A Bayesian Clinical Trial Simulation for Expected Value of Sample Information Analysis in Systemic Lupus Erythematosus

Watson, P. and Brennan, A.

Introduction

Designing clinical trials in Systemic Lupus Erythematosus (SLE) has been challenging for investigators. Value of Information can be used to optimise the design of clinical trials and can be used to improve data collection for use in cost-effectiveness analysis.

Objective

To develop a simulation model for clinical trial outcomes in SLE to sample future trial results in Expected Value of Sample Information analysis. To validate the simulation against an SLE observational cohort and SLE clinical trials.

Method

An individual patient simulation model generates a representative sample of 30,000 SLE patients. Individuals are entered into the trial if they meet the trial inclusion criteria. The patient's disease activity profile is updated at 3 monthly intervals until the trial duration has elapsed. Patients can withdraw from the trial in any period due to lack of efficacy, adverse event or other reason. The simulation is flexible to allow alternative inclusion criteria, sample size, duration and primary endpoint definition.

Results

Most of the simulation outcomes are not statistically significantly different from the observational cohort of SLE patients. Prednisone dose is underestimated in the simulation compared with the first year of the observational cohort, but provides a more stable estimate suitable for clinical trials. The clinical trial outcomes are similar to recently complete Phase III trials.

Discussion

The simulation provides realistic estimates of clinical trial outcomes based on comparisons with observation cohorts and clinical trials. The simulation will be used in the future in value of information analyses to evaluate alternative trial designs.

Introduction

It is important to consider cost-effectiveness constraints when planning and designing data collection in drug development programmes. Budget constrained healthcare systems have sought methods to incorporate costs and health benefits into decision making about treatment access. Institutions have been established at the national level, such as the National Institute for Health and Clinical Excellence (NICE) in the UK, PBAC in Australia, and Canadian Agency for Drugs and Technologies in Health (CADTH) in Canada. These use Cost per Quality Adjusted Life Years (QALYs), to assess the value of new treatments compared with current standard care. Economic evaluation is an important hurdle in drug development, and data collected in Phase III trials are often used to estimate effectiveness in cost-effectiveness analyses. Therefore, uncertainty in the outcomes of clinical trials directly impacts on the uncertainty in the economic value of the new treatments compared with the old.

Systemic Lupus Erythematosus (SLE) is a multi-system autoimmune disorder with variable manifestations, characterised by periods of active disease and remission. The disease pathways and patient outcomes are extremely heterogeneous across patients. Inflammation can affect multiple organ systems including, skin, musculoskeletal, neuropsychiatric, renal, cardiovascular, respiratory, and haematological systems. Active disease is managed with immunosuppressive therapy. Persistent disease activity, prednisone dose, and cytotoxic treatments have all been associated with severe long term outcomes of organ damage and mortality.

Until recently, there were few Randomised Controlled Trials (RCTs) in SLE, no cost-effectiveness analyses, and no treatment had received a license for SLE for over 20 years. In the last 10 years a strong pipeline of new biologics treatments has generated a great deal of interest and excitement among clinicians. Given the complex and heterogeneous nature of the disease, designing clinical trials has been challenging. A large number of clinical trials in SLE have failed to meet their primary endpoint, despite clinicians' belief in the effectiveness of the treatment, suggesting that the trial may not have demonstrated the efficacy of the drugs (1;2). This is a complex disease and very few trials have been successful, therefore there is much uncertainty and many outstanding questions about how best to design future clinical trials.

There is limited information available about SLE trials and how alternative design decisions will affect the trial outcomes. Currently, there are several questions that investigators may be interested in when designing clinical trials in SLE.

1. What sample size is required?
2. What duration of follow-up is needed to demonstrate long term benefits of treatment?
3. What inclusion criteria should be imposed in the Phase III trial?
4. How should the primary endpoint of the trial be specified?

Analytical models can be used to assist with the design of clinical trials by simulating potential outcomes of trials under alternative designs. This approach is less costly than testing alternative trial designs with real treatments and patients.

Expected Value of Sample Information (VOI) analyses are used to assess the value of future research programmes by estimating the expected reduction in uncertainty for the decision maker after data collection. VOI has been used to design future clinical trials (3-6). A decision model with uncertain parameters, θ, which assesses two or more interventions, D, generates net benefit for each intervention given the parameters of the model, NB(D,θ). The Expected Value of Sample Information (EVSI) is described by:

$$EVSI(\theta) = E_{X_\theta}\left[\max_D\{E_{\theta|X}(NB(D,\theta))\}\right] - \max_D\{E_\theta(NB(D,\theta))\}$$

where EVSI is expressed by the difference between the expected value of a decision made after collecting information X_θ from a trial, and the expected value of a decision made with current information. These methods can be employed before embarking on a Phase III trial, to identify the design which minimises the uncertainty in the economic evaluation outcomes. Sampled estimates of realistic trial outcomes, for a range of trial designs, are required to complete EVSI analysis.

It is often possible to sample from an efficacy parameter generated from previous clinical trials in EVSI. However, there are several scenarios where this approach will not be satisfactory to address all the questions above. For example, a prior parameter based on a single study may underestimate variability in sampled populations, which may be important in simulating heterogenous diseases like SLE. Secondly, the impact of follow-up cannot be investigated if the investigator wishes to extend the duration of the Phase III trial beyond that used in the Phase II trial. Thirdly, the inclusion criteria could be altered in Phase III to target

a different patient profile, which will impact on the distribution of the clinical trial outcomes. Finally, the trial investigators may wish to collect clinical outcomes that were not included in the Phase II trial.

