**Disease Trajectory modeling of SLE clinical endpoints using a Latent Variable Model: Analysis of Pooled Patient-Level Placebo (Standard-of-Care) Data**

Kosalaram Goteti², Jonathan French⁴, Ramon Garcia⁵, Ying Li¹, Florence Casset-Semanaz¹, Aida Ayedmir¹, Robert Townsend¹, Cristina Vazquez Mateo¹, Matthew Studham¹, Oliver Gunther³, Amy Kao⁴, Marc Gastonguay³, Pascal Girard³, Lisa Benincosa¹, Karthik Venkatakrishnan¹

¹EMD Serono Research and Development Institute, Inc., Billerica, MA, USA (an affiliate of Merck KGaA, Darmstadt, Germany)
²Metrum Research Group, Tarryville, CT, US
³Merck Institute of Pharmacometrics, Lausanne, Switzerland
⁴Merck KGaA, Darmstadt, Germany

**INTRODUCTION**

- Systemic lupus erythematosus (SLE) is a multifactorial disease in a heterogeneous population.
- Response rates for current SOC are high (40%-50%) in an unselected patient population.
- High placebo+SoC (placebo added to standard-of-care therapy) response is likely one reason for many failed clinical trials in SLE and complicates the assessment of the actual treatment effect of a new investigational agent.
- Therefore, identification of patient characteristics predictive of poor response to current SOC is desirable to define enrichment strategies for efficient POC studies.
- Precision medicine strategies
- Increased probability of success

**Objective:** To understand the time course of clinical endpoints and identify clinically important covariates, we developed a mathematical framework to describe the SLE disease trajectory with important covariates.

**Methods:**

- A cross-functional team was formed for this effort to enable multi-disciplinary integration of disease area knowledge, pharmacometrics and statistical expertise.

**RESULTS**

- Several latent variable models, including linear, mono-exponential and Emax, were explored to define the SLE disease burden over time. Covariate screening was performed through LASSO and Random Forest (RF) models of the posterior median random effects; variable importance was assessed using multiple algorithms (Boruta, Permutation, Random-split, Shapley methods).
- A full covariate model was fit, integrating the learnings from the covariate screening with visual predictive checks.
- All modeling activities were performed in the Bayesian framework using R and Stan.

**CONCLUSION**

- The mono-exponential model best described the disease burden of the latent process over time.
- Machine Learning algorithms identified similar covariates in Stage II as in Stage I.
- Some of the baseline covariates that were identified to be important via LASSO and RF and confirmed through the LV model were SLEDAI total score, number of organ systems with a BILAG score of A or B, and low complement levels.
- The model developed using the 60% data set predicted both the 20% validation and 20% test datasets well for both the component and composite clinical endpoints.

**METHODS**

- The model was developed using the 60% data set predicted both the 20% validation and 20% test datasets well for both the component and composite clinical endpoints.

**RESULTS**

- The mono-exponential model best described the disease burden of the latent process over time.
- Machine Learning algorithms identified similar covariates in Stage II as in Stage I.
- Some of the baseline covariates that were identified to be important via LASSO and RF and confirmed through the LV model were SLEDAI total score, number of organ systems with a BILAG score of A or B, and low complement levels.
- The model developed using the 60% data set predicted both the 20% validation and 20% test datasets well for both the component and composite clinical endpoints.