


Evaluation of structured data from electronic health records to identify clinical classification criteria attributes for systemic lupus erythematosus

Theresa L Walunas ,^{1,2} Anika S Ghosh,² Jennifer A Pacheco,³ Vesna Mitrovic,² Andy Wu,² Kathryn L Jackson,² Ryan Schusler,⁴ Anh Chung,⁴ Daniel Erickson,⁴ Karen Mancera-Cuevas,⁴ Yuan Luo,⁵ Abel N Kho,^{1,2} Rosalind Ramsey-Goldman⁴

To cite: Walunas TL, Ghosh AS, Pacheco JA, *et al.* Evaluation of structured data from electronic health records to identify clinical classification criteria attributes for systemic lupus erythematosus. *Lupus Science & Medicine* 2021;**8**:e000488. doi:10.1136/lupus-2021-000488

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/lupus-2021-000488>).

Received 2 February 2021
Revised 12 April 2021
Accepted 13 April 2021

ABSTRACT

Objective Our objective was to develop algorithms to identify lupus clinical classification criteria attributes using structured data found in the electronic health record (EHR) and determine whether they could be used to describe a cohort of people with lupus and discriminate them from a defined healthy control cohort.

Methods We created gold standard lupus and healthy patient cohorts that were fully adjudicated for the American College of Rheumatology (ACR), Systemic Lupus International Collaborating Clinics (SLICC) and European League Against Rheumatism/ACR (EULAR/ACR) classification criteria and had matched EHR data. We implemented rule-based algorithms using structured data within the EHR system for each attribute of the three classification criteria. Individual criteria attribute and classification criteria algorithms as a whole were assessed over our combined cohorts and the overall performance of the algorithms was measured through sensitivity and specificity.

Results Individual classification criteria attributes had a wide range of sensitivities, 7% (oral ulcers) to 97% (haematological disorders) and specificities, 56% (haematological disorders) to 98% (photosensitivity), but all could be identified in EHR data. In general, algorithms based on laboratory results performed better than those primarily based on diagnosis codes. All three classification criteria systems effectively distinguished members of our case and control cohorts, but the SLICC criteria-based algorithm had the highest overall performance (76% sensitivity, 99% specificity).

Conclusions It is possible to characterise disease manifestations in people with lupus using classification criteria-based algorithms that assess structured EHR data. These algorithms may reduce chart review burden and are a foundation for identifying subpopulations of patients with lupus based on disease presentation to support precision medicine applications.

INTRODUCTION

SLE is a complex systemic autoimmune disease with a broad array of clinical and laboratory manifestations that make

Key messages

What is already known about this subject?

► Currently, rule-based algorithms exist for identifying people with lupus in claims and electronic health record data. These algorithms are based on diagnosis codes and laboratory results.

What does this study add?

► This study demonstrates that classification criteria attributes for lupus, particularly those determined with laboratory-based tests, can be identified in medical record data using rule-based algorithms that rely on structured data. In addition, it was possible to distinguish people in a well-characterised lupus cohort from people in a well-characterised healthy cohort using the classification criteria rules for defining 'definite lupus'.

How might this impact on clinical practice or future developments?

► These tools may be effective for describing the presentation of lupus using medical record data in the absence of manual chart review and support identification of patients for clinical trials, subpopulation analyses and population health management.

it challenging to diagnose and treat.^{1 2} Delayed identification can have profound impacts on people with SLE, and previous studies have shown that existing disease-related damage is one of the critical variables predicting long-term disease-related damage and severity of disease.³ Disease heterogeneity, in addition to making SLE identification difficult, also presents a challenge for using precision medicine strategies to develop therapeutic regimens and to ensure that the right patient gets the right care at the right time. To improve care for persons with SLE, it is critical to understand how the disease presents and identify subpopulations of patients with similar



© 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.

For numbered affiliations see end of article.

Correspondence to

Dr Theresa L Walunas; t-walunas@northwestern.edu

attributes so that clinicians can develop individualised approaches to manage disease with support of readily available tools such as the electronic health record (EHR).

Clinical classification criteria for SLE represent an evidence-based set of disease manifestations for people with SLE developed by clinical experts to describe the presentation of the disease for research applications. Currently, there are three validated classification criteria in use by the rheumatology community. First is the American College of Rheumatology (ACR) criteria initially developed in 1982 and enhanced in 1997.^{4 5} The Systemic Lupus International Collaborating Clinics (SLICC) criteria, developed in 2012, expanded the use of individual laboratory results to detect and describe autoimmune phenomena.⁶ Finally, the combined European League Against Rheumatism and ACR (EULAR/ACR) classification criteria effort, published in 2019, focused on developing a criteria set with high sensitivity and specificity that included disease attributes common in early-onset lupus.^{7–12} Historically, assessment of these classification criteria includes manual chart review by expert clinicians to define disease in patients participating in research studies. As of yet, they have not been adapted to mine EHR data, which could reduce chart review burden as well as provide a foundation for clinical and research applications.