Therefore, sampling from Phase II trial outcomes would either limit the scope of the EVSI or rely on other means to estimate prior parameters, such as elicitation (7). However in there are a number of longitudinal cohort studies of SLE, which provide a rich source of data to describe the patterns of disease in individual patients. Furthermore, data analysis facilitates detailed description of dependencies and correlation between clinical outcomes which would be challenging to elicit from clinical experts. Therefore, we developed a Bayesian Clinical Trial Simulation (BCTS to simulate the natural history of the disease, utilising observational cohort data to estimate the effect of new treatments in modifying disease progression.

This article presents a detailed description of the SLE BCTS model. The utility of the simulation depends on the reliability of the simulation output. The SLE BCTS model must generate realistic patient disease outcomes at the individual and cohort level so that the different clinical trial designs can be assessed. The aims of the article are to describe the methods used to generate clinical trial outcomes and to validate these against observational cohort data and published clinical trials. In the next section the structure of the simulation is presented, followed by a description of the data sources used, a description of the simulation process and the tests used to assess the validity of the simulation.

Method

Simulation design

The simulation is designed to generate clinical trial outcomes based on prior knowledge of the disease and the efficacy of the new treatment. The simulation is intended to be used after a drug has successfully completed a Phase II trial as an aid to designing the Phase III trial. The simulated clinical trial is a two arm RCT comparing a new biologic drug and standard of care against placebo and standard of care. The clinical trial is modelled as an individual patient simulation where patient status is updated in three monthly intervals. The simulation generates measures of the disease that are useful in describing the efficacy of the treatment and are used in the cost-effectiveness model for SLE.

The simulation generates a large SLE population from which patients are recruited into the trials. This ensures that the simulation captures heterogeneity in sampled patient profiles, which may impact on the variability of trial outcomes.

Modelling the disease

Disease Activity is measured in the trial using the SELENA revision of the SLE Disease Activity Index (SLEDAI) (8). Other measures of disease Activity such as the British Isles Lupus Assessment Group (BILAG) (9) and the Physicians Global Assessment (PGR) may also be used in clinical trials as either the primary endpoint or as part of a composite responder index. However, these are not included in the simulations. The SLEDAI is a 24 item index describing the overall burden of the disease across multiple organ systems (see Table 3 for list of symptoms). Each symptom of the SLEDAI has been weighted by severity. The total SLEDAI score can be used as a composite measure of the overall severity of the disease.

In the simulation each of the 24 items of the SLEDAI are modelled independently. Although this approach increases the complexity of the simulation it enables the details of organ involvement to be retained. The simulated trial uses a random effects logistic regression model to describe the probability of a symptom being present at any 3 monthly visit.

$$\Pr(y_{ij} = 1 | x_{ij}, \zeta_j) = \frac{\exp(\boldsymbol{x}_{ij}\boldsymbol{\beta} + \zeta_j)}{1 + \exp(\boldsymbol{x}_{ij}\boldsymbol{\beta} + \zeta_j)}$$

Where y_{ij} indicates the item of the SLEDAI, $\boldsymbol{x}_{ij}$ is a vector of covariate data for patient characteristics and symptoms present at the previous visit, and $\boldsymbol{\beta}$ are estimated coefficients for the impact of the covariates on the probability. The function includes a random effects parameter describing each patient's unobservable propensity to the symptom ζ_j, which varies between patients j but is constant with time i. In the simulation, the random effects parameters are assigned to patients by sampling from a multivariate normal distribution, with a covariance matrix describing the variability between patients and correlation between the items of the SLEDAI. This incorporates heterogeneity in the simulation and generates correlation between items that are more commonly observed within a patient.

The prednisone dose is assumed to be conditional on disease activity. The simulation must reflect the proportion of patients receiving prednisone and the respective dose. The daily dose of prednisone is best described as count data due to clusters around certain dose types. A

hurdle model was used to estimate prednisone dose, comprising a logistic function to determine the probability that prednisone was prescribed, and a zero truncated negative binomial model to estimate the level of prednisone dose prescribed.

Modelling Efficacy

In most practical settings statistical analysis of the Phase II data would estimate the treatment effect. The logistic model presented above would be structurally similar to the models used in the natural history model, but extended to include an additional parameter Δ, indicating the revised probability for treated patients.

$$\Pr\left(y_{ij} = 1 \middle| x_{ij}, \zeta_j\right) = \frac{\exp\left(\boldsymbol{x}_{ij}\boldsymbol{\beta} + D_{ij}\Delta + \zeta_j\right)}{1 + \exp\left(\boldsymbol{x}_{ij}\boldsymbol{\beta} + D_{ij}\Delta + \zeta_j\right)}$$

However, in this study the treatment was a hypothetical new biologic drug and the distribution of the log odds ratio must be assumed rather than estimated. Data from a recently completed Phase II trial was sought from two publications (10;11). The odds ratio for response was distributed between the 24 items of the SLEDAI and weighted by the prevalence of each item. The published data showed that there was evidence that the treatment effect varied between organ systems and the simulation was adjusted to reflect this pattern. The methods used to describe efficacy are not intended to be used in future applications but have been designed to estimate realistic distributions for the log odds ratio for treatment effect.

