

## ORIGINAL RESEARCH

# Validation of a machine learning approach to estimate Systemic Lupus Erythematosus Disease Activity Index score categories and application in a real-world dataset

Pedro Alves,<sup>1</sup> Jigar Bandaria,<sup>1</sup> Michelle B Leavy ,<sup>2</sup> Benjamin Gliklich,<sup>3</sup> Costas Boussios,<sup>1</sup> Zhaohui Su,<sup>4</sup> Gary Curhan<sup>2</sup>

**To cite:** Alves P, Bandaria J, Leavy MB, *et al.* Validation of a machine learning approach to estimate Systemic Lupus Erythematosus Disease Activity Index score categories and application in a real-world dataset. *RMD Open* 2021;**7**:e001586. doi:10.1136/rmdopen-2021-001586

Received 27 January 2021  
Accepted 7 May 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Data Science, OM1 Inc, Boston, Massachusetts, USA

<sup>2</sup>Research, OM1 Inc, Boston, Massachusetts, USA

<sup>3</sup>Research, Noble and Greenough School, Dedham, Massachusetts, USA

<sup>4</sup>Biostatistics, OM1 Inc, Boston, Massachusetts, USA

**Correspondence to**  
Ms Michelle B Leavy;  
mleavy@om1.com

## ABSTRACT

**Objective** Use of the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) in routine clinical practice is inconsistent, and availability of clinician-recorded SLEDAI scores in real-world datasets is limited. This study aimed to validate a machine learning model to estimate SLEDAI score categories using clinical notes and to apply the model to a large, real-world dataset to generate estimated score categories for use in future research studies.

**Methods** A machine learning model was developed to estimate an individual patient's SLEDAI score category (no activity, mild activity, moderate activity or high/very high activity) for a specific encounter date using clinical notes. A training cohort of 3504 encounters and a separate validation cohort of 1576 encounters were created from the OM1 SLE Registry. Model performance was assessed using the area under the receiver operating characteristic curve (AUC), calculated using a binarised version of the outcome that sets the positive class to be those records with clinician-recorded SLEDAI scores >5 and the negative class to be records with scores ≤5. Model performance was evaluated by categorising the scores into the four disease activity categories and by calculating the Spearman's R value and Pearson's R value.

**Results** The AUC for the two categories was 0.93 for the development cohort and 0.91 for the validation cohort. The model had a Spearman's R value of 0.7 and a Pearson's R value of 0.7 when calculated using the four disease activity categories.

**Conclusion** The model performs well when estimating SLEDAI score categories using unstructured clinical notes.

## INTRODUCTION

Composite disease activity measures, such as the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI), are key outcome measures for assessing systemic lupus erythematosus (SLE) clinical status and response to treatment.<sup>1</sup> The SLEDAI has

## Key messages

### What is already known about this subject?

- ▶ Machine learning models may be used to estimate disease activity scores in rheumatological conditions, such as systemic lupus erythematosus and rheumatoid arthritis.

### What does this study add?

- ▶ A machine learning model performs well when estimating Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) score categories for patients with systemic lupus erythematosus (SLE) using unstructured clinical notes data.
- ▶ This study appears to be one of the first uses of machine learning to estimate SLE disease activity score categories from unstructured clinical notes.

### How might this impact on clinical practice or future developments?

- ▶ The model could be used to estimate SLEDAI score categories in real-world data sources, making these data more useful for research.
- ▶ The approach used to develop this model could be applied to disease activity measures for other rheumatological conditions.

been used extensively in clinical trials and, to some extent, in clinical practice. It has strong psychometric properties, is simple to score and can be used to identify patients in remission both on and off-therapy.<sup>2</sup> High disease activity as assessed by the SLEDAI is associated with more severe disease and damage.<sup>3</sup> Yet, despite its relative advantages, use of SLEDAI in clinical practice is inconsistent,<sup>4</sup> and availability of clinician-recorded SLEDAI scores in real-world datasets is limited.

Increasing the availability of SLEDAI scores in real-world datasets would yield new

opportunities for comparisons to trials data and research on patient treatment patterns and outcomes in a real-world setting. To date, statistical methods of imputation have been used to address missing disease activity scores in research studies,<sup>5</sup> but patients in real-world datasets may have no recorded disease activity scores, limiting the utility of this approach. Even for patients with some recorded disease activity level data, adding additional estimates at different timepoints would provide a more complete view of the patient's response to treatment and outcomes over time.