Adoption of EHR systems in the USA was incentivised in 2009¹³ and as of 2019, 95% of hospitals and 80% of ambulatory practices are estimated to use EHR systems to document clinical care.^{14 15} For patients with long-term chronic conditions, such as SLE, the EHR is a rich longitudinal data source describing care, procedures provided, diagnoses identified, medications prescribed and laboratory test results. These data are the foundation for assessing whether clinical classification criteria attributes can be identified in the EHR to characterise people with SLE, and whether people with and without lupus can be distinguished using classification criteria.

While algorithms have been published that identify people with lupus in claims¹⁶ and EHR data based on diagnosis codes and laboratory results,^{17 18} to date, no algorithms exist that can help describe people with SLE and determine whether they satisfy classification criteria in the absence of manual chart review. We sought to determine whether it was possible to build rule-based algorithms for the ACR, SLICC and EULAR/ACR classification criteria for SLE using structured data easily identified in the EHR. Using a cohort of patients with lupus and a cohort of healthy patients from a general medicine population, both of which had manually adjudicated classification criteria attributes and in-depth medical records, we first determined whether we could identify individual attributes of the classification criteria in structured EHR data and then characterised which attributes were more difficult to identify correctly. We then assessed each full classification criteria algorithm over our cohorts

to determine whether they could distinguish between the lupus and healthy cohorts based on satisfying the overall definition of ‘definite lupus’ as defined for each classification criteria set.

METHODS

Development of rule-based algorithms for SLE classification criteria

We created three rule sets to identify the attributes of three published classification criteria: the ACR 1982/1997 classification criteria,^{4 5} the SLICC classification criteria⁶ and the joint EULAR/ACR classification criteria.^{7–10} For each attribute of each classification criteria, we developed a rule-based algorithm to detect the attribute using structured EHR data. For clinical attributes (such as alopecia or neurological conditions), we used diagnosis codes (International Classification of Diseases, Ninth and Tenth Revisions; ICD-9/10). For laboratory result attributes (such as ANA tests or complement levels), we used laboratory test results. Some attributes required additional terminologies to define, including procedure codes (current procedural terminology) and medication orders. [Table 1](#) provides data types underlying each attribute. Detailed information about underlying terminology codes for each attribute is provided in online supplemental table 1 and in online supplemental file 2 for attributes (arthritis, oral ulcers, serositis and ANA) had the same definition for all criteria. Nine attributes (acute cutaneous lupus, chronic cutaneous lupus, alopecia, leucopenia, haemolytic anaemia, thrombocytopenia, anti-dsDNA antibodies, antiphospholipid antibodies, anti-Smith antibodies) were common to SLICC and EULAR/ACR and the remaining 17 attributes were unique components of one classification criteria. Notably, renal disease and neurological disease are present in all three criteria but have different definitions and are represented as individual attributes.

Data sources

The Chicago Lupus Database (CLD), established in 1991, is a rheumatologist adjudicated (RRG) registry of 1052 patients with possible or definite lupus according to the revised 1982/1997 ACR classification criteria.^{4 5} The CLD has laboratory data, symptoms and patient demographics based on each known visit. If a patient was referred, history information from the notes is documented. Patients in the CLD have consented to research use of their medical records. The Northwestern Medicine Electronic Data Warehouse (NMEDW) is the primary data repository for all EHRs of patients who receive care within Northwestern Medicine (NM). Established in 2007, the NMEDW contains records for over 6.6 million patients.

Gold standard SLE cohort

To create our gold standard SLE cohort we identified patients in the CLD who also had medical records

Table 1 Individual classification criterion as described by domain, classification criteria they belong to and underlying medical record data types used to identify each criterion

Criterion	Classification criteria domain	Lupus classification criteria			Underlying medical record data types
		ACR 1997	SLICC 2012	EULAR/ACR 2019	
Acute cutaneous lupus	Clinical		X	X	Diagnosis codes (ICD-9/10), medications
Alopecia	Clinical		X	X	Diagnosis codes (ICD-9/10)
Arthritis (synovitis)	Clinical	X	X	X	Diagnosis codes (ICD-9/10)
Chronic cutaneous lupus	Clinical		X	X	Diagnosis codes (ICD-9/10)
Delirium	Clinical			X	Diagnosis codes (ICD-9/10)
Discoid rash	Clinical	X			Diagnosis codes (ICD-9/10)
Fever	Clinical			X	Diagnosis codes (ICD-9/10)
Haematological disorder	Clinical	X			Diagnosis codes (ICD-9/10), laboratory results
Haemolytic anaemia	Clinical		X	X	Diagnosis codes (ICD-9/10), laboratory results
Leucopenia	Clinical		X	X	Diagnosis codes (ICD-9/10), laboratory results
Malar rash	Clinical	X			Diagnosis codes (ICD-9/10), medications
Neurological disorder (ACR)	Clinical	X			Diagnosis codes (ICD-9/10)
Neurological disorder (SLICC)	Clinical		X		Diagnosis codes (ICD-9/10)
Oral ulcers	Clinical	X	X	X	Diagnosis codes (ICD-9/10)
Photosensitivity	Clinical	X			Diagnosis codes (ICD-9/10)
Renal disorder (ACR)	Clinical	X			Diagnosis codes (ICD-9/10), laboratory results
Renal disorder (SLICC)	Clinical		X		Diagnosis codes (ICD-9/10), laboratory results
Renal disorder (EULAR/ACR)	Clinical			X	Laboratory results
Seizure	Clinical			X	Diagnosis codes (ICD-9/10)
Serositis	Clinical	X	X	X	ICD-9/10 codes, procedure codes (CPT)
Thrombocytopenia	Clinical		X	X	Diagnosis codes (ICD-9/10), laboratory results
Anti-dsDNA Ab	Immunological		X	X	Laboratory results
Antinuclear Ab	Immunological	X	X	X	Laboratory results
Antiphospholipid Ab	Immunological		X	X	Laboratory results
Anti-Smith Ab	Immunological		X	X	Laboratory results
Direct Coombs test	Immunological		X		Laboratory results
Immunological disorders	Immunological	X			Diagnosis codes (ICD-9/10), laboratory results
Low complement	Immunological		X		Laboratory results
C3	Immunological			X	Laboratory results
C4	Immunological			X	Laboratory results

