The Journal of Rheumatology 2023;50:512–8
doi:10.3899/jrheum.220100
First Release December 15 2022

The Journal of Rheumatology

Evaluating the Threshold Score for Classification of Systemic Lupus Erythematosus Using the EULAR/ACR Criteria

Sindhu R. Johnson[1], Juan P. Diaz Martinez[2], Laura Whittall-Garcia[3], Murray B. Urowitz[3], Dafna D. Gladman[3], and Zahi Touma[3]

ABSTRACT. ***Objective.*** To evaluate whether a change in the European Alliance of Associations for Rheumatology (EULAR)/American College of Rheumatology (ACR) systemic lupus erythematosus (SLE) classification criteria threshold score affects accurate classification of SLE cases compared to disease-based control subjects. We evaluated a range of threshold scores to determine the score that maximizes the accurate classification of early SLE.

Methods. We conducted a cross-sectional study comparing SLE cases and control patients. A EULAR/ACR criteria score was calculated using baseline information. Sensitivity, specificity, positive likelihood ratios (+LRs), and negative likelihood ratios (–LRs) with 95% CIs were used to evaluate operating characteristics. Threshold scores of 6 to 12 were evaluated in subjects with early disease (ie, disease duration of ≤ 5 years). +LRs > 10 and –LRs < 0.1 provide evidence to rule in or rule out SLE.

Results. A total of 2764 patients were included: 1980 SLE cases who fulfilled either the ACR or Systemic Lupus International Collaborating Clinics criteria and 784 control subjects. The EULAR/ACR SLE criteria had a sensitivity of 98% (95% CI 97-98), a specificity of 99% (95% CI 98-100), a +LR of 95.5 (95% CI 48.0-190), and a –LR 0.03 (95% CI 0.02-0.03). The criteria operated well in those with early disease, in women, in men, and in White, Black, Chinese, and Filipino people. A score of 10 maximized the accurate classification of patients with early disease (+LR 174.4, 95% CI 43.8-694.6; –LR 0.03, 95% CI 0.02-0.04). An increase in the threshold score from 10 to 11 resulted in significant worsening in the –LR (threshold score 10: –LR 0.03, 95% CI 0.02-0.03 vs threshold score 11: –LR 0.05, 95% CI 0.04-0.06).

Conclusion. The EULAR/ACR SLE classification criteria threshold score of 10 performs well, particularly among those with early disease and across sexes and ethnicities.

Key Indexing Terms: classification criteria, epidemiology, systemic lupus erythematosus

Systemic lupus erythematosus (SLE) is a heterogeneous disease characterized by immune activation, target organ inflammation, and damage.[1] Classification criteria are used to identify homogeneous groups of patients for inclusion into observational studies and trials.[2] The reduction in disease heterogeneity improves the ability to make valid inferences within studies and improves generalizability across studies.[2,3] The European Alliance of Associations for Rheumatology (EULAR)/American College of Rheumatology (ACR) classification criteria for SLE[4,5] were developed with a balanced use of data-driven and expert-based consensus methods.[6-10]

The EULAR/ACR SLE classification criteria constitute a defined system that produces a measure of the relative probability that a particular case (ie, a combination of clinical features) has SLE. This system is comprised of an additive point system with hierarchical clustering of clinical and serologic features. Antinuclear antibody positivity (titer of ≥ 1/80 on HEp-2 cells or an equivalent test) is required as an entry criterion. A score of 10 is the threshold above which experts would classify cases as SLE for the purpose of research studies. Each criterion has been defined carefully to ensure appropriate face/content validity and to improve the reliability of application (ie, precision). As part of the development process, this system was validated against a large number of cases, including many cases that are not clear-cut SLE. In the validation cohort, this system had 96% sensitivity and 94% specificity compared to other disease-based controls.[11,12]

Independent external validation is an expectation of both EULAR and ACR of all criteria sets and is the final requirement after endorsement. Published validation studies with only SLE cases are limited as they only report sensitivity.[13-15] Control subjects are required for estimation of specificity and negative

This study was, in part, supported by the Schroeder Arthritis Institute, University Health Network, Toronto, Canada.

[1]S.R. Johnson, MD, PhD, FRCPC, Division of Rheumatology, Schroeder Arthritis Institute, Krembil Research Institute, Department of Medicine, Toronto Western Hospital, Mount Sinai Hospital, and Institute of Health Policy, Management and Evaluation, University of Toronto; [2]J.P. Diaz Martinez, MSc, Centre for Prognosis Studies in Rheumatic Diseases, Toronto Western Hospital, University of Toronto Lupus Clinic; [3]L. Whittall-Garcia, MD, M.B. Urowitz, MD, D.D. Gladman, MD, FRCPC, Z. Touma, MD, PhD, Division of Rheumatology, Schroeder Arthritis Institute, Krembil Research Institute, and Department of Medicine, Toronto Western Hospital, University of Toronto, Toronto, Ontario, Canada.