Modelling Safety/Withdrawal

Safety outcomes are important in Phase III trials and are used in submissions to regulatory bodies such as the Food and Drugs Agency (FDA). A broad range of adverse events are often observed and affect a large proportion of patients in both arms of the clinical trials. Furthermore, it is difficult to distinguish between disease related symptoms and treatment induced symptoms. Infections are frequently observed in SLE patients due to the background therapies they receive, which complicates the detection of infections in patients treated with the new drug. Adverse events are often graded by severity in the clinical trial to compare across treatment arms. However, monitoring adverse events systematically in a multi-centre RCT can be challenging. This also means that it is difficult to incorporate adverse events into the simulation and safety outcomes are not currently simulated in the SLE BCTS.

Withdrawal from treatment is an important parameter in cost-effectiveness analyses and consequently simulated in the clinical trial. A systematic review was conducted to identify SLE trials. The review identified 12 clinical trials of which 5 were completed with sufficiently large sample sizes and reported reasons for withdrawal from the trial (1;2;10;12-14). The number of patients who withdrew due to adverse events, lack of efficacy and other administrative causes was extracted from these studies. The rate of withdrawal from adverse events and other causes in the simulation are estimated as a weighted average of rates observed in recently published clinical trials (Table 1). The probability of withdrawing due to lack of efficacy is conditional on SLEDAI score, and therefore patients who experience an increase in SLEDAI score of greater than 4 units are eligible for withdrawal due to lack of efficacy. It is assumed that 40% of these patients will withdraw from the trial. The value of 40% was obtained through an iterative process which calibrated outcomes of the simulation against the trial results reported in Table 1.

Source of Inputs

Hopkins Lupus Cohort

The Hopkins Lupus Cohort is a longitudinal study of SLE patients which collects demographic details, disease characteristics and their past medical history before cohort entry (15). Patients in the Hopkins Lupus cohort visit the clinic at least every 3 months. Data from the Hopkins Lupus Cohort collected between 1987 and 2010 were used in the statistical analyses. The data includes detailed description of the SLEDAI score, prednisone dose and records organ damage events included in the SLICC/ACR DI. The Hopkins Lupus cohort is used to specify parameters for the baseline characteristics of the large simulated SLE population.

Extensive statistical analyses of this cohort have generated a large set of statistical models to predict the probability each of the 24 items of the SLEDAI, whether prednisone is prescribed and what dose of prednisone is prescribed. The functional forms of the SLEDAI and prednisone dose regression models are described above. The regression point estimates and covariance matrices are sampled in the simulation from a multivariate normal distribution to reflect uncertainty in the estimates of the natural history of the disease. The treatment effect parameters have been appended onto the multivariate normal distributions, where the variance is taken from the phase II trial data and the model assumes zero covariance with other parameters from the natural history model. A total of 227 parameters are used to

simulate the SLEDAI and prednisone, as a consequence the parameter estimates are not reported here.

The Simulation Process

The stages of the clinical trial simulation are illustrated in a model schematic in Figure 1. In the first phase of the simulation an SLE population was generated to represent a real-life population of SLE patients, and the heterogeneity between patients. The inclusion criteria for the trial were specified and patients were selected from the SLE population according to those criteria. Once a cohort of patients was selected to meet the sample size of the trial the patients were randomised and their baseline disease profile was recorded.

In the second phase of the simulation the patients entered into the clinical trial and were monitored over time. The simulation used the statistical models developed from the Hopkins Lupus Cohort to update patient status every 3 months. The process of updating the patient's status was repeated until the duration of follow-up for the trial is reached.

In the final phase of the simulation the outcomes of the patient were recorded and summarised to evaluate the efficacy of the new treatment.

Validating Clinical Trial Outcomes

The SLE BCTS will be useful in developing future SLE clinical trial designs if it is able to simulate realistic outcomes and reflect the uncertainty of trial. A validation of the SLE BCTS was performed to evaluate whether the simulation predicts realistic SLE outcomes. The validity of the simulation can be assessed by comparing the outcomes of the simulation against real life observations from observational data and clinical trials. The simulation output was compared against the Hopkins Lupus cohort, and three successful clinical trials reporting outcomes using a SLEDAI index.

The validation exercise aims to observe whether the clinical trial simulation meets the following assessments, which are used to indicate how well the simulation fits to reality.

1. 95% of simulation outcomes should be within the 95% confidence interval of observed Hopkins Lupus outcomes.
2. The distribution of simulated SLEDAI scores and prednisone doses should have a similar distribution to the Hopkins data.

3. Observed clinical trial endpoints should fall within the distribution of simulated trial endpoints and the discrepancy between mean simulated output and observed output should not be greater than the observed margin of error at the 5% significance level.

The criteria were assessed with the following analyses from three simulation runs. each of 1000 iterations. The accumulation of organ damage and mortality events were not be tested in this analysis.

Simulation 1- The Hopkins Lupus Cohort

The simulation was run to reproduce outcomes from the Hopkins Lupus cohort for the first 5 years of individual's follow-up. The SLE population used in this simulation was based on the baseline characteristics of the Hopkins SLE population. Summary statistics for the baseline profile of patients are reported in Table 2. The analysis was only conducted on placebo patients to compare the natural history outcomes for patients without treatment. Statistical tests at each observation period were performed and the proportion of observation more extreme than the observed data was evaluated.