ACR, American College of Rheumatology; CPT, current procedural terminology; EULAR, European League Against Rheumatism; ICD-9/10, International Classification of Diseases versions 9 and 10; SLICC, Systemic Lupus International Collaborating Clinics.

in the NMEDW between 2007 and 2019. There are 885 patients in the CLD who have definite lupus as determined by the ACR classification criteria. After removing patients who did not have medical records in the NMEDW, 818 patients remained. Both ACR

and SLICC classification criteria have mechanisms to define 'definite lupus' that require satisfying ≥ 4 attributes. Given that attributes accumulate over time, are not always identified when patients first present with lupus, and to ensure sufficient data depth for analysis,¹⁹

we included patients with ≥ 4 encounters documented in the NMEDW, reducing the final SLE cohort size to 472 patients.

Gold standard healthy cohort

To create our gold standard healthy cohort, we selected 500 patients in the NMEDW who had received care in the NM general medicine clinic between 2007 and 2019 and had ≥ 4 encounters documented in the NMEDW. This cohort was frequency matched for sex, race and age to our SLE cohort, and did not have a diagnosis of lupus as determined by ICD-9/10 codes (710.0 or M32.1) and chart review. We required ≥ 4 encounters to reduce concerns of poor attribute detection due to insufficient information present in the EHR.¹⁹ Age range at time of record retrieval was between 18 and 45 to reflect the age range of first lupus diagnosis in the SLE cohort (29.6 ± 11.4 years). All 500 patient records were manually chart reviewed for the ACR, SLICC and EULAR/ACR classification criteria by medical students (AW and RS) who were trained in the chart abstraction process by the clinical lupus expert (RRG), who adjudicated the CLD. Chart abstractors reviewed 10 records together to align data abstraction definitions. Challenging cases were referred to the clinical expert (RRG) for final decision-making.

Assessment of classification criteria algorithms

We assessed and validated the performance of algorithms for the individual attributes that made up the classification criteria and the overall performance of each classification criteria using our gold standard SLE and healthy cohorts. The ACR classification criteria have 11 individual attributes, divided into clinical (eight attributes) and immunological domains (three attributes) and to be identified with definite SLE, a patient must be documented with ≥ 4 attributes.^{4 5} The SLICC classification criteria have 17 attributes divided into clinical (11 attributes) and immunological (six attributes) domains. Definite SLE is defined as having at least one clinical, one immunological domain criteria and ≥ 4 criteria overall or an identification of lupus nephritis as determined by a renal biopsy in combination with a positive ANA or anti-dsDNA test.⁶ Finally, the EULAR/ACR classification criteria have 21 individual attributes divided into 10 domains based on organ system. Attributes within the domains are ordered by degree of severity and scored with more severe attributes receiving higher scores. To be classified with SLE, patients must have a positive ANA test and score 10 or more points across any number of domains,^{7 10} and biopsy-proven lupus nephritis in the presence of positive ANA test qualifies a patient as having definite lupus.

For this study, we focused on structured EHR data that could be found in multiple EHR environments without difficulty to support reusability and increase potential portability of the algorithms. Given that we did not mine text data for biopsy results, we were unable to determine biopsy-proven lupus nephritis in the EHR so this aspect

of the SLICC and EULAR/ACR classification criteria was not examined for this study.

Analysis of electronically specified classification criteria

To assess the capability of our algorithms to identify clinical and immunological attributes of the disease using structured data from the EHR and determine whether persons with and without SLE could be distinguished with the full SLE classification algorithm, we combined our gold standard SLE and healthy patient cohorts. We assessed the sensitivity and specificity over each individual attribute and over the full classification criteria using our combined gold standard patient cohorts.

Patient and public involvement

Patients and the public were not involved in the design and conduct of this study.

RESULTS

Population demographics for our SLE and healthy cohorts are presented in [table 2](#). Both cohorts were 92% female, and had similar racial composition (approximately 50% white, 30% African-American and 20% other racial categories). Average age at onset of SLE was 30 years and average age of healthy patients when the data were extracted was 35.