In recent years, significant progress has been made in classifying medical text and applying machine learning to the features produced. These techniques could be applied to the rich clinical data that exist in rheumatology clinical notes to estimate disease activity scores. This approach has been used in rheumatoid arthritis, where disease activity scores have been estimated from electronic medical records (EMR) data using machine learning, text features and laboratory values with good performance.<sup>6</sup> Because SLE disease activity is also measured using clinical and laboratory variables, it is reasonable to consider using the information contained in the clinical record to estimate disease activity at specific timepoints for patients with SLE to increase the value of real-world data for clinical and research purposes.

The objectives of this study were to validate a machine learning model to estimate SLEDAI score categories for patients with SLE using clinical notes and to apply the model to a large, real-world dataset to generate estimated SLEDAI score categories for use in future research studies.

## METHODS

### Participants

We used the OM1 SLE Registry, which includes over 41 000 patients with SLE from over 800 rheumatologists.<sup>7</sup> Practices are geographically distributed across the USA. Clinical data are drawn from EMRs and include medication history and prescription information, laboratory results and diagnoses as documented by a physician, primarily from outpatient or emergency room settings. Disease activity scores such as the SLEDAI are captured in the registry when they are documented in the patient record in structured form. Unstructured, physician-documented notes are available as well. The registry includes data from 2013 to 2020. To be eligible for inclusion in the registry, patients must be followed by a rheumatologist and have either multiple diagnosis codes for SLE or prescriptions for SLE treatments plus recorded SLE-specific patient-reported outcome measures. All data in the registry are deidentified. This study was submitted for Institutional Review Board approval and determined to be exempt.

Patients in the SLE Registry with at least one clinician-recorded SLEDAI and at least one text-based clinical note associated with the date of the clinician-recorded SLEDAI score were identified and randomly assigned to

either the model training cohort (69%) or the model validation cohort (31%).

### Modelling strategy

#### Dependent variable

The model is trained to estimate four SLEDAI score categories: no activity (SLEDAI score of 0), mild activity (SLEDAI scores of 1–5), moderate activity (SLEDAI scores of 6–10) and a combined high/very high activity (SLEDAI scores of  $\geq 11$ ). These categories align with published SLEDAI cutpoints for disease activity.<sup>8</sup> The trained model generates an estimated SLEDAI (eSLEDAI) score category for a specific encounter on a specific date. Because scores are estimated for a specific encounter on a specific date, an individual patient may have multiple timepoints for which they have eSLEDAI score categories and additional timepoints for which they have clinician-recorded SLEDAI scores.

#### Explanatory variables

We derived the explanatory features from clinical notes. In order to ensure that there was sufficient information to estimate the score category, analysed clinical notes were restricted to those that have components of medical history and physical examination. In the event that more than one clinical note exists on a single date, clinical notes from the same day were appended together as a single note for purposes of score category estimation. Explanatory variables that were used in the development of the estimation model were related to references of signs, symptoms, treatment and medications from the nine systems assessed by the SLEDAI (central nervous system, vascular, renal, musculoskeletal, serosal, dermal, immunologic, constitutional and haematologic).

### Modelling strategy

The SLEDAI estimation model used in this study was a multivariable ordinal regression model with the explanatory features generated from the clinical notes. Our modelling approach involved multiple steps: (1) processing of the body of clinical notes in order to de-noise and standardise their content; (2) identification of terms and phrases that are used by the physicians to indicate SLEDAI-related signs and symptoms to form the set of explanatory variables; and (3) model development and validation.

First, standard text processing steps such as word tokenisation (a process of separating a piece of text into smaller units), text lemmatisation (a process that reduces various inflectional and related forms of a word to a common root word) and removal of stop-words were followed (extremely common words that have little value for modelling).<sup>9 10</sup> The text processing was done in Python language using the NLTK package. N-grams features were generated from the processed text, and the features were TFIDF (Term frequency Inverse Document Frequency) transformed to generate the input for the model training. The text processing for feature

generation and model development was performed using the scikit-learn package in Python. The entire process of text processing, feature generation and model development was performed in a Python environment where the package version was fixed for reproducibility. An initial estimate of the feature coefficients was obtained by fitting the data to a logistic regression model.

