Abstract 603 Table 3 Multivariable Models of the Impact of Disease Activity States Since Cohort Entry on Annual Direct and Indirect Costs

Model A		
	Annual Direct Costs, coefficient (95%CI)	**Annual Indirect Costs, coefficient (95%CI)**
Active state*	1161 (743, 1579)	3390 (1424, 5356)
Disease duration	333 (249, 417)	1346 (652, 2040)
White race/ethnicity	-2049 (-3356, -742)	-
Residing outside North America	-	-13657 (-19202, -8112)
Model B		
Remission Off-Treatment**	-1296 (-1800, -792)	-3353 (-5382, -1323)
Remission On-Treatment	-987 (-1550, -424)	-3508 (-5761, -1256)
LDA-TC	-1037 (-1853, -222)	-3229 (-5681, -778)
mLLDAS	-1307 (-2194, -420)	-3822 (-6309, -1334)
Disease duration	330 (245, 415)	1353 (662, 2044)
White race/ethnicity	-1996 (-3319, -674)	-
Residing outside North America	-	-13569 (-19040, -8097)
Difference between disease activity state coefficients (95%CI)		
Remission On vs Remission Off-Treatment	309 (-304, 921)	-156 (-1680, 1369)
LDA-TC vs Remission Off-Treatment	259 (-660, 1117)	123 (-1812, 2058)
LDA-TC vs Remission On-Treatment	-50 (-924, 824)	279 (-1400, 1959)
mLLDAS vs Remission Off-Treatment	-11 (-902, 881)	-469 (-2259, 1321)
mLLDAS vs Remission On-Treatment	-320 (-1255, 616)	-313 (-2741, 2115)
mLLDAS vs LDA-TC	-270 (-1365, 826)	-592 (-3056, 1872)

*Reference group for active state in Model A is all other disease activity states
** Reference group for all disease activity states in Model B is active state
LDA-TC: Low disease activity – Toronto Cohort; mLLDAS: modified Lupus Low Disease Activity State

$1550, -$424; IC -$3508, 95%CI -$5761, -$1256), LDA-TC (DC -$1037, 95%CI -$1853, -$222; IC -$3229, 95%CI -$5681, -$778) and mLLDAS (DC -$1307, 95%CI -$2194, -$420; IC - $3822, 95%CI -$6309, $-1334) (table 3, Model B). There were no differences in costs between remission and LDA.

Conclusions Remission and LDA are associated with lower costs, likely mediated through the known association of these DAS with more favourable clinical outcomes.

604 PREDICTING ADVERSE PREGNANCY OUTCOMES IN WOMEN WITH SYSTEMIC LUPUS ERYTHEMATOSUS: EXTERNAL VALIDATION OF THE PROMISSE MODEL USING MULTIPLE INDEPENDENT COHORTS

[1]Melissa Fazzari, [2]Marta Guerra, [3]Marta Mosca, [3]Dina Zucchi, [4]Jill Buyon, [5]Anna Brode, [2,6]Jane Salmon, [1]Mimi Kim*. *[1]Albert Einstein College of Medicine, New York, USA; [2]Hospital for Special Surgery, New York, USA; [3]University of Pisa, Pisa, Italy; [4]New York University School of Medicine, New York, USA; [5]Hackensack Meridian School of Medicine, New Jersey, USA; [6]Weill Cornell Medical College, New York, USA*

10.1136/lupus-2022-lupus21century.25

Background Nearly 20% of pregnancies in patients with Systemic lupus erythematosus (SLE) result in an adverse pregnancy outcome (APO); early identification of women with SLE who are at high risk of APO is vital. We previously examined several regression and machine learning (ML) predictive models for APO using data from the PROMISSE Study, a large multi-center, multi-ethnic/racial study of APO in women with mild/moderate SLE and/or aPL. Penalized logistic regression (LASSO), as well as several "black box" ML algorithms (Random Forest, Support Vector Machine, and Super Learner) each achieved good internal cross-validated performance, with area under the receiver operating curve (AUC) of 0.77-0.78. The goal of this study was to externally validate the performance of these promising APO risk models using three independent, external cohorts.

Methods The PROMISSE data set used to develop the initial APO prediction models consisted of N=385 pregnancies, 71 APO events (18.4%), and 32 known or potential APO risk factors that are routinely measured in clinical practice early in pregnancy. APO was defined as preterm delivery due to placental insufficiency or preeclampsia, fetal or neonatal death, or fetal growth restriction. Three independent prospective cohorts were provided by a team of international investigators with expertise in SLE pregnancy (Bronx, NY: N=96; NYC, NY: N=62; Pisa, Italy: N=152). Patient demographics were summarized for each cohort and missing data handled using multiple imputation with chained equations. Using the APO risk models developed with the PROMISSE data, we computed for each cohort: 1) the standard deviation (SD) of predicted risk scores to summarize the degree of heterogeneity in patient characteristics and 2) the area under the receiver operating curve (AUC) to summarize the ability of each model to discriminate patients with and without APO.