The authors declare no conflicts of interest relevant to this article.

Address correspondence to Dr. S.R. Johnson, Division of Rheumatology, Ground Floor, East Wing, Toronto Western Hospital, 399 Bathurst Street, Toronto, ON M5T 2S8, Canada. Email: Sindhu.Johnson@uhn.ca.

Accepted for publication October 13, 2022.

predictive value. In addition, there have been calls to revise the criteria or the threshold of 10 for classification.[16-18] Cui and colleagues[17] opined that "rheumatologists should be informed of exact probability of illness in patients with underlying SLE who are below the threshold (ie, total score <10) so as to provide better decision-making, evaluation and follow-up." Rheumatologists and researchers should use classification criteria with the proper level of confidence in categorization derived from not only sensitivity and specificity, but also the positive likelihood ratio (+LR) and negative likelihood ratio (–LR). Finally, the ability to accurately identify patients with early disease is important so that these patients may have access to earlier intervention and trials of innovative therapies with the goal of preventing damage or other ominous outcomes.[19]

The primary objective of this study was to evaluate whether a change in the threshold score would affect accurate classification of SLE compared to disease-based control subjects. Our secondary objectives were to evaluate whether a change in threshold score would affect accurate classification in subsets of patients with SLE, specifically early disease, across sexes and ethnicities.

METHODS

Subjects. We conducted a cross-sectional study across multiple clinics at University of Toronto–affiliated hospitals. The University of Toronto Lupus Cohort encompasses an inception cohort (ie, joined the clinic within the first 12 months from SLE diagnosis) and prevalent patients (ie, joined the clinic after 12 months from SLE diagnosis). All patients with SLE met the revised 1997 ACR classification criteria for SLE, or they met 3 criteria and had a supportive skin or kidney biopsy.[20,21] Consecutive control subjects from the general and specialty rheumatology clinics of Toronto Western Hospital and Mount Sinai Hospital, Toronto, Canada, were included. No additional inclusion or exclusion criteria (eg, age ranges and presence or absence of specific diagnoses) were used. All patients with systemic sclerosis (SSc) met the ACR/EULAR classification criteria for SSc.[22] All other diagnoses were physician based.

Data collection. All clinical, patient-reported, and serologic data were obtained from the patient chart and electronic medical record. Standardized abstraction forms were used to collect sex, disease duration, ethnicity, diagnosis, clinical manifestations, and serology. Clinical manifestations were attributed to SLE if there was no more likely alternative explanation.[4,8] For each patient, the EULAR/ACR criteria score was calculated based on the baseline clinical, laboratory, and renal biopsy information. The baseline information was obtained from the first 2 visits—both visits occurring within 1 to 4 months—as some of the tests ordered at the first visit were recorded only at the second visit. Data were entered into a computerized database. Data quality was maintained using logic and data entry checks.

Analysis. Descriptive statistics were used to summarize the data. Sensitivity, specificity, +LR, and –LR for the EULAR/ACR classification criteria were estimated with 95% CIs. CIs for +LRs and –LRs are based on formulas provided by Simel et al.[23]

The operating characteristics of threshold scores of 6 to 12 were evaluated for the total cohort; in subjects with early disease, defined as a disease duration of 5 years or less; and across sexes and ethnicities. LRs above 10 and below 0.1 were considered to provide strong evidence to rule in or rule out diagnoses. An LR close to 1 suggests the criteria were of little value. The optimal threshold is one that maximizes the +LR and minimizes the –LR. Receiver operating characteristic (ROC) curves were used to illustrate the performance of the EULAR/ACR SLE classification criteria in the full cohort, in those with early disease, in men, in women, and in patients of White or Black ethnicity.

Ethics. University Health Network Research Ethics Board (REB) approval was obtained prior to the conduct of this study (REB No. 17-5926).

Patient and public involvement. Patients were involved with data collection and will be involved in choosing which results to share, when to share results, and in what format study results will be disseminated to the relevant wider patient communities.

RESULTS

Patients. We included 2764 patients with 1980 SLE cases who fulfilled either the ACR or Systemic Lupus International Collaborating Clinics (SLICC) criteria and we included 784 controls. Out of 2764 patients, 427 (15%) were male. Control subject diagnoses included systemic autoimmune rheumatic diseases (ie, rheumatoid arthritis, systemic sclerosis, and Sjogren syndrome), infections, metabolic