An Electronic Health Record Based Algorithm for Predicting Systemic Lupus Erythematosus Flares: Integrating Clinical Factors and Social Determinants of Health

Yu Huang[1], **Lixia Yao**[2], zhengkang Fan[1], Jingchuan Guo[1] and Jiang Bian[1]

[1]University of Florida, Gainesville, FL, [2]Polygon Health Analytics LLC, Chalfont, PA

BACKGROUND

Systemic lupus erythematosus (SLE) follows a relapsing-remitting course, characterized by periodic disease activity flares. These flares, along with prolonged disease activity, lead to progressive multi-system organ damage, diminished quality of life, elevated healthcare costs, and increased mortality. Notably, significant disparities persist in healthcare access and outcomes among SLE patients, with a strong association with social determinants of health (SDoH). For instance, individuals from racial and ethnic minorities and socially disadvantaged backgrounds exhibit higher rates of SLE disease activities, experience more frequent flares, endure worse outcomes. Unfortunately, existing studies on SLE flares and disease activities have not yet incorporated considerations of SDoH and patients' social risks. In response, this study integrates patients' SDoH with their clinical characteristics, leveraging an electronic health record (EHR)-based machine learning algorithm to predict SLE flares.

This holistic approach supports real-time clinical decision-making, addresses social risks and SDoH, and aims to mitigate health disparities associated with SLE.

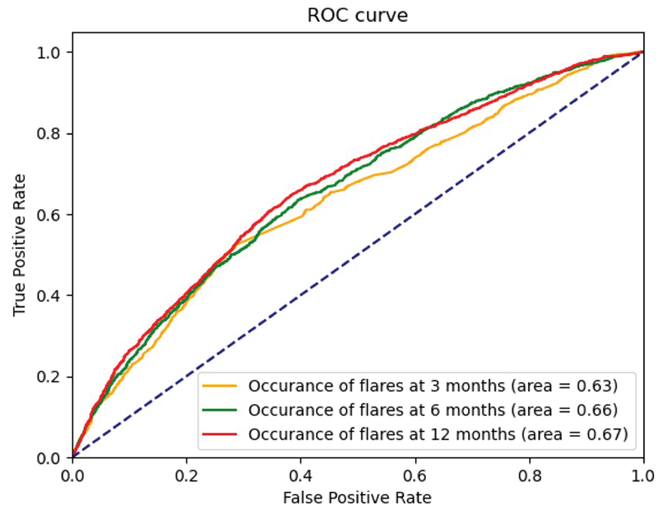
METHODS

We identified adult SLE patients using OneFlorida+ data from 2013 to 2021 and spatiotemporally linked the cohort with 486 neighborhood-level SDoH indicators (e.g., neighborhood unemployment rate and median income). The cohort included patients who met the following criteria: (1) were 18 years of age or older; (2) had recorded diagnoses of SLE (ICD-9 codes 710.0 or ICD-10 code M32.* and ≥ 1 rheumatology visit). The index date is decided based on a randomly selected encounter between the first recorded SLE diagnosis and the penultimate encounter; (3) had at least 1 year of available data before the index date; and (4) had at least 1 encounter every 12 months during the study period.

To quantify the occurrence of flare, we used a validated surrogate measure, SLE Disease Activity Score 2000 (SLEDAI-2K), extracted from EHR data for each 30-day period over the baseline year and follow-up year. The median score for the first 12 months established a baseline disease activity level, with a flare defined as an increase of at least 4 from the baseline score. We incorporated SDoH indicators and patients' clinical characteristics (e.g., medical diagnoses, prescriptions, and lab results) collected at baseline year and employed a gradient boosting machine to predict flare occurrence at 3 months, 6 months and 12 months after baseline.

Table 1: Demographic and SLEDAI-2K criteria attributes of the Study Cohort

	Total (n=33151)	
	median/n	IQR/percent
Age at first-ever recorded SLE diagnosis	45.51	(34.46, 57.11)
Sex		
Female	29650	89.44%
Male	3500	10.56%
Other	1	0.00%
SLEDAI-2K criteria attributes at follow-up		
Seizure	1821	5.49%
Psychosis	2764	8.34%
Organic brain syndrome	3882	11.71%
Visual disturbance	1222	3.69%
Cranial nerve disorder	225	0.68%
Lupus headache	5825	17.57%
CVA	1350	4.07%
Vasculitis	396	1.19%
Arthritis	1188	3.58%
Myositis	652	1.97%
Urinary casts	568	1.71%
Hematuria	1788	5.39%
Proteinuria	1355	4.09%
Pyuria	4178	12.60%
Rash	4435	13.38%
Alopecia	123	0.37%
Mucosal ulcers	775	2.34%
Pleurisy	210	0.63%
Pericarditis	407	1.23%
Low complement	258	0.78%
Increased DNA binding	319	0.96%
Fever	1883	5.68%
Thrombocytopenia	1672	5.04%
Leukopenia	2680	8.08%
SLEDAI-2K criteria attributes at baseline		
Seizure	2009	6.06%
Psychosis	3177	9.58%
Organic brain syndrome	4526	13.65%
Visual disturbance	1513	4.56%
Cranial nerve disorder	261	0.79%
Lupus headache	6763	20.40%
CVA	1566	4.72%
Vasculitis	558	1.68%
Arthritis	2093	6.31%
Myositis	973	2.94%
Urinary casts	667	2.01%
Hematuria	2126	6.41%
Proteinuria	1863	5.62%
Pyuria	4623	13.95%
Rash	5726	17.27%
Alopecia	190	0.57%
Mucosal ulcers	925	2.79%
Pleurisy	271	0.82%
Pericarditis	585	1.76%
Low complement	325	0.98%
Increased DNA binding	595	1.79%
Fever	2226	6.71%
Thrombocytopenia	1928	5.82%
Leukopenia	2844	8.58%
Outcome - flare occurrence (if ever, yes/no)	12815	38.66%
Outcome - flare occurrence (count)		
Count=1	6538	19.72%
Count=2	2902	8.75%
Count=3	1467	4.43%
Count>3	1908	5.76%



RESULTS

We identified 33,151 eligible SLE patients; the mean age was 45 (std: 16) years, 89% were females, and 39%, 26%, and 25% were non-Hispanic White, non-Hispanic Black, and Hispanic race/ethnicity, respectively. Of the cohort, over one-third (39%) developed at least one flare during the follow-up year; 19%, 9%, 4%, and 6% experienced 1, 2, 3, and 3+ time(s) flares during the follow up year. Our algorithm showed a good utility for predicting individuals' risk of flares, achieving a C statistic (standard deviation) of 0.63 (0.01), 0.66 (0.01), and 0.67 (0.01) at the 3- month, and 6- month, and 12-month period, respectively (Figure).

CONCLUSIONS

By incorporating both SDoH and clinical data, our EHR-based algorithm shows promise for identifying likely flares in real-world patients with SLE. Future studies are needed to identify key SDoH causally associated with flares occurrence to identify intervention targets that improve both health outcomes and health disparities in patients with SLE.

DISCLOSURES

Y. Huang: None; Z. Fan: None; J. Guo: None; J. Bian: None; L. Yao is an employee at Polygon Health Analytics LLC.