Next, we evaluated the clinical notes to identify terms and phrases that have correlation to the SLEDAI score and are clinically important for the calculation of the SLEDAI score. The terms and phrases that we identified were based on the 24 variables used in calculation of SLEDAI score (eg, arthritis, rash, vasculitis, fever) as well as terms that indicate worsening, improvement or change in these variables. Terms and phrases that had a correlation with SLEDAI scores were used as predictive features in the model. This set of predictive features was used to train the model on the training cohort. In the final step, the model was validated on the validation cohort.

Model performance was assessed using the area under the receiver operating characteristic curve (AUC). The AUC was calculated using a binarised version of the outcome that sets the positive class to be those records with clinician-recorded SLEDAI scores greater than 5 (the threshold at which SLEDAI scores are considered to correspond to ‘mild’ vs ‘moderate’ disease activity), and the negative class to be records with scores less than or equal to 5. Model performance was also evaluated by categorising the scores into the four disease activity categories and calculating the Spearman’s R value and Pearson’s R value. Correlation coefficients were used to evaluate performance because there is ordinality between the four eSLEDAI score categories (ie, it is more accurate to categorise ‘no activity’ as ‘mild activity’ than to categorise ‘no activity’ as ‘high/very high activity’). Sensitivity and specificity for the four eSLEDAI categories were also calculated. The predictors were reviewed for clinical suitability for estimating a SLEDAI score category, and the distribution of eSLEDAI score categories was compared with the distribution of clinician-recorded SLEDAI scores.

### Application of the model to the SLE Registry

The trained model was applied to all SLE Registry patients who did not meet the eligibility criteria for the development and validation cohorts but for whom encounter notes with sufficient detail existed. Descriptive analyses were conducted to examine the demographic and clinical characteristics, including age, sex, race, medical history and treatment history, of the SLE Registry patients with clinician-recorded SLEDAI scores, with eSLEDAI score categories, and with both clinician-recorded scores and eSLEDAI score categories. Analyses were also conducted to examine the relation of clinician-recorded SLEDAI scores and eSLEDAI score categories to healthcare resource utilisation, pain medication prescriptions and corticosteroid prescriptions, as previous studies have shown that resource utilisation including medication

usage increases with increased disease activity.<sup>11 12</sup> While a comparison of the eSLEDAI score category to a clinician-recorded SLEDAI score cannot be done for these SLE Registry encounters, the performance of the model can still be assessed based on other types of medical data. In these analyses, we compared trends in healthcare resource utilisation, pain medication prescriptions and corticosteroid prescriptions across the four estimated SLEDAI categories to trends across clinician-recorded SLEDAI scores as independent measures of the model’s performance in a real-world dataset.

## RESULTS

### Participants and characteristics

The model training cohort consisted of 3504 encounters from 1130 patients with SLE, while the validation cohort consisted of 1576 encounters from 500 distinct patients. Demographics and clinical characteristics of patients in the training set are presented and discussed in [table 1](#).

### Model performance

The AUC was 0.93 for the development cohort and 0.91 for the validation cohort for the binary outcome ([figure 1](#)). Sensitivity and specificity for the four eSLEDAI score categories were calculated for the validation cohort and are presented in [table 2](#).

The model had a Spearman’s R value of 0.7 and a Pearson’s R value of 0.7 when evaluating the model performance by the four estimated SLEDAI score categories. The confusion matrix for the Spearman data is presented in [table 3](#).

Features that were used in the development of the model were related to signs and symptoms from the nine systems assessed by the SLEDAI. The top predictors were signs/symptoms that are specifically used in determining SLEDAI, plus additional terms that relate to the severity and/or frequency of these signs/symptoms.

### Estimation of SLEDAI score categories in the SLE Registry

The model was used to generate eSLEDAI score categories for 62263 encounters from 21393 SLE Registry patients that do not have clinician-recorded SLEDAI scores, but have clinical notes with sufficient information. The distribution of score categories for encounters with clinician-recorded SLEDAI scores in the validation cohort was similar to the distribution of score categories for encounters with eSLEDAI score categories in the SLE Registry ([figure 2](#)). Specifically, for clinician-recorded SLEDAI scores, 31.7% of clinical notes were categorised as no activity, 31.8% of notes as mild activity, 21.5% of notes as moderate activity and 15.1% of notes as high/very high activity. For eSLEDAI score categories, 32.1% of notes were classified as no activity, 26.6% of notes as mild activity, 23.1% of notes as moderate activity and 18.2% of notes as high/very high activity.