Results of manual chart adjudication for classification criteria attributes for both cohorts are also described in [table 2](#). For the adjudicated SLE cohort, we observed differences in how many persons were identified as having definite lupus across the classification criteria. For the ACR criteria, 471 of 472 (99.8%) met the definition of definite lupus while 468 of 472 (99.2%) met the SLICC criteria and 452 of 472 (95.8%) satisfied the EULAR/ACR criteria. While no patients with SLE, as determined by a diagnosis code, were included in the healthy cohort, five patients in this group did satisfy the criteria for definite lupus based on chart review (2 of 500 (0.4%) for the ACR criteria, 2 of 500 (0.4%) for the SLICC criteria and 5 of 500 (1.0%) for the EULAR/ACR criteria). Two patients satisfied all classification criteria, and three patients were identified only by EULAR/ACR. None were found to have SLE by chart review. However, one had a history of pre-eclampsia, one had fibromyalgia, three had other rheumatological diseases and all had at least one laboratory result near the normal cut-off. Of the 30 attributes defined across the three classification criteria, all were identified in our gold standard SLE cohort except for delirium, an attribute of the EULAR/ACR classification criteria, which was not part of the original chart review for this cohort since it was not an attribute of the ACR 1982/1997 classification criteria. There was a wide range of occurrence rates for the attributes, ranging from 12 (direct Coombs test) to 455 (ANA) per 472 patients. Within the healthy cohort, nine attributes were not identified through chart review: photosensitivity, chronic cutaneous lupus, delirium, fever, anti-Smith,

Table 2 Basic demographics and prevalence of classification criteria attributes in gold standard case and control patient cohorts as determined by chart adjudication

	Gold standard case cohort n=472	Gold standard control cohort n=500
Sex (% female)	434 (92)	458 (92)
Mean age**	29.6 (SD 11.4)	35 (SD 6.1)
Race (%)		
White	210 (45)	255 (51)
African-American	128 (27)	144 (29)
Other	134 (28)	101 (20)
Satisfy lupus classification criteria		
ACR	471	2
SLICC	468	1
EULAR/ACR	452	5
Classification criteria attributes (%)		
Acute cutaneous lupus	407 (87)	0 (0)
Alopecia	96 (20)	16 (3)
Arthritis (synovitis)	409 (87)	41 (8)
Chronic cutaneous lupus	120 (25)	0 (0)
Delirium	ND	0 (0)
Discoid rash	97 (21)	1 (0.2)
Fever	81 (17)	0 (0)
Haematological disorder	280 (59)	10 (2)
Haemolytic anaemia	44 (9)	1 (0.2)
Leucopenia	432 (92)	6 (1)
Malar rash	250 (53)	1 (0.2)
Neurological disorder (ACR)	27 (6)	9 (2)
Neurological disorder (SLICC)	102 (22)	9 (2)
Oral ulcers	220 (47)	2 (0.4)
Photosensitivity	324 (69)	0 (0)
Renal disorder (ACR)	165 (35)	12 (2)
Renal disorder (SLICC)	178 (38)	11 (2)
Renal disorder (EULAR/ACR)	101 (21)	9 (2)
Seizure	31 (7)	13 (3)
Serositis	174 (37)	9 (2)
Thrombocytopenia	58 (13)	2 (0.4)
Anti-dsDNA Ab	329 (70)	3 (0.6)
Antinuclear Ab	455 (96)	16 (3)
Antiphospholipid Ab	127 (27)	2 (0.4)
Anti-Smith Ab	105 (22)	0 (0)
Direct Coombs test	12 (3)	0 (0)
Immunological disorders	361 (76)	2 (0.4)
Low complement	325 (69)	0 (0)
C3	265 (56)	0 (0)
C4	266 (56)	0 (0)

*For patients with SLE, age is mean age at first diagnosis of lupus; for healthy controls, age is mean age at time of data extraction.

ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; ND, not determined; SD, Standard Deviation; SLICC, Systemic Lupus International Collaborating Clinics.

low complement, C3, C4 and the Coombs test. For those attributes that were identified, occurrence rates were low, with a range of 1 (malar rash, discoid rash and haemolytic anaemia) to 41 (arthritis) occurrences

per 500 patients. Absence of attribute identification by chart review is not the same as a clinical determination that an attribute is not present and does not eliminate the possibility that a patient may have a given attribute,