Simulation 2 – The Phase II Clinical Trial

The simulation was set up to run a 1 year clinical trial of 321 patients. The baseline characteristics of the patients were designed to match with the demographic profile of the belimumab Phase II trial. Treatment efficacy was estimated from summary statistics reported in Furie et al. (2009). Summary statistics for the baseline profile of patients is reported in Table 2. The mean percentage change in SELENA SLEDAI score from baseline to week 52, and the percentage of patients who experienced a change in SELENA SLEDAI score of 4 units or more in each treatment arms, were reported for each simulation run. The odds ratio for treatment effect was estimated. The simulated outcomes are compared with the published trial outcomes from Wallace and Furie et al (10;11).

Simulation 3 – The Phase III Clinical Trial

Summary statistics for the baseline profile of patients is reported in Table 2. The simulation was set up to run a 1 year clinical trial of 800 patients. The baseline characteristics of the patients were designed to match with the demographic profiles of the belimumab Phase III trials. Prior estimates for treatment efficacy were estimated from summary statistics reported in Furie et al. (2009). The percentage of patients who experienced a change in SELENA

SLEDAI score of 4 units or more in the treatment and placebo arms were reported for each simulation run. The odds ratio for treatment effect was calculated.

Results

The Hopkins Lupus Cohort

The simulation replicates most of the outcomes from the Hopkins Lupus Cohort with good accuracy. The proportion of patients with each SLEDAI item was recorded at every visit for five years and compared with those observed in the first five years of the Hopkins lupus Cohort. Graphical plots indicated that the simulated outcomes were close to the Hopkins data and that most of the Hopkins observations fell within the 95% percentile plots to indicate the upper and lower limits of the simulation (results not reported here). The simulation demonstrated clear downward trend in the proportion of patients which is observed in the Hopkins cohort. Table 3 reports summary results for statistical tests on the difference between the simulated proportion at each time point and the observed Hopkins data. In eighteen of the items of the SLEDAI the proportion of simulated outcomes that were statistically significantly different from the Hopkins data was less than 5%. Two items of the SLEDAI had more than 10% of simulated outcomes that were statistically significantly different.

Figure 2 illustrates a histogram of SLEDAI scores generated in the simulation and compares this with a histogram of observed SLEDAI score from the Hopkins cohort. The diagram shows that the simulation produces a very similar distribution of SLEDAI score to those observed in the Hopkins Lupus Cohort.

Figure 3 illustrates a histogram of Prednisone doses generated in the simulation and compares this with a histogram of observed Prednisone dose in the Hopkins cohort. The simulation accurately simulates the proportion of patients with zero Prednisone dose. The distribution of positive prednisone doses in the simulation is much smoother than those observed in the Hopkins Cohort. The real-life observations cluster around certain values, which is difficult to replicate in a simulation. prednisone dose in the first year of the simulation is significantly lower than that observed in the Hopkins cohort (results not reported here). This is due to the aggressive treatment of patient’s when they are referred to doctors at the Hopkins cohort. This pattern does not need to be replicated in the simulation.

Phase II trial results

Figure 4 reports the percentage change in SLEDAI score from baseline to week 52 reported in the belimumab Phase II trial and the Phase II simulation model. The thin line depicts a normal distribution fitted to summary statistics reported in Wallace et al. (2009) (10). The bars indicate the distribution of 1000 simulated trial results. The two graphs describe the outcomes for placebo and treatment. In both diagrams the simulation tends to produce lower percentage change in SLEDAI scores than those observed in the Phase II trial. However, the Phase II results are within the simulated range of possible trial outcomes.

Figure 5 reports the proportion of patients who achieve a 4 unit reduction in the SLEDAI score in the Phase II trial. The mean difference in the proportion of responders and the odds ratio of response are also illustrated. The thin line represents a normal distribution fitted to the summary statistics of responders reported in Furie et al. (2009) (11) and the bars represents the distribution of simulated outcomes. The graphs show that the simulation tends to predicted greater proportions of responders than those observed in the Phase II trial. However, the difference between placebo and treatment and odds ratio have the same distribution as that fitted to the Phase II trial summary statistics.

Phase III belimumab trial

The Phase III belimumab trial and simulation results are illustrated in Figure 6. The thin line represents a normal distribution fitted to relevant summary statistics from two Phase III trials (12;13). The bars represent the distribution of simulated observations. The simulation output fits well with the summary statistics for the Phase III trial. The simulation tends to predict slightly lower proportions of responding patients than that observed in the Phase III trial.

Discussion

Analysis of the SLE BCTS has shown that the simulation accurately predicts most outcomes of the Hopkins Cohort. Eighteen items of the SLEDAI were well matched to the cohort observations and followed time trends in the data. Six of the items of the SLEDAI estimated proportions of patients with symptoms considered extreme, which suggests that there may be some bias in the incidence of these symptoms in the simulation. However, the overall distribution of simulated composite SLEDAI scores and prednisone dose were well matched to the Hopkins data. The simulation generates clinical trial endpoints that are similar to those that have been observed in large SLE clinical trials. Overall the validation process has illustrated that the simulation outcomes are not substantially different from real-life patient

outcomes observed in clinical trials and the Hopkins Lupus Cohort. Therefore, the simulation can be useful in predicting realistic outcomes from clinical trial endpoints.