The SLE Registry contains 7017 encounters from 2286 patients (of 41000 total patients or 5.6%) that have

**Table 1** Demographic and clinical characteristics of SLE Registry patients

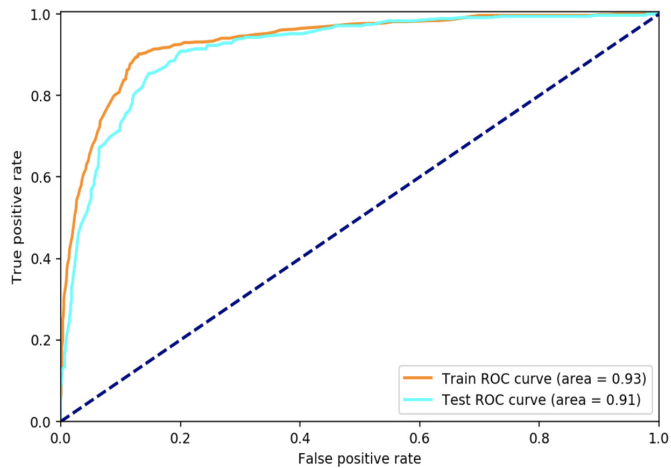
	Patients with clinician-recorded scores (N=2286)	Patients with estimated score categories (N=21 393)	Patients with both clinician-recorded and estimated score categories (N=1762)	Patients in model training cohort (N=1130)
Sex				
Female	2139 (93.6%)	19691 (92.0%)	1650 (93.6%)	1053 (93.2%)
Male	147 (6.4%)	1702 (8.0%)	112 (6.4%)	77 (6.8%)
Race				
Black	602 (26.3%)	4230 (19.8%)	466 (26.4%)	300 (26.5%)
White	1284 (56.2%)	12236 (57.2%)	1031 (58.5%)	645 (57.1%)
Other	60 (2.6%)	645 (3.0%)	50 (2.8%)	32 (2.8%)
Unknown	340 (14.9%)	4282 (20.0%)	215 (12.2%)	153 (13.5%)
Age, years	50.0 (14.5)	51.2 (15.0)	48.3 (14.1)	50.4 (14.5)
Duration of follow-up, years	6.5 (1.4)	6.4 (1.4)	6.6 (1.3)	6.7 (1.1)
Medical history (n, %)				
Constitutional manifestations	818 (35.8%)	6271 (29.3%)	530 (30.1%)	414 (36.6%)
Musculoskeletal manifestations	739 (32.3%)	5953 (27.8%)	518 (29.4%)	377 (33.4%)
Dermatologic manifestations	260 (11.4%)	1890 (8.8%)	128 (7.3%)	120 (10.6%)
Renal manifestations	375 (16.4%)	3291 (15.4%)	203 (11.5%)	191 (16.9%)
Neuropsychiatric manifestations	464 (20.3%)	3640 (17.0%)	355 (20.1%)	248 (21.9%)
Pulmonary manifestations	101 (4.4%)	849 (4.0%)	58 (3.3%)	45 (4.0%)
Gastrointestinal manifestations	40 (1.7%)	344 (1.6%)	24 (1.4%)	24 (2.1%)
Cardiac manifestations	40 (1.7%)	307 (1.4%)	18 (1.0%)	22 (1.9%)
Haematologic manifestations	212 (9.3%)	1540 (7.2%)	119 (6.8%)	98 (8.7%)

Continued

**Table 1** Continued

	Patients with clinician- recorded scores (N=2286)	Patients with estimated score categories (N=21 393)	Patients with both clinician- recorded and estimated score categories (N=1762)	Patients in model training cohort (N=1130)
Treatment history (n, %)				
Azathioprine	206 (9.0%)	1112 (5.2%)	117 (6.6%)	104 (9.2%)
Belimumab	389 (17.0%)	866 (4.0%)	177 (10.0%)	199 (17.6%)
Cyclophosphamide	4 (0.2%)	21 (0.1%)	0 (0.0%)	1 (0.1%)
Ciclosporin	2 (0.1%)	25 (0.1%)	3 (0.2%)	1 (0.1%)
Hydroxychloroquine	1455 (63.6%)	10939 (51.1%)	974 (55.3%)	713 (63.1%)
Methotrexate	227 (9.9%)	1638 (7.7%)	155 (8.8%)	121 (10.7%)
Mycophenolate	310 (13.6%)	1572 (7.3%)	166 (9.4%)	146 (12.9%)
Rituximab	70 (3.1%)	238 (1.1%)	27 (1.5%)	45 (4.0%)
Sirolimus	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Tacrolimus	4 (0.2%)	41 (0.2%)	6 (0.3%)	2 (0.2%)
Ustekinumab	0 (0.0%)	9 (0.0%)	0 (0.0%)	0 (0.0%)
Glucocorticoids	1296 (56.7%)	9191 (43.0%)	912 (51.8%)	661 (58.5%)
Prescription NSAIDs	570 (24.9%)	4300 (20.1%)	408 (23.2%)	282 (25.0%)

NSAIDs, non-steroidal anti-inflammatory drugs.