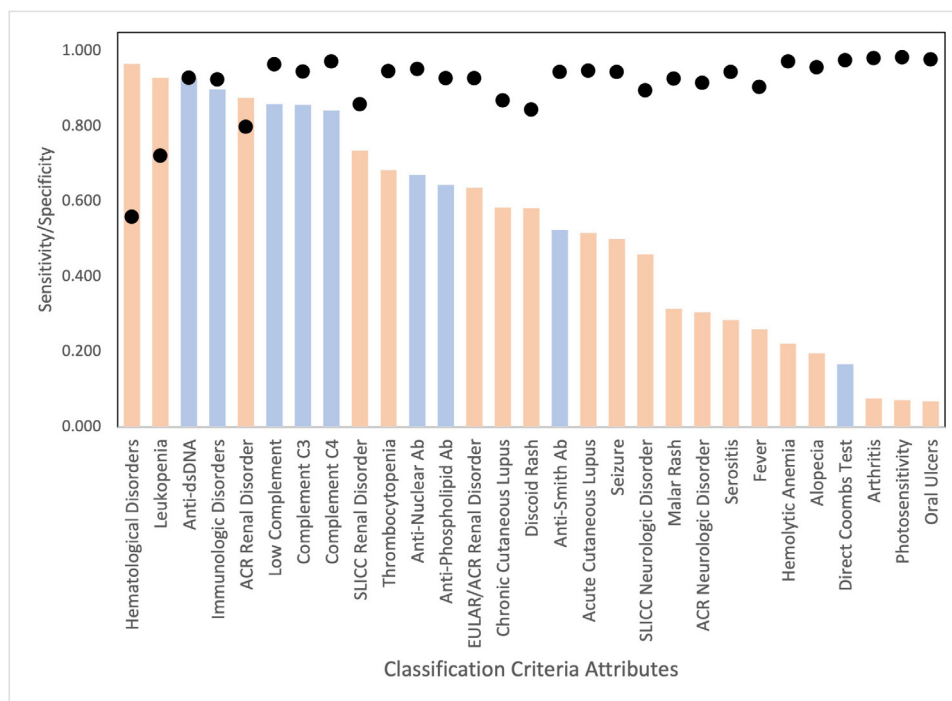


Figure 1 Sensitivity and specificity of classification criteria attributes in electronic health record (EHR) data. The algorithms for 29 attribute components of the ACR, SLICC and EULAR/ACR classification criteria were assessed for sensitivity and specificity of attribute detection in EHR data relative to chart adjudication results for the same patients. For each attribute, sensitivity is displayed via coloured bars and specificity by the black points. Sensitivity bars for clinical attributes are shown in red and immunological attributes are shown in blue. ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; SLICC, Systemic Lupus International Collaborating Clinics.

but it is not documented in our records because care was received elsewhere.

We used our algorithms to identify the 30 individual attributes that comprise the ACR, SLICC and EULAR/ACR classification criteria in the EHR data for our combined SLE and healthy cohorts (see [table 1](#) for attribute data types and online supplemental table 1 for full definitions) and compared the results to manual chart adjudication. [Figure 1](#) shows the sensitivity and specificity for the clinical and immunological domain attributes. Overall, the sensitivity of the individual attributes had a wide range from 7% (oral ulcers) to 97% (haematological disorders) with a median sensitivity of 58%, while the range of specificity was narrower: 56% (haematological disorders) to 98% (photosensitivity), with a median sensitivity of 94%. The lowest sensitivity criteria were primarily those based on diagnosis codes, and several criteria were very difficult to detect, including photosensitivity, oral ulcers and arthritis, and the direct Coombs test. The highest sensitivity criteria were based on laboratory results, reflecting that laboratories are both billed for and fully documented in the structured EHR data due to electronic laboratory reporting that directly populates the EHR with test results.

[Table 3](#) describes the sensitivity and specificity of detection of the clinical and immunological domain attribute groups in EHR data for each classification criteria. For all three classification criteria, the median sensitivity of the immunological domain attributes was higher than the clinical domain

attributes. Within the ACR criteria the median sensitivity and specificity of the clinical domain criteria were 31% (range 7%–97%) and 93% (range 56%–98%), and for the immunological domain criteria they were 78% (range 67%–90%) and 94% (range 93%–95%). Within the SLICC criteria the median sensitivities and specificities of the clinical and immunological domain criteria were 46% (range 7%–93%) and 95% (range 72%–98%), and 66% (range 17%–93%) and 94% (range 93%–98%), respectively. Finally, while the EULAR/ACR criteria are grouped into systemic domains to calculate a score to determine definite lupus, we grouped them into clinical and immunological domains (see [table 1](#)) to support comparison to ACR and SLICC criteria. For the EULAR/ACR criteria, the median sensitivity and specificity were 39% (range 7%–93%) and 94% (range 72%–98%) for the clinical domain, and 76% (range 65%–94%) and 95% (range 93%–98%) for the immunological domain. Overall, across all the attributes present in all three classification criteria, the median sensitivity and specificity of detection of the clinical attributes were 46% (range 7%–97%) and 93% (range 56%–98%), and the median sensitivity and specificity of detection of the immunological attributes were 84% (range 17%–93%) and 95% (range 93%–98%). Taken together, these data suggest that the immunological criteria that are primarily based on laboratory results are more accurately detected in EHR data and may make a stronger contribution to overall classification criteria scoring than the clinical domain attributes.

Table 3 Sensitivity and specificity of algorithms to identify classification criteria attributes in structured EHR data grouped by SLE classification criteria and attribute domain as specified by classification criteria

Classification criteria	Clinical domain attributes			Immunological domain attributes		
	Median sensitivity (%)	Range	Median specificity (%)	Range	Median sensitivity (%)	Range
ACR 1997	31	7%–97%	93	56%–98%	78	67%–90%
SLICC 2012	46	7%–93%	95	72%–98%	66	17%–93%
EULAR/ACR 2019	39	7%–93%	94	72%–98%	76	54%–93%
All	46	7%–97%	93	56%–98%	84	17%–93%

ACR, American College of Rheumatology; EHR, electronic health record; EULAR, European League Against Rheumatism; SLICC, Systemic Lupus International Collaborating Clinics.