Results The three external cohorts and the PROMISSE development cohort showed distributional differences in previously identified APO risk factors (table 1). Non-Hispanic White comprised 49.3% of the PROMISSE, compared to 98.7% in Pisa, 27.4% in NYC, and 0% in the Bronx. LAC positivity varied from 8.1% in PROMISSE to 22.6% in the NYC cohort, while PGA > 1 varied from 10.6% in the development cohort to 4.4% in the Bronx, NY cohort. Current antihypertensive use was 8.6% in PROMISSE, higher in the Bronx cohort (12.6%), and lower in the NYC (4.8%) and Pisa (5.3%) cohorts. APO rates were the same in PROMISSE and Pisa (18.4%) and higher in the Bronx (24%) and NYC cohorts (25.8%). Prediction risk score SD indicated similar levels of heterogeneity within each external cohort compared to the PROMISSE cohort. Model performance in external validation cohorts varied depending on the algorithm used. As expected, AUCs in the external cohorts were generally lower than cross-validated internal estimates, but still indicated satisfactory performance of the different models with the independent data sets (table 2). Super Learner, the highest performing algorithm in PROMISSE, performed well across all three external cohorts, with a minimum AUC of 0.63 in the NYC cohort and a maximum of 0.71 in the Pisa cohort (table 2). LASSO also maintained consistent external performance with minimum AUC of 0.60 and maximum of 0.66. Overall, performance was highest using data from the Pisa cohort, which was the largest and most complete of the three external validation data sets.

Conclusions Penalized regression and ML approaches using variables obtained early in pregnancy show potential in discriminating pregnancies with high APO risk from those pregnancies with lower risk. This study provides confirmation of the geographic transportability of the best performing algorithms developed with PROMISSE. While Super Learner

showed the most satisfactory performance across external cohorts, LASSO also performed well and yielded a parsimonious model that may be easier and more efficient to use as a risk assessment tool in practice. Data from additional external cohorts from the US and abroad will be obtained in the future for further validation and refinement of our APO prediction models.

Acknowledgments This work was supported by NIH grant R21 AR076612

Trial Registration ClinicalTrials. gov Identifier: NCT00198068

Lay summary Nearly 20% of pregnancies in patients with Systemic lupus erythematosus (SLE) result in an adverse pregnancy outcome (APO); early identification of women with SLE who are at high risk of APO is vital. We previously explored several regression and machine learning methods to predict APO using data from the PROMISSE Study, a large multi-center, multi-ethnic/racial study of APO in women with mild/moderate SLE and/or aPL. We sought to determine which of the best performing algorithms in PROMISSE continued to perform well using data from other SLE pregnancy cohorts in the US and abroad. Most models showed satisfactory performance across cohorts in the ability to differentiate patients who did and not have an APO using variables measured early in pregnancy, indicating their potential for use in clinical practice to manage pregnant SLE patients.

Abstract 604 Table 1 Patient demographics by SLE Pregnancy Cohort

	Development	External validation data sets		
	PROMISSE	Bronx, NY	NYC, NY	Pisa, Italy
	n=385, 71 APOs	n=96, 23 APOs	n=62, 16 APOs	n=152, 28 APOs
	event rate=18.4%	event rate = 24.0%	event rate = 25.8%	event rate=18.4%
Maternal age , years	31 (28,34)	29 (24,34)	33 (29,33.5)	32 (28,36)
Non-Hispanic White (%)	49.3	0.0	27.4	98.7
Platelet count, x 10^9 cells/L	243 (204,296)	235 (209,259)	228 (188,274)	204 (188, 238)
Diastolic BP, mmHg	67 (60,73)	72.5 (63,79)	70 (64,75)	70 (61,75)
LAC + (%)	8.1	19.1	22.6	15.9
PGA > 1 (%)	10.6	4.4	11.1	7.9
SLE disease activity score	2 (0,4)	1.5 (0,2)	2 (0,5)	2 (0,4)
Low C3 (%)	20.2	27.6	24.2	52.6
aCL IgG + (%)	6.1	2.1	9.7	11.8
aCL IgM + (%)	1.8	0.0	16.1	1.3
Current glucocorticoid use (%)	39.7	50.5	23.1	58.9
Current anti hypertensives use (%)	8.6	12.6	4.8	5.3
Current hydroxychloroquine use (%)	64.7	54.3	84.6	63.4