The simulation generates a detailed disease activity profile for the individual patients and incorporates correlations between items of the SLEDAI to generate realistic disease outcomes. The SLEDAI is not an exhaustive disease activity index, and it is believed that two or more disease activity indices could be employed in clinical trials to more sensitively capture variations in the disease (11). Furie et al. (2009) used data from the Phase II trial to develop an appropriate composite endpoint for the belimumab Phase III trials that defined responders by an improvement in the SLEDAI and no worsening in the BILAG or PGA. This is a stricter requirement for response that that used in the simulation, because the other indices are sensitive to symptoms not captured in the SLEDAI.

Other measures of disease activity were not included in the simulation for three reasons. Firstly, incorporating more than one disease activity index would increase both the computational complexity of the simulation and the analytical complexity of the statistical analyses required to describe correlations between the indices. Secondly, there are few longitudinal cohort studies that have monitored multiple disease activity indices over a long period. Thirdly, the current cost-effectiveness model for SLE relates the SLEDAI score to long term outcomes of the disease. Therefore, only the trial outcomes of SLEDAI can be applied to the cost-effectiveness model. However, this is a limitation of the simulation, and consequently the simulation is not able to simulate alternative primary endpoint definitions, which is an area of clinical trial design in SLE where many questions and uncertainties can be found.

Despite not generating disease outcomes for multiple disease activity indices, the simulation is a very complex model. This detailed approach allows the simulation to address questions relating to organ involvement, concomitant medications and long terms outcomes. However, an important drawback of this approach is that the simulation is relatively slow to generate outcomes. In these analyses, the comparison with the Hopkins Lupus Cohort (Simulation 1) took approximately 6 hours to run 1000 iterations on a standard desktop computer. Consequently, extensive analyses of multiple trial design questions will be slow to run and may need to be performed using a high performance computing server. The speed of the simulation also means that the model structure is not necessarily appropriate to use in a cost-effectiveness analyses, which would need to follow the cohort of patients until death.

Future adaptations of the simulation should be driven by the requirements of clinical investigators. If there is sufficient interest in testing alternative primary endpoints for future clinical trials then additional statistical analyses of the BILAG score will be required and a longitudinal cohort study of BILAG scores would be sought. The dataset would need to collect the BILAG score and SLEDAI simultaneously to appropriately describe the correlations between these scores.

The simulation can be used to simulate trial outcomes and can be used in an EVSI to generate sampled outcomes from clinical trials. The SLE BCTS will be used to simulate sampled data from short and long clinical trials in SLE to estimate the reduction in uncertainty in the long term extrapolation of treatment benefit and to value this reduction in uncertainty. It will be possible to estimate the reduction in uncertainty for alternative designs of clinical trial, to include testing alternative duration of follow-up, sample size, and inclusion criteria. EVSI is known to take a long time to compute, and will be problematic to apply to a complex disease like SLE. Our further research aims to pursue this challenge and develop a tool to evaluate alternative trial designs.

Reference List

(1) Merrill JT, et al. Activity of Abatacept in SLE: Results of a 12 month Phase II Exploratory Study. Ann Rheum Dis 2009 68(suppl3):70 . 2009.
Ref Type: Abstract

(2) Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis Rheum 2010 Jan;62(1):222-33.

(3) Stevenson MD, Oakley JE, Jones ML, Brennan A, Compston JE, McCloskey EV, et al. The Cost-Effectiveness of an RCT to Establish Whether 5 or 10 Years of Bisphosphonate Treatment Is the Better Duration for Women With a Prior Fracture. Medical Decision Making 2009;29(6):678-89.

(4) Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. Health Econ 1996 Nov;5(6):513-24.

(5) Brennan A, Kharroubi SA. Expected value of sample information for Weibull survival data. Health Economics 2007;16(11):1205-25.

(6) Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Medical Decision Making 2004;24(2):207-27.

(7) Bojke L, Claxton K, Bravo-Vergel Y, Sculpher M, Palmer S, Abrams K. Eliciting distributions to populate decision analytic models. Value Health 2010 Aug;13(5):557-64.

(8) Petri M, Kim MY, Kalunian KC, Grossman J, Hahn BH, Sammaritano LR, et al. Combined oral contraceptives in women with systemic lupus erythematosus. N Engl J Med 2005 Dec 15;353(24):2550-8.

(9) Isenberg DA, Rahman A, Allen E, Farewell V, Akil M, Bruce IN, et al. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology (Oxford) 2005 Jul;44(7):902-6.

(10) Wallace DJ, Stohl W, Furie RA, Lisse JR, McKay JD, Merrill JT, et al. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. Arthritis Rheum 2009 Sep 15;61(9):1168-78.

(11) Furie RA, Petri MA, Wallace DJ, Ginzler EM, Merrill JT, Stohl W, et al. Novel evidence-based systemic lupus erythematosus responder index. Arthritis Rheum 2009 Sep 15;61(9):1143-51.

(12) Navarra SV, Guzman RM, Gallacher AE, Hall S, Levy RA, Jimenez RE, et al. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. Lancet 2011 Feb 26;377(9767):721-31.

(13) Navarra S, et al. Belimumab, a BLyS-Specific Inhibitor, Reduced Disease Activity, Flares and Prednisone Use in PAtients with Active SLE: Efficacy and Safety Results From the Phase 3 BLISS-52 Study. Arthritis and Rheumatism 2009 (S10):60 . 2010.
Ref Type: Abstract

(14) Fortin PR, Abrahamowicz M, Ferland D, Lacaille D, Smith CD, Zummer M. Steroid-sparing effects of methotrexate in systemic lupus erythematosus: A double-blind, randomized, placebo-controlled trial. Arthritis Care and Research 1994; 59(12)(pp 1796-1804), 2008;15.

