and the remainder of the population may suggest that stratification by region may be appropriate to control heterogeneity and maximize treatment response in the design of pivotal MRCTs in SLE.

Given the well-recognized heterogeneity of SLE disease biology,[41] we anticipate that machine learning

methodologies[42] could further support identification of prognostic and predictive factors at the molecular level by incorporating patient-level deep biological profiling data, such as transcriptomic signatures.[43]

Our analysis represents a first step in the development of a quantitative model-based framework to enable clinical trial simulations to optimize SLE trial designs. Simulations with population disease trajectory models, as presented herein, can help improve trial design (e.g., optimize the duration and sample size of proof-of-concept trials) and assist in objectively estimating the probability of success ahead of initiating phase III trials. Although based on patient-level data from placebo + SoC arms of phase II and III SLE trials, in principle, the structural models should be applicable for population exposure-response analysis of longitudinal data in phase II proof-of-concept trials to inform phase III trial design decisions (e.g., endpoint selection, dose selection, and stratification factors for appropriate control of heterogeneity). In a previously published population exposure-response analysis of efficacy of the type I interferon receptor antibody anifrolumab in SLE, repeated measures logistic regression was used to estimate the underlying exposure-response for the time course of SRI-4 response rate in phase II to inform phase III dose selection.[44] Our analysis borrows longitudinal data from across multiple endpoints in a latent variable model, which, in principle, should provide a fundamental understanding of sources of variability and exposure-response relationships in the underlying disease process. Given that the leave-one-study-out cross-validation evaluations showed generally acceptable model performance, the final model can be applied as a Bayesian prior for analyzing emerging data from phase II trials. It should also be possible to simulate an investigational agent's expected long-term performance (i.e., 1-year duration) in phase III trials, provided the assumption of conserved longer term disease trajectory, beyond the duration of the phase II trial, is reasonably supported by biological considerations.

Drug development in SLE has been challenged by low success rates,[45] discordance in trial performance across primary endpoints, and the lack of reliable short-term outcome biomarkers of efficacy. Population disease trajectory models using patient-level data from placebo + SoC arms of diverse phase II/III SLE trials, as developed in this analysis, may help establish model informed drug development frameworks to support principled decision making and increase the probability of success in SLE clinical drug development. Simulations from the model developed herein could be helpful for drug developers to investigate the design of futility or interim analyses to evaluate predictive performance for early termination of a study if treatment response is unlikely to be significantly different from SoC at an earlier timepoint (e.g., after 6 months). With multiple clinical endpoints in SLE, their relative performance characteristics (e.g., sensitivity and variability), can be assessed in silico using the developed model to optimize trial design. Obvious application of such simulations could be to assist future protocol design, such as the role of interim analysis as an optimization tool,[46,47] the relevance or not of run-in phases,[48] treatment duration for optimal characterization of treatment effect size,[49] and potential impact of comorbidities. Additional applications may extend to adaptive protocols or enrichment trial designs to increase the overall probability of success of a trial. We hope that the current framework will stimulate applications to address these questions for future investigational agents likely to be developed for SLE. To this end, we envision this disease trajectory model serving as a first step for future model-informed precision medicine development in SLE.

AUTHOR CONTRIBUTIONS

K.G., J.F., R.G., M.G., P.G., Y.L., A.K., F.C.S., M.S., O.G., R.T., C.V.M., A.A., L.B., and K.V. wrote the manuscript. K.G., J.F., R.G., F.C.S., A.K., Y.L., L.B., O.G., A.A., M.G., P.G., M.S., C.V.M., R.T., and K.V. designed the research. J.F., R.G., K.G., K.V., Y.L., F.C.S., and A.K. performed the research. K.G., J.F., Y.L., F.C.S., A.K., R.G., and K.V. analyzed the data. J.F. and R.G. contributed new analytical tools.

ACKNOWLEDGMENTS

The authors would like to thank Lena Klopp-Schulz and Jamie Shaw for their valuable discussions and input.

FUNDING INFORMATION

This study was funded by EMD Serono.

CONFLICT OF INTEREST

J.F. and R.G. are employees of Metrum Research Group and were paid consultants of EMD Serono at the time this research was conducted. K.G., Y.L., F.C.S., A.A., R.T., C.V.M., M.S., O.G., A.K., P.G., L.B., and K.V. are employees of EMD Serono/Merck KGaA. As an Associate Editor for CPT: Pharmacometrics & Systems Pharmacology, Jonathan French was not involved in the review or decision process for this paper.