**Figure 1** AUC of SLEDAI prediction model. The AUC was calculated using a binarised version of the outcome that sets the positive class to be those records with clinician-recorded SLEDAI scores greater than 5 (the threshold at which SLEDAI scores are considered to correspond to ‘mild’ vs ‘moderate’ disease activity), and the negative class to be records with scores less than or equal to 5. AUC, area under the receiver operating characteristic curve; ROC, receiver operating characteristic; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index.

clinician-recorded SLEDAI scores from routine clinical rheumatology practice. Of the 2286 patients with encounters with clinician-recorded scores, 1762 had eSLEDAI score categories estimated for other encounters. Demographics and clinical characteristics of patients with clinician-recorded SLEDAI scores, patients with eSLEDAI score categories, patients with both clinician-recorded scores and eSLEDAI score categories and patients in the model training cohort are presented in [table 1](#). The majority (93.6% with clinician-recorded scores, 92.0% with eSLEDAI score categories, 93.6% with both scores and 93.2% in the model training cohort) were female, and the mean age as of the score was 49.9 years for patients with clinician-recorded scores, 51.2 years for patients with eSLEDAI score categories, 48.3 years for patients with both scores and 50.4 years for patients in the training cohort.

Healthcare resource utilisation, pain medication prescriptions and corticosteroid prescriptions were also examined for encounters with clinician-recorded scores and eSLEDAI score categories. Inpatient visits and emergency room (ER) visits that occurred within

a 12-month window ( $\pm 6$  months) of the SLEDAI score date are presented in [table 4](#). Encounters with clinician-recorded scores and those with estimated scores have similar trends in terms of healthcare resource utilisation. Pain medication prescriptions that occurred within a 90-day window ( $\pm 45$  days) of the SLEDAI score date are presented in [table 4](#). Pain medication prescriptions increased with higher score categories, and similar trends are observed among encounters with clinician-recorded scores and those with eSLEDAI score categories (30.2% of high/very high activity clinician-recorded scores and 23.8% of high/very high activity eSLEDAI score categories). Corticosteroid prescriptions that occurred within a 90-day window ( $\pm 45$  days) of the SLEDAI score date are presented in [figures 3 and 4](#). Corticosteroid prescriptions increased with higher score categories, and similar trends are observed among encounters with clinician-recorded scores and those with eSLEDAI score categories (36.8% of high/very high activity clinician-recorded scores and 29.6% of high/very high activity eSLEDAI score categories).

## DISCUSSION

While the SLEDAI is widely used in clinical trials, its use in routine clinical practice is inconsistent, and availability of clinician-recorded SLEDAI scores in real-world datasets is limited. The lack of SLEDAI scores in real-world datasets reduces the utility of these datasets for addressing questions about real-world treatment patterns and outcomes and for supporting new research, including identification of eligible patients for clinical trials. This study demonstrated that a machine learning model can be used to estimate eSLEDAI score categories. This machine learning model performs well when estimating the four SLEDAI score categories using unstructured clinical notes from a real-world dataset. The results align with objective data, such as corticosteroid prescriptions, pain medication prescriptions and healthcare resource utilisation, providing further confidence in the estimated score categories. The approach used to develop this model could be applied to other rheumatology disease activity scores, and future models may build on this effort by combining structured clinical data with unstructured notes to further improve performance.

This study appears to be one of the first uses of machine learning to estimate four SLE disease activity

**Table 2** Sensitivity and specificity for the four eSLEDAI categories

	Sensitivity	Specificity
No activity vs (mild activity, moderate activity, high/very high activity)	0.921	0.587
Mild activity vs (no activity, moderate activity, high/very high activity)	0.134	0.872
Moderate activity vs (no activity, mild activity, high/very high activity)	0.243	0.930
High/Very high activity vs (no activity, mild activity, moderate activity)	0.621	0.888
(No activity, mild activity) vs (moderate activity, high/very high activity)	0.678	0.924

eSLEDAI, estimated Systemic Lupus Erythematosus Disease Activity Index.