All three classification criteria have a definition of ‘definite lupus’ (see the Methods section). We assessed whether determination of ‘definite lupus’ based on EHR data could distinguish patients in our lupus and healthy cohorts for each classification criteria (table 4). Both the SLICC and EULAR/ACR criteria have a path to defining definite lupus through a positive renal biopsy in the presence of a positive ANA test. Since natural language processing of free text notes is required to identify renal biopsy results, we did not include this path in our assessment of either algorithm. Based on structured EHR data, 301 patients satisfied the ACR criteria, compared with 471 determined by chart adjudication, while for the SLICC criteria, 358 out of 468 satisfied both the domain requirements and met the requirement of four or more criteria attributes overall. The EULAR/ACR criteria require a positive ANA test as an entry criterion. Of 452 patients who satisfied the EULAR/ACR entry criterion and scoring based on chart adjudication, only 269 could be identified based on structured EHR data. Of the five healthy cohort members who satisfied one or more classification criteria by chart review, three were also identified using EHR data. Two members of the healthy cohort who did not satisfy classification criteria by chart adjudication were identified using EHR-based algorithms. Chart review determined that these patients did not have SLE; however, both were ANA positive, had borderline laboratory results and neurological presentations.

We observed similar performance characteristics across all three classification criteria-based algorithms using structured EHR data. The sensitivity and specificity of the algorithms ranged from 59% to 76% and from 97% to 99%, respectively.

DISCUSSION

SLE is a complex disease with highly variable presentation that makes it difficult to identify and characterise. We examined whether rule-based algorithms for SLE classification criteria could be used to detect attributes of the criteria in the structured EHR data of persons with lupus and distinguish them from our healthy cohort. Our results, based on an SLE cohort with linked medical records that had been fully adjudicated for the ACR classification criteria, demonstrate that all three existing classification criteria (ACR, SLICC and EULAR/ACR) have the potential to describe persons with lupus in EHR data and distinguish them from a known healthy patient cohort and that the algorithms have high overall sensitivity and specificity. Thus, classification criteria-based algorithms may be a foundation for characterisation and identification of people with lupus.

The overall performance of the algorithms to discriminate between people with and without lupus was similar, consistent with the significant attribute overlap between the ACR, SLICC and EULAR/ACR criteria definitions and how those attributes are assessed. The SLICC-based algorithm demonstrated higher sensitivity than those based

Table 4 Overall performance of classification criteria-based algorithms to distinguish lupus cases and healthy controls using structured data from electronic health records

Classification criteria	Case detection	Control detection	Sensitivity	Specificity
ACR 1997	301/471 (64%)	3/500 (0.6%)	64%	99%
SLICC 2012	357/468 (76%)	3/500 (0.6%)	76%	99%
EULAR/ACR 2019	269/452 (59%)	3/500 (0.6%)	59%	97%
Average	66%	0.6%	66%	98%

ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; SLICC, Systemic Lupus International Collaborating Clinics.

on ACR and EULAR/ACR, likely due to the stronger reliance of the SLICC criteria on individual laboratory attributes compared with clinical attributes, which were easier to detect in EHR data and had generally higher sensitivity and specificity than clinical attributes. In particular, SLICC scores individual laboratory attributes, while the ACR criteria incorporate them in composites (immunological and haematological disorders) thus reducing their power to define ‘definite lupus’ compared with clinical attributes in this context. The EULAR/ACR classification criteria require a positive ANA test as an entry criterion before the rest of the attributes are assessed. Depending on policies in the care environment, the ANA test may not be repeated if evidence of a historic positive test is present.^{20–22} Some patients in our SLE cohort have a long history of lupus and their ANA tests were performed prior to entering the CLD or receiving care documented in the EHR. The ANA test for these patients was documented in their clinical notes and could not be detected as a laboratory result, reducing the number of patients who were assessed for the full set of criteria and the overall sensitivity of the EULAR/ACR-based algorithm in our data set. Given that one of the stated goals of the EULAR/ACR criteria was to include attributes that were found earlier in the development of disease, such as fever,⁹ this problem may be less relevant when assessing the information of newly identified patients with lupus. Importantly, while the SLICC algorithm had the best overall performance in this study and may be a good choice for general studies using EHR data, all three algorithms are viable for use in medical records and different populations of patients with lupus, research applications and EHR documentation approaches may make a given algorithm preferable in different environments.