Data are summarized as median (IQR), unless otherwise indicated; BP=blood pressure

Abstract 604 Table 2 AUC (95% CI) of all algorithms based on internal and external assessments

	Development	External Validation data sets		
	PROMISSE	Bronx, NY	NYC, NY	Pisa, Italy
LASSO	0.77 (0.71,0.83)*	0.60 (0.46,0.73)	0.63 (0.47,0.80)	0.66 (0.53,0.79)
Support vector machine	0.77 (0.70,0.84)*	0.61 (0.47, 0.74)	0.58 (0.41, 0.74)	0.73 (0.63,0.83)
Random Forest	0.77 (0.71,0.83)*	0.68 (0.55,0.81)	0.57 (0.46,0.80)	0.67 (0.56, 0.79)
Super Learner	0.78 (0.72,0.84)*	0.66 (0.53,0.79)	0.63 (0.43,0.76)	0.71 (0.56, 0.81)

*Based on 5x10-fold cross-validation

605 THE SYSTEMIC LUPUS ERYTHEMATOSUS INTERNATIONAL COLLABORATING CLINICS (SLICC), AMERICAN COLLEGE OF RHEUMATOLOGY (ACR), AND LUPUS FOUNDATION OF AMERICA (LFA) DAMAGE INDEX REVISION – ITEM GENERATION PHASE

[1]Burak Kundakci, [2]Ann E Clarke, [3]Sindhu R Johnson, [4]Hermine I Brunner, [5]Jiacai Cho, [6]Nathalie Costedoat-Chalumeau, [7]Ellen M Ginzler, [8]John G Hanly, [7]Abida Hasan, [9]Murat Inanç, [7]Naureen Kabani, [10]Kaitlin Lima, [11]Livia Lindoso, [12]Anselm Mak, [10]Rosalind Ramsey- Goldman, [13]Guillermo Ruiz-Irastorza, [11]Clovis A Silva, [14]Farah Tamirou, [11]Vitor C Trindade, [15]Évelyne Vinet, [1]Ian N Bruce, [2]Megan RW Barber*. *[1]Centre for Epidemiology Versus Arthritis, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, Manchester, UK; [2]Division of Rheumatology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; [3]Division of Rheumatology, Department of Medicine, Schroeder Arthritis Institute, Krembil Research Institute, Toronto Western and Mount Sinai Hospitals; Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario Canada; [4]Division of Rheumatology, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Department of Pediatrics; Cincinnati, Ohio, USA; [5]National University Health System (NUHS), Singapore, Singapore; [6]Cochin Hospital, Internal Medicine Department, Centre de référence maladies auto-immunes et systémiques rares d'île de France, Paris, France; [7]SUNY Downstate Health Sciences University, Department of Medicine, Brooklyn, NY, USA; [8]Division of Rheumatology, Queen Elizabeth II Health Sciences Center (Nova Scotia Rehabilitation Site) and Dalhousie University, Halifax, Nova Scotia, Canada; [9]Istanbul University Faculty of Medicine, Istanbul, Turkey; [10]Northwestern University Feinberg School of Medicine, Chicago, USA; [11]Faculdade de Medicina da Universidade de São Paulo (FMUSP), Brazil; [12]Division of Rheumatology, Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; [13]Autoimmune Diseases Research Unit, Biocruces Bizkaia Health Research Institute, Hospital Universitario Cruces, UPV/EHU, Barakaldo, País Vasco, Spain; [14]Rheumatology Department, Cliniques universitaires Saint- Luc, Brussels, Belgium; [15]McGill University Faculty of Medicine, Division of Rheumatology, Montreal, QC, Canada; *Designates presenting author*

10.1136/lupus-2022-lupus21century.26

Background The SLICC, ACR and LFA embarked on a data- and expert-driven project to develop a revised systemic lupus erythematosus (SLE) organ damage index (SDI). The methodological approach includes 5 phases: updating the construct of damage (I), item generation (II), item reduction (III), item weighting and threshold determination (IV), and the assessment of validation and reliability (V). In phase I, a consensus statement was developed to define the construct of damage in SLE[1]. In the Item Generation phase, we aimed to develop and agree on a candidate list of items that reflect the construct of damage in SLE and are appropriate to be included in a new damage index including consideration of relevant items from adult, paediatric and young adult SLE. In this analysis, we compare the two approaches to initial item generation that were employed in a parallel process, namely a literature review and a Delphi exercise.

Methods Item generation included a literature review and 3-part Delphi exercise. A group of lupus experts conducted a