**Table 3** Confusion matrix for Spearman's rank correlation coefficient

	Predicted			
	No activity	Mild activity	Moderate activity	High/Very high activity
Actual				
No activity	280	18	5	1
Mild activity	224	41	23	17
Moderate activity	41	42	50	73
High/Very high activity	6	24	25	90

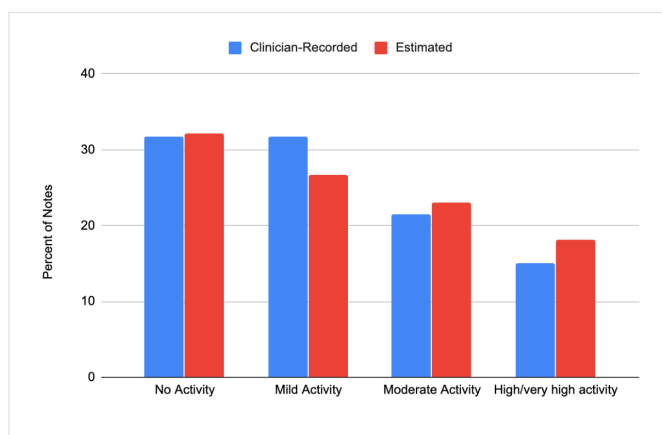
score categories from unstructured text. Other efforts have developed models to predict high disease activity in SLE,<sup>13</sup> to predict future disease activity scores in rheumatoid arthritis<sup>6</sup> and to estimate disease severity in rheumatoid arthritis using administrative claims data.<sup>14 15</sup> This effort is unique in that it focuses on SLE, estimates four disease activity score categories at a specific point in time as opposed to predicting future scores and uses clinical narrative data to assess disease activity. Assessing disease activity is a routine part of patient care for SLE, as disease activity is an essential factor in determining the appropriate intervention strategies and therapies to achieve disease remission. These automated estimates also provide a means to assist in tracking disease progression, improvement and remission over time and can generate insights into the effectiveness of interventions. Importantly, objective measures of disease activity allow for standardisation of patient outcome measures across clinicians and patients and facilitate care management, population health management and research.

Yet, calculation and recording of SLEDAI scores is challenging in routine clinical practice settings, leading to the need for new strategies to estimate these scores at the individual patient and population level. As shown in

table 1, only 5.6% of patients in a real-world setting had SLEDAI scores recorded in their EMR. While a clinician-recorded SLEDAI score will remain the gold standard, generating eSLEDAI score categories based on the clinical narrative dramatically increases the number of available endpoints for tracking and understanding patient outcomes when clinician-recorded SLEDAI scores are not available. For patients for whom there are multiple SLEDAI values available, statistical imputation could be considered as an alternative approach, but this was the case for a small minority of patients. Statistical methods for imputing missing data, like the multiple imputation method, usually assume missing at random which may not hold because patients with and without clinician-recorded SLEDAI values have different characteristics. In addition, the multiple imputation method is applicable only when the percent of missing data is less than 40%.<sup>16</sup>

Multiple applications exist for a machine learning model to estimate SLEDAI score categories. First, use of the model within real-world datasets would expand the population of patients that could be included in research studies that require disease activity scores. This would be particularly useful for building large cohorts of patients who receive treatment outside specialised SLE clinics where the SLEDAI score is rarely recorded.<sup>4</sup> Estimating score categories for these patients and including these patients in real-world studies could generate useful information to better correlate trial outcomes with real-world patient outcomes. Beyond research uses, generation of an estimated score category at the individual patient level could help clinicians track a patient's response to treatment over time and support care management. Consideration of the performance characteristics of the model, particularly AUC, as well positive predictive value and negative predictive value in appropriate use cases, will be critical.