Our study has several limitations to consider. First, by opting to work with structured EHR data to optimise future conversion of the algorithm to common data models such as the Patient-Centered Outcomes Research Network or Observational Medical Outcomes Partnership common data models^{23 24} that do not include free text data, we lost information from notes for clinical care and more complex procedures (such as renal biopsies) that are an important part of lupus care. Second, the primary role of the EHR is to document clinical care and support billing for rendered care. We believe that the reduction in sensitivity for some clinical criteria (such as photosensitivity,

oral ulcers and arthralgia) is primarily due to these criteria not generally being used as diagnoses for billing purposes and instead being documented in clinical notes. Likewise, the direct Coombs test is an infrequent test used in the context of haematological diagnoses that may end up documented in clinical notes instead of laboratory results, particularly if the test occurred at a different healthcare institution. We are exploring use of natural language processing to identify low sensitivity clinical attributes that are not often billed for (arthritis, oral ulcers), attributes derived from procedure notes (renal biopsy), and to detect historic laboratory data (ANA, direct Coombs test, autoantibody laboratory results) captured in clinical notes to improve identification of these attributes and ensure better description of lupus presentation when free text is available. Third, in addition to where data are located in the record, there could be differences in how data are entered by different care providers, what tests are used at different institutions, or depth of documentation for patients who are part of a registry versus those who are not. Thus, the algorithm may perform differently at different sites. Fourth, SLE is a complex disease with a wide range of presentations and severities. Patients may see many different care providers and get care at multiple healthcare locations depending on emergency needs or changes in insurance status. This may disproportionately impact specific subgroups of patients with regard to earlier identification, particularly non-white patients and those who are the beneficiaries of public insurance, given that previous work has shown that non-white patients and those with public insurance are much more likely to receive care in more than one location.²⁵ This is a single site study, and data for all the care patients in our cohorts received may not be represented in one EHR database. More patients may satisfy our algorithms if data from more than one organisation were present. We are addressing this limitation by applying our algorithms to a clinical data research network that represents the majority of Chicago medical centres²⁶ to explore whether a broader picture of care improves algorithm performance. Fifth, only cases from the CLD defined as ‘definite lupus’ were included in our SLE cohort, and our healthy cohort was selected for absence of a lupus diagnosis by both clinical data and chart review. While these algorithms can be used to describe people with lupus and were able to distinguish them from our healthy cohort,

further study will be required to determine whether it can be used to differentiate people with and without lupus in broader rheumatological and general patient populations where it is unlikely lupus-focused laboratories will be run. We caution against using them to identify people with lupus in EHR data for general or rheumatological populations until further study has been performed. In particular, in future studies, we will evaluate laboratory thresholds and neurological attribute diagnosis codes which may have been responsible for identification of attributes in patients in the healthy cohort who did not have SLE. Finally, while classification criteria represent an evidence-based consensus on what experts believe to be the most important descriptors for lupus, the heterogeneity of the disease combined with its relative rarity means that some important but rarely seen attributes are not included in any classification criteria and may limit detection of some people with lupus.

We have demonstrated that rule-based algorithms using structured EHR data to identify features of lupus from classification criteria can be developed and used to describe people with lupus, and that all three existing classification criteria are effective for this task. This work suggests that it may be possible to characterise the spectrum of disease in people with lupus as described through care documented in medical records. Thus, these algorithms have the potential to help identify classification criteria attributes in patients with lupus for inclusion in clinical trials, reduce chart review burden for patients with lupus participating in clinical research, provide a foundation for exploring lupus subtypes and to create tools that can improve population health management, screening and care for people with lupus.

Author affiliations

¹Division of General Internal Medicine and Geriatrics, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

²Center for Health Information Partnerships, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

³Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

⁴Division of Rheumatology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

⁵Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

Twitter Theresa L Walunas @TheresaWalunas

Contributors TLW was involved in the conception and design of the project, acquisition, analysis and interpretation of data and drafting the manuscript. ASG, JAP, VM and KLJ were involved in acquisition, analysis and interpretation of data. AW, RS, AC, DE and KMC were involved in the acquisition and management of data. YL and ANK were involved in the acquisition and interpretation of data. RRG was involved in the conception and design of the project, acquisition, analysis and interpretation of data and critical review of the manuscript. All authors were involved in manuscript review and revision and approval of the final draft.

Funding The authors received funding support provided by grants from the National Institute of Arthritis and Musculoskeletal Disease (5R21AR072262 and P30AR072579) and the National Human Genome Research Institute (U01HG008657).

Competing interests ANK is a strategic advisor for Datavant.

Patient consent for publication Not required.

Ethics approval This study was governed by the Northwestern University Institutional Review Board (NU IRB) under study protocols STU00205559 (SLE Phenotype Development) and STU00009193 (Chicago Lupus Database, CLD). Participants in the Chicago Lupus Database gave informed consent for the inclusion of their data in the database and linkage of their medical records to data collected for the CLD. Our healthy control population was governed under the SLE Phenotype Development protocol. We received a waiver of consent from the NU IRB to analyse the data from this cohort given impracticality of consent and low risk of the study. Medical record review for this study was approved by the NU IRB.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement The data used and analysed during the current study are available from the corresponding author on reasonable request. Note that row-level data are access controlled and cannot be provided publicly due to data use restrictions. Algorithms for identification of classification criteria attributes can be found in the Phenotype KnowledgeBase (PheKB) <https://phekb.org/>.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