ORCID

Kosalaram Goteti https://orcid.org/0000-0002-8971-8283
Jonathan French https://orcid.org/0000-0001-9649-169X
Marc Gastonguay https://orcid.org/0000-0002-5725-5835
Karthik Venkatakrishnan https://orcid.org/0000-0003-4039-9813

REFERENCES

1. Romero-Diaz J, Isenberg D, Ramsey-Goldman R. Measures of adult systemic lupus erythematosus: updated version of British Isles lupus assessment group (BILAG 2004), European consensus lupus activity measurements (ECLAM), systemic lupus activity measure, revised (SLAM-R), systemic lupus activity questionnaire for population studies (SLAQ), systemic lupus erythematosus disease activity index 2000 (SLEDAI-2K), and systemic lupus international collaborating clinics/American College of Rheumatology Damage Index (SDI). *Arthritis Care Res*. 2011;63(Suppl 11):S37-S46.
2. Luijten KM, Tekstra J, Bijlsma JW, Bijl M. The systemic lupus erythematosus responder index (SRI); a new SLE disease activity assessment. *Autoimmun Rev*. 2012;11:326-329.
3. Furie R, Morand EF, Bruce IN, et al. 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. *Arthritis Rheumatol*. 2021;73:2059-2068.
4. Franklyn K, Lau CS, Navarra SV, et al. Definition and initial validation of a lupus low disease activity state (LLDAS). *Ann Rheum Dis*. 2016;75:1615-1621.
5. Mahieu MA, Strand V, Simon LS, Lipsky PE, Ramsey-Goldman R. A critical review of clinical trials in systemic lupus erythematosus. *Lupus*. 2016;25:1122-1140.
6. Manzi S, Raymond S, Tse K, et al. Global consensus building and prioritisation of fundamental lupus challenges: the ALPHA project. *Lupus Sci Med*. 2019;6:e000342.
7. Furie RA, Morand EF, Bruce IN, 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-e219.
8. Morand EF, Furie R, Tanaka Y, et al. Trial of Anifrolumab in active systemic lupus erythematosus. *N Engl J Med*. 2020;382:211-221.
9. Thanou A, Chakravarty E, James JA, Merrill JT. Which outcome measures in SLE clinical trials best reflect medical judgment? *Lupus Sci Med*. 2014;1:e000005.
10. Touma Z, Moghaddam B, Su J, Katz P. Cognitive function trajectories in association with the depressive symptoms trajectories in systemic lupus erythematosus over time. *Arthritis Care Res*. 2021;73:1436-1443.
11. Yip TC, Saria S, Petri M, Magder LS. Predictors of the start of declining eGFR in patients with systemic lupus erythematosus. *Lupus*. 2021;30:15-24.
12. Alliance LR. Current projects: drivers of response for SRI-4 and BICLA in systemic lupus erythematosus (SLE) clinical trials. https://www.lupusresearch.org/for-researchers/research-partners/. Accessed November 13, 2022.
13. Kalunian KC, Kim M, Xie X, Baskaran A, Daly RP, Merrill JT. Impact of standard of care treatments and disease variables on outcomes in systemic lupus erythematosus trials: analysis from the Lupus Foundation of America collective data analysis initiative. *Eur J Rheumatol*. 2016;3:13-19.
14. Kim M, Merrill J, Kalunian K, et al. Brief report: longitudinal patterns of response to standard of care therapy for systemic lupus erythematosus: implications for clinical trial design. *Arthritis Rheumatol*. 2017;69:785-790.
15. Kim M, Merrill JT, Kalunian K, Hanrahan L, Izmirly P. Evaluating duration of response to treatment in systemic lupus erythematosus clinical trials. *Lupus Sci Med*. 2018;5:e000266.
16. FDA. Best practices for development and application of disease progression models. 2021. https://www.fda.gov/drugs/news-events-human-drugs/best-practices-development-and-application-disease-progression-models-11192021. Accessed July 11, 2021.
17. Yin PT, Desmond J, Day J. Sharing historical trial data to accelerate clinical development. *Clin Pharmacol Ther*. 2019;106:1177-1178.
18. Team RC. R: A language and environment for statistical computing. 2021. www.R-project.org/. Accessed November 13, 2022.
19. Team SD. Stan Modeling Language Users Guide and Reference Manual 2.28. 2021. https://mc-stan.org. Accessed November 13, 2022.
20. Team SD. RStan: the R interface for Stan. 2021. https://mc-stan.org/rstan/. Accessed November 13, 2022.
21. Hutmacher MM, French JL. Extending the latent variable model for extra correlated longitudinal dichotomous responses. *J Pharmacokinet Pharmacodyn*. 2011;38:833-859.
22. Hu C, Xu Z, Mendelsohn AM, Zhou H. Latent variable indirect response modeling of categorical endpoints representing change from baseline. *J Pharmacokinet Pharmacodyn*. 2013;40:81-91.
23. Kursa MB, Jankowski A, Rudnicki WR. Boruta – a system for feature selection. *Fund Inform*. 2010;101:271-285.
24. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. 2017. ArXiv abs/1705.07874.
25. Terranova N, French J, Dai H, et al. Pharmacometric modeling and machine learning analyses of prognostic and predictive factors in the JAVELIN gastric 100 phase III trial of avelumab. *CPT Pharmacometrics Syst Pharmacol*. 2022;11:333-347.
26. Ripley B. *Pattern Recognition and Neural Networks*. Cambridge University Press; 1996.
27. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput*. 2017;27:1413-1432.
28. Watanabe S. A widely applicable Bayesian information criterion. *J Mach Learn Res*. 2013;14:867-897.
29. Harrell FE Jr. *Regression Modeling Strategies: with Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Springer; 2015.
30. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21:128-138.
31. Della Pasqua O, Santen GW, Danhof M. The missing link between clinical endpoints and drug targets in depression. *Trends Pharmacol Sci*. 2010;31:144-152.
32. Santen G, Danhof M, Della PO. Sensitivity of the Montgomery Asberg depression rating scale to response and its consequences for the assessment of efficacy. *J Psychiatr Res*. 2009;43:1049-1056.
33. Stohl W, Hilbert DM. The discovery and development of belimumab: the anti-BLyS-lupus connection. *Nat Biotechnol*. 2012;30:69-77.
34. Chen C, Jonsson S, Yang S, Plan EL, Karlsson MO. Detecting placebo and drug effects on Parkinson's disease symptoms by longitudinal item-score models. *CPT Pharmacometrics Syst Pharmacol*. 2021;10:309-317.
35. D'Agate S, Wilson T, Adalig B, et al. Model-based meta-analysis of individual international prostate symptom score trajectories