This approach has some limitations. First, the model relies on data recorded in the clinical notes, and notes without sufficient text and clinical details are excluded. Estimation of the SLEDAI score categories requires that adequate time was spent with the patient during the visit and that the visit was documented with sufficient thoroughness. In addition, the model was trained and validated using EMR data from multiple rheumatology practices in the USA. Prior to use in other data sources, the model will need to be modified to address variations

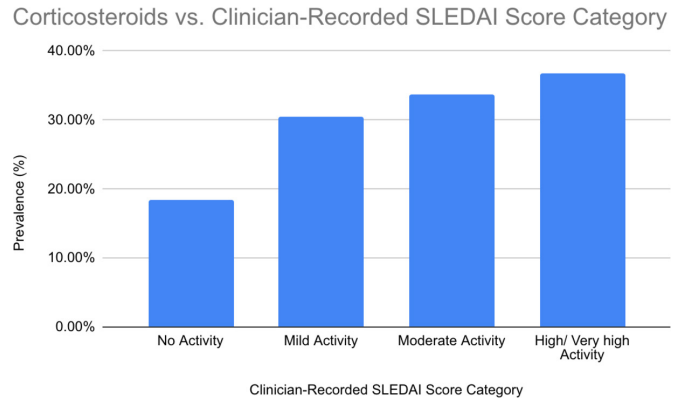


**Figure 2** Comparison of clinician-recorded SLEDAI score categories versus eSLEDAI score categories. The distribution of SLEDAI score categories for encounters with clinician-recorded SLEDAI scores in the validation cohort and the distribution of SLEDAI score categories for encounters with eSLEDAI score categories are shown here. eSLEDAI, estimated SLEDAI; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index.

**Table 4** Healthcare resource utilisation and pain medication prescriptions by SLEDAI score category for clinician-recorded scores and eSLEDAI score categories

	SLEDAI score category					Total (N=7017)
	No activity (N=1900)	Mild activity (N=2483)	Moderate activity (N=1752)	High/Very high activity (N=882)		
<i>Clinician-recorded SLEDAI</i>						
No of inpatient days	Mean (SD) 0.3 (1.7)	0.4 (3.1)	0.5 (3.1)	0.7 (5.0)	0.4 (3.1)	
	Median (Q1–Q3) 0 (0–0)	0 (0–0)	0 (0–0)	0 (0–0)	0 (0–0)	
Any pain medication prescriptions	Yes 279 (14.7%)	477 (19.2%)	498 (28.4%)	266 (30.2%)	1520 (21.7%)	
<i>eSLEDAI</i>						
No of inpatient days	Mean (SD) 0.5 (4.7)	0.8 (4.9)	0.7 (4.1)	0.7 (8.3)	0.6 (5.4)	
	Median (Q1–Q3) 0 (0–0)	0 (0–0)	0 (0–0)	0 (0–0)	0 (0–0)	
Any pain medication prescriptions	Yes 2433 (11.4%)	2600 (15.6%)	2598 (18.6%)	2451 (23.8%)	10082 (16.2%)	

Note: Inpatient days and ER visits are included if they occur within a 12-month window (±6 months) of the SLEDAI score date. Pain medication prescriptions are included if they occur within a 90-day window (±45 days) of the SLEDAI score date. ER, emergency room; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index.

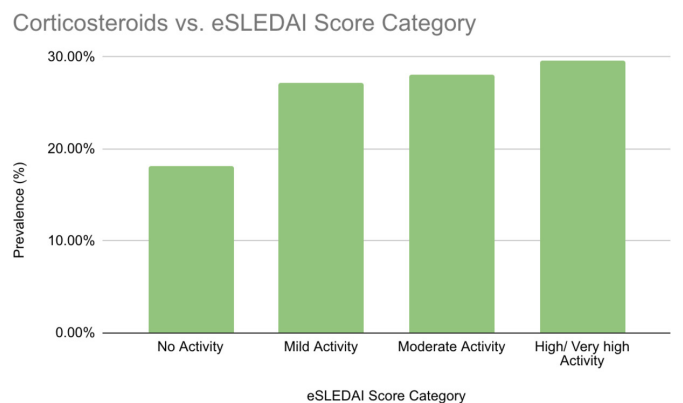


**Figure 3** Corticosteroid prescriptions by SLEDAI score category for clinician-recorded SLEDAI scores. Corticosteroid prescriptions that occurred within a 90-day window (±45 days) of the SLEDAI score date are presented here by SLEDAI score category for clinician-recorded scores. SLEDAI, Systemic Lupus Erythematosus Disease Activity Index.

in other care settings and documentation practices. Validity should be confirmed before applying the model in other data sources.

By developing a model to estimate SLEDAI score categories, this study addressed the lack of SLE disease activity scores in real-world data sources. Use of the model to estimate SLEDAI score categories could make real-world data sources more useful for research and potentially support patient care management in clinical practice.