(15) Petri M. Lupus in Baltimore: evidence-based 'clinical pearls' from the Hopkins Lupus Cohort. Lupus 2005;14(12):970-3.

Table 1: Withdrawal from SLE clinical trials

Author (year)	Treatment	Sample size	No. Withdrawn from the trial			Rate of withdrawal from the trial		
			Efficacy	AE	Other	Efficacy	AE	Other
Fortin (2008)	MTX	41	7	5	2	0.171	0.122	0.049
	Placebo	45	10	0	2	0.222	0	0.044
Wallace (2009)	Belimumab	336	3	20	42	0.009	0.059	0.125
	Placebo	113	1	5	14	0.009	0.044	0.124
Merril (2010)	Abatacept	118	21	7	9	0.178	0.059	0.076
	Placebo	57	12	1	1	0.211	0.018	0.018
Merril (2010)	Rituximab	169	NA	19	30	NA	0.112	0.178
	Placebo	88	NA	3	11	NA	0.148	0.125
Navarra (2010)	Belimumab 1mg	578	24	31	42	0.042	0.054	0.073
	Placebo	287	16	19	26	0.056	0.066	0.091
Navarra (2011)	Belimumab 1mg	546	29	41	84	0.053	0.075	0.154
	Placebo	273	20	23	46	0.073	0.084	0.168
					Treatment	0.049	0.069	0.116
					Placebo	0.064	0.069	0.117

Figure 1: Model schematic

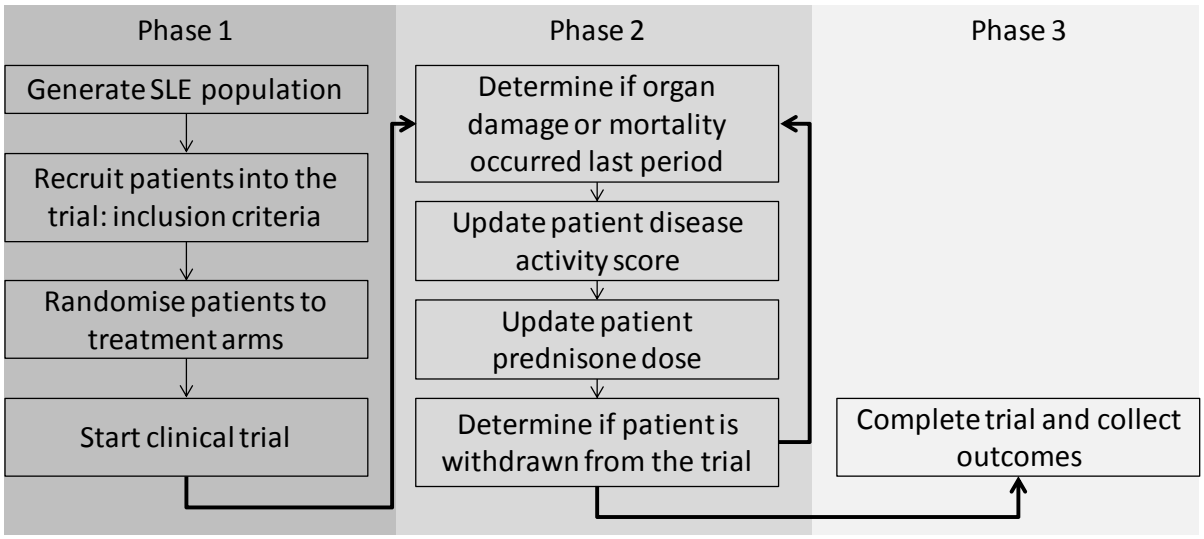

Table 2: Baseline characteristic of patients for each simulation

	Hopkins Simulation	Phase II Simulation	Phase III simulation
Number of patients	1354	321	842
Age mean (sd)	37.8 (13.05)	40.6 (11.0)	37.8 (11.8)
Disease duration mean (sd)	4.83 (6.3)	8.98 (8.1)	6.37 (6.2)
Women %	92.9	95.2	94.0
Black %	38.8	26.2	9.0

Table 3: Simulation results for the SLEDAI items and proportion of simulated outcomes within the 95% confidence interval of Hopkins results

SLEDAI Item	Proportion of events in Hopkins Lupus Cohort across all observations up to 5 years	Average simulated proportions of events across all observations up to 5 years	Proportion of simulated observations significantly different from Cohort
Seizure	0.002	0.003	0.017
Psychosis	0.001	0.002	0.014
Organic Brain Syndrome	0.004	0.006	0.030
Visual Disturbance	0.004	0.004	0.017
Cranial Nerve Disorder	0.005	0.004	0.005
Lupus Headache	0.009	0.010	0.029
CVA	0.001	0.009	0.029
Vasculitis	0.013	0.015	0.026
Arthritis	0.098	0.079	0.217*
Myositis	0.006	0.006	0.040
Urinary Casts	0.001	0.001	0.005
Hematuria	0.048	0.049	0.034
Proteinuria	0.076	0.087	0.084*
Pyuria	0.029	0.027	0.019
New rash	0.089	0.089	0.023
Alopecia	0.072	0.060	0.051*
Mucosal Ulcers	0.043	0.041	0.018
Pleurisy	0.029	0.024	0.071*
Pericarditis	0.008	0.008	0.012
Low Complement	0.298	0.281	0.027
Increased DNA binding	0.283	0.266	0.044
Fever	0.036	0.022	0.016
Thrombocytopenia	0.067	0.054	0.212*
Leukopenia	0.006	0.008	0.079*