in patients with benign prostatic hyperplasia with moderate or severe symptoms. *Br J Clin Pharmacol*. 2020;86:1585-1599.
36. Cummings JL. Challenges to demonstrating disease-modifying effects in Alzheimer's disease clinical trials. *Alzheimers Dement*. 2006;2:263-271.
37. ICH. E17 general principles for planning and design of multi-regional clinical trials guidance for industry. 2018. Published online March 28, 2022. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e17-general-principles-planning-and-design-multi-regional-clinical-trials. Accessed November 13, 2022.
38. Venkatakrishnan K, Gupta N, Smith PF, et al. Asia-inclusive clinical Research and Development enabled by translational science and quantitative clinical pharmacology: toward a culture that challenges the status quo. *Clin Pharmacol Ther*. 2022.
39. Venkatakrishnan K, Benincosa LJ. Diversity and inclusion in drug development: rethinking intrinsic and extrinsic factors with patient centricity. *Clin Pharmacol Ther*. 2021;112:204-207.
40. Venkatakrishnan K, Cook J. Driving access to medicines with a totality of evidence mindset: an opportunity for clinical pharmacology. *Clin Pharmacol Ther*. 2018;103:373-375.
41. Ruiz-Cerdá ML, Irurzun-Arana I, González-Garcia I, et al. Towards patient stratification and treatment in the autoimmune disease lupus erythematosus using a systems pharmacology approach. *Eur J Pharm Sci*. 2016;94:46-58.
42. Terranova N, Venkatakrishnan K, Benincosa LJ. Application of machine learning in translational medicine: current status and future opportunities. *AAPS J*. 2021;23:74.
43. Allen ME, Rus V, Szeto GL. Leveraging heterogeneity in systemic lupus erythematosus for new therapies. *Trends Mol Med*. 2021;27:152-171.
44. Chia YL, Santiago L, Wang B, et al. Exposure-response analysis for selection of optimal dosage regimen of anifrolumab in patients with systemic lupus erythematosus. *Rheumatology (Oxford)*. 2021;60:5854-5862.
45. Murphy G, Isenberg DA. New therapies for systemic lupus erythematosus – past imperfect, future tense. *Nat Rev Rheumatol*. 2019;15:403-412.
46. Santen G, van Zwet E, Bettica P, Gomeni RA, Danhof M, Della Pasqua O. From trial and error to trial simulation III: a framework for interim analysis in efficacy trials with antidepressant drugs. *Clin Pharmacol Ther*. 2011;89:602-607.
47. Goyal N, Gomeni R. A latent variable approach in simultaneous modeling of longitudinal and dropout data in schizophrenia trials. *Eur Neuropsychopharmacol*. 2013;23:1570-1576.
48. Boessen R, Heerspink HJL, de Zeeuw D, Grobbee DE, Groenwold RHH, Roes KCB. Improving clinical trial efficiency by biomarker-guided patient selection. *Trials*. 2014;15:103.
49. Quartagno M, Walker AS, Carpenter JR, Phillips PP, Parmar MK. Rethinking non-inferiority: a practical trial design for optimising treatment duration. *Clin Trials*. 2018;15:477-488.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

[Correction added on 1 December 2022, after first online publication: Supporting information has been updated in this version].

How to cite this article: Goteti K, French J, Garcia R, et al. Disease trajectory of SLE clinical endpoints and covariates affecting disease severity and probability of response: Analysis of pooled patient-level placebo (Standard-of-Care) data to enable model-informed drug development. *CPT Pharmacometrics Syst Pharmacol*. 2023;12:180-195. doi:10.1002/psp4.12888