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

Theresa L Walunas <http://orcid.org/0000-0002-7653-3650>

REFERENCES

- Lisnevskaja L, Murphy G, Isenberg D. Systemic lupus erythematosus. *Lancet* 2014;384:1878–88.
- Murphy G, Lisnevskaja L, Isenberg D. Systemic lupus erythematosus and other autoimmune rheumatic diseases: challenges to treatment. *Lancet* 2013;382:809–18.
- Bruce IN, O’Keeffe AG, Farewell V, *et al*. Factors associated with damage accrual in patients with systemic lupus erythematosus: results from the systemic lupus international collaborating clinics (SLICC) inception cohort. *Ann Rheum Dis* 2015;74:1706–13.
- Tan EM, Cohen AS, Fries JF, *et al*. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982;25:1271–7.
- Hochberg MC. Updating the American College of rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40:1725.
- Petri M, Orbai A-M, Alarcón GS, *et al*. Derivation and validation of the systemic lupus international collaborating clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum* 2012;64:2677–86.
- Aringer M, Costenbader K, Daikh D, *et al*. 2019 European league against Rheumatism/American College of rheumatology classification criteria for systemic lupus erythematosus. *Arthritis Rheumatol* 2019;71:1400–12.
- Aringer M, Dörner T, Leuchten N, *et al*. Toward new criteria for systemic lupus erythematosus—a standpoint. *Lupus* 2016;25:805–11.
- Johnson SR, Khanna D, Daikh D, *et al*. Use of consensus methodology to determine candidate items for systemic lupus erythematosus classification criteria. *J Rheumatol* 2019;46:721–6.
- Aringer M, Costenbader K, Daikh D, *et al*. 2019 European league against Rheumatism/American College of rheumatology classification criteria for systemic lupus erythematosus. *Ann Rheum Dis* 2019;78:1151–9.
- Mosca M, Costenbader KH, Johnson SR, *et al*. Brief report: how do patients with newly diagnosed systemic lupus erythematosus present? A multicenter cohort of early systemic lupus erythematosus to inform the development of new classification criteria. *Arthritis Rheumatol* 2019;71:91–8.
- Tedeschi SK, Johnson SR, Boumpas D, *et al*. Developing and refining new candidate criteria for systemic lupus erythematosus

- classification: an international collaboration. *Arthritis Care Res* 2018;70:571–81.
- 13 Blumenthal D. Launching HITECH. *N Engl J Med* 2010;362:382–5.
 - 14 Henry J, Pylypchuk Y, Searcy T, *et al.* Adoption of electronic health record systems among U.S. Non-Federal acute care hospitals: 2008–2015. *onc data brief No. 35*. Washington, D.C: Office of the National Coordinator for Health Information Technology, 2016.
 - 15 Office of the National Coordinator for Health IT. Office-Based physician health it adoption and use: national electronic health records survey, 2015. Available: <https://dashboard.healthit.gov/datadashboard/documentation/physician-health-it-adoption-use-data-documentation.php> [Accessed 20 Dec 2016].
 - 16 Gómez-Puerta JA, Barbhayya M, Guan H, *et al.* Racial/Ethnic variation in all-cause mortality among United States Medicaid recipients with systemic lupus erythematosus: a Hispanic and Asian paradox. *Arthritis Rheumatol* 2015;67:752–60.
 - 17 Barnado A, Casey C, Carroll RJ, *et al.* Developing electronic health record algorithms that accurately identify patients with systemic lupus erythematosus. *Arthritis Care Res* 2017;69:687–93.
 - 18 Jorge A, Castro VM, Barnado A, *et al.* Identifying lupus patients in electronic health records: development and validation of machine learning algorithms and application of rule-based algorithms. *Semin Arthritis Rheum* 2019;49:84–90.
 - 19 Rasmussen LV, Thompson WK, Pacheco JA, *et al.* Design patterns for the development of electronic health record-driven phenotype extraction algorithms. *J Biomed Inform* 2014;51:280–6.
 - 20 Yazdany J, Schmajuk G, Robbins M, *et al.* Choosing wisely: The American College of Rheumatology's top 5 list of things physicians and patients should question. *Arthritis Care Res* 2013;65:329–39.
 - 21 Choi MY, Fritzler MJ. Autoantibodies in SLE: prediction and the p value matrix. *Lupus* 2019;28:1285–93.
 - 22 Choi MY, Fritzler MJ. Challenges and advances in SLE autoantibody detection and interpretation. *Curr Treatm Opt Rheumatol* 2019;5:147–67.
 - 23 PCORnet. PCORnet common data model, 2019. Available: <https://pcornet.org/data/>
 - 24 Overhage JM, Ryan PB, Reich CG, *et al.* Validation of a common data model for active safety surveillance research. *J Am Med Inform Assoc* 2012;19:54–60.
 - 25 Walunas TL, Jackson KL, Chung AH, *et al.* Disease outcomes and care fragmentation among patients with systemic lupus erythematosus. *Arthritis Care Res* 2017;69:1369–76.
 - 26 Kho AN, Hynes DM, Goel S, *et al.* Capricorn: Chicago area patient-centered outcomes research network. *J Am Med Inform Assoc* 2014;21:607–11.