* Proportion of simulation runs considered extreme (statistically significantly different 5% significant level) in comparison with Hopkins Cohort

Figure 2: A histogram of SLEDAI scores from the Simulation and Hopkins cohort

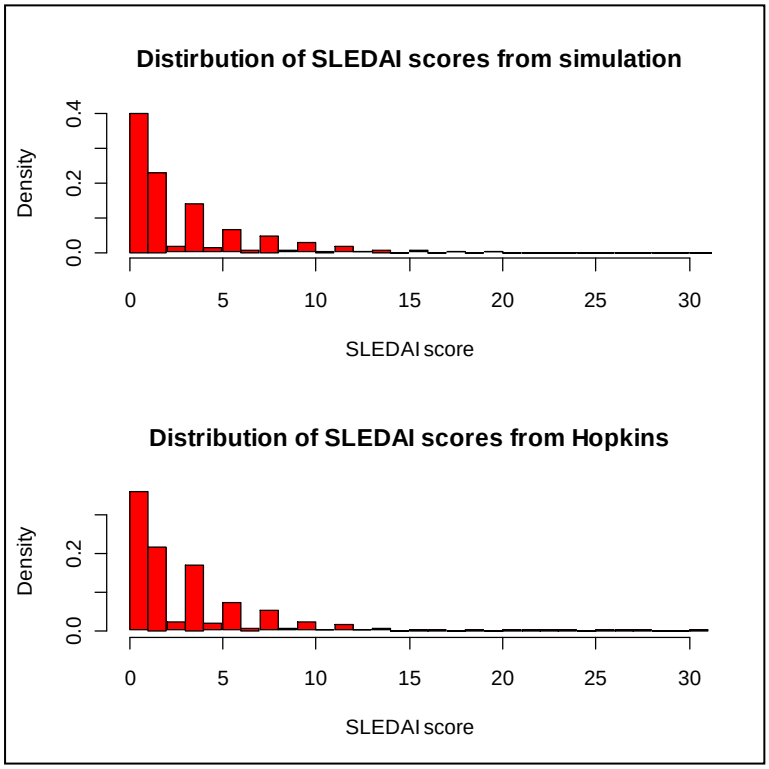

Figure 3: A histogram of Prednisone dose from the Simulation and Hopkins cohort

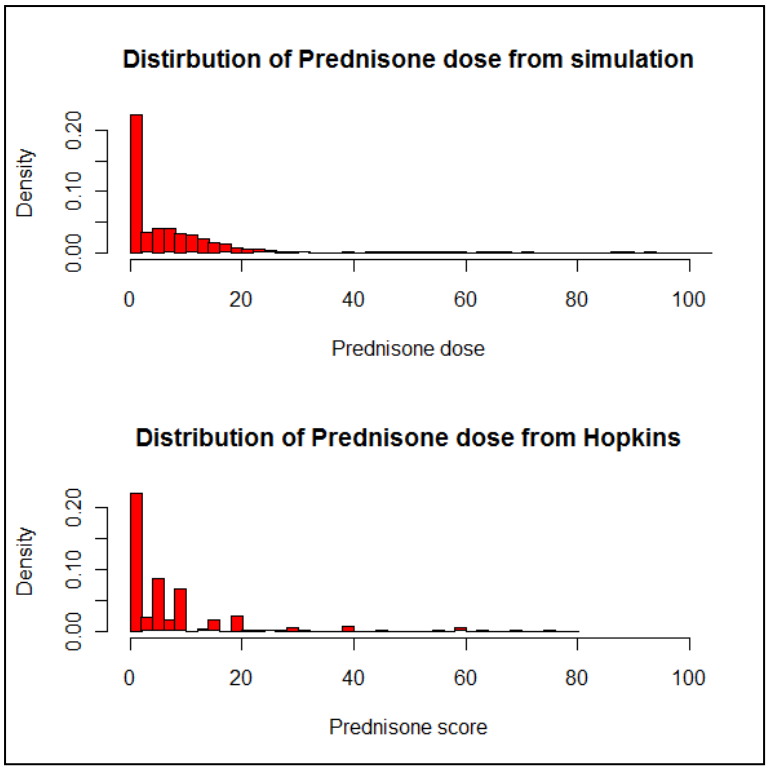

Figure 4: Distribution of percentage change in SLEDAI for simulated trials (red bars) and estimate distribution of Phase II results (Blue line) for placebo and treatment arm