**Contributors** PA—study concepts and design, manuscript and figure preparation, manuscript editing, final approval of the article; JB—study concepts and design, manuscript and figure preparation, manuscript editing, final approval of the article; ML—study concepts and design, manuscript and figure preparation, manuscript editing, final approval of the article; BG—study concepts and design, manuscript and figure preparation, manuscript editing, final approval of the article; CB—study concepts and design, manuscript and figure preparation, manuscript editing, final approval of the article; ZS—study concepts and design, manuscript and figure preparation, manuscript editing, final approval of the article; GC—study concepts



**Figure 4** Corticosteroid prescriptions by SLEDAI score category for eSLEDAI score categories. Corticosteroid prescriptions that occurred within a 90-day window (±45 days) of the estimated score category date are presented here by eSLEDAI score categories. eSLEDAI, estimated SLEDAI; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index.



and design, manuscript and figure preparation, manuscript editing, final approval of the article.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** The authors indicated are employees of OM1, which is involved in issues related to the topic of this manuscript.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** No data are available. This study was conducted using deidentified participant data compiled from multiple sources. Restrictions apply to the availability of these data, which were used under certain permissions for this study. Please contact the corresponding author (Michelle Leavy, ORCID ID 0000-0003-1927-7248) with any questions.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

**ORCID iD**

Michelle B Leavy <http://orcid.org/0000-0003-1927-7248>

**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:S37–46.
- 2 van Vollenhoven R, Voskuyl A, Bertsias G, et al. A framework for remission in SLE: consensus findings from a large international Task force on definitions of remission in SLE (DORIS). *Ann Rheum Dis* 2017;76:554–61.

- 3 Koelmeyer R, Nim HT, Nikpour M, et al. High disease activity status suggests more severe disease and damage accrual in systemic lupus erythematosus. *Lupus Sci Med* 2020;7:e000372.
- 4 Castrejón I, Tani C, Jolly M, et al. Indices to assess patients with systemic lupus erythematosus in clinical trials, long-term observational studies, and clinical care. *Clin Exp Rheumatol* 2014;32:S-85–95.
- 5 Mongin D, Lauper K, Turesson C, et al. Imputing missing data of function and disease activity in rheumatoid arthritis registers: what is the best technique? *RMD Open* 2019;5:e000994.
- 6 Norgeot B, Glicksberg BS, Trupin L, et al. Assessment of a deep learning model based on electronic health record data to forecast clinical outcomes in patients with rheumatoid arthritis. *JAMA Netw Open* 2019;2:e190606.
- 7 SLE Registry. OM1, 2021. Available: <https://www.om1.com/solutions/real-world-evidence/clinical-registries/immunology-and-ra-2/>
- 8 Abrahamowicz M, Fortin PR, du Berger R, et al. The relationship between disease activity and expert physician's decision to start major treatment in active systemic lupus erythematosus: a decision aid for development of entry criteria for clinical trials. *J Rheumatol* 1998;25:277–84.
- 9 Rajaraman A, Ullman JD. Data mining. In: Rajaraman A, Ullman JD, eds. *Mining of massive datasets*. Cambridge: Cambridge University Press, 2011: 1–17.
- 10 Manning CD, Raghavan P, Schütze H. *Introduction to information retrieval*. Cambridge: Cambridge University Press, 2008.
- 11 Kan H, Guerin A, Kaminsky MS, et al. A longitudinal analysis of costs associated with change in disease activity in systemic lupus erythematosus. *J Med Econ* 2013;16:793–800.
- 12 Garris C, Jhingran P, Bass D, et al. Healthcare utilization and cost of systemic lupus erythematosus in a US managed care health plan. *J Med Econ* 2013;16:667–77.
- 13 Hoi A, Nim HT, Koelmeyer R, et al. Algorithm for calculating high disease activity in SLE. *Rheumatology* 2021;7.
- 14 Chandran U, Reps J, Stang PE, et al. Inferring disease severity in rheumatoid arthritis using predictive modeling in administrative claims databases. *PLoS One* 2019;14:e0226255.
- 15 Speyer CB, Li D, Guan H, et al. Comparison of an administrative algorithm for SLE disease severity to clinical SLE disease activity index scores. *Rheumatol Int* 2020;40:257–61.
- 16 Jakobsen JC, Gluud C, Wetterslev J, et al. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Med Res Methodol* 2017;17:162.