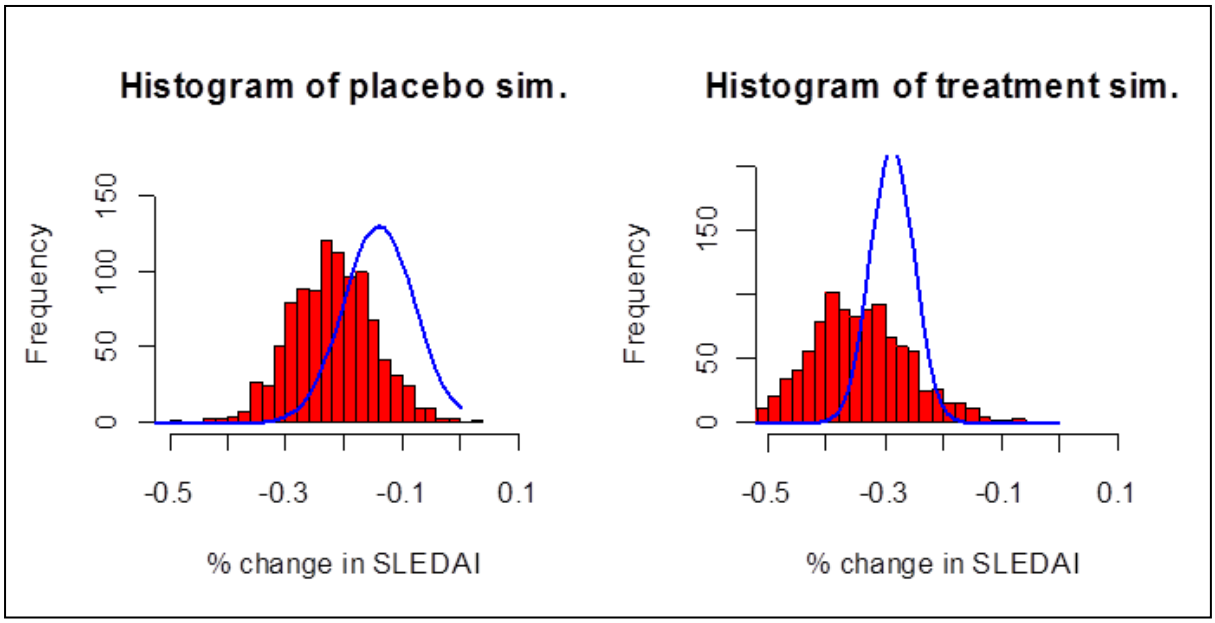

Figure 5: Distribution of proportion of patients with a ≥4 unit reduction in SLEDAI for simulated trials (red bars) and estimate distribution of Phase II results (Blue line) for placebo and treatment arm

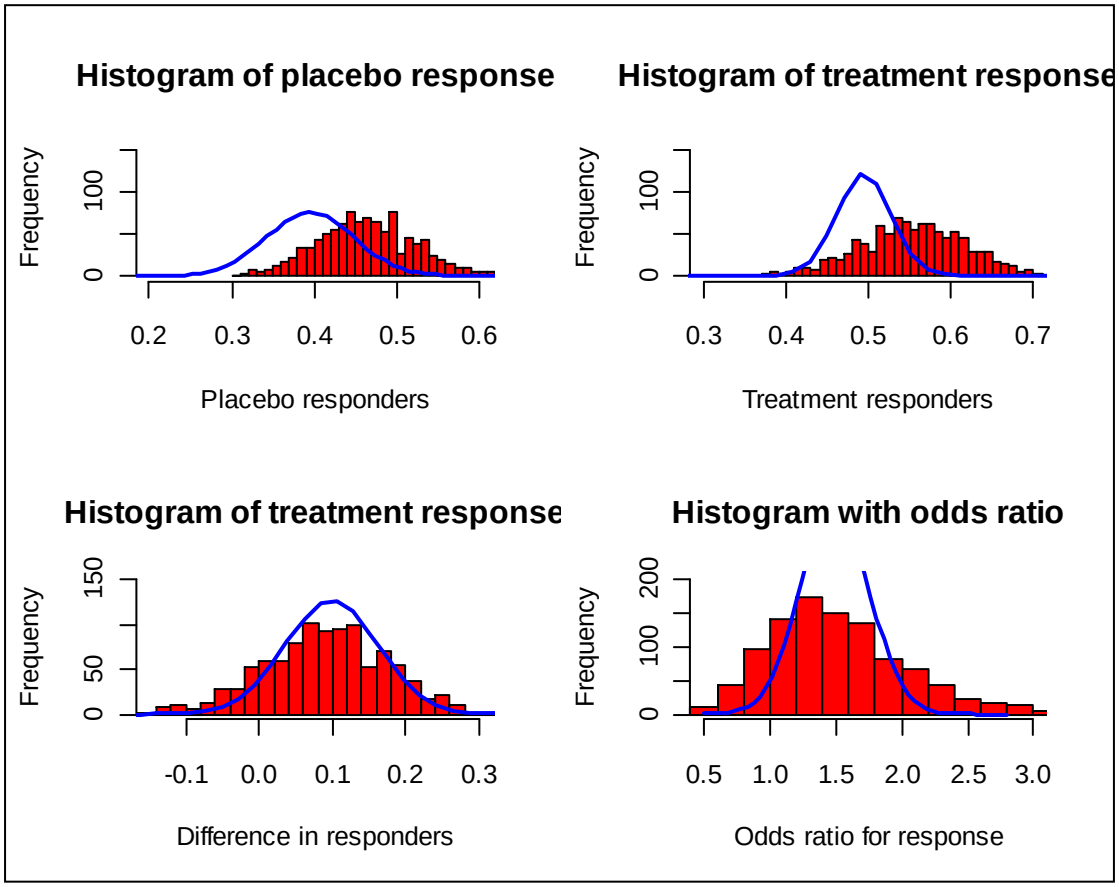

Figure 6: Distribution of proportion of patients with a ≥4 unit reduction in SLEDAI for simulated trials (red bars) and estimate distribution of Phase III results (Blue line) for placebo and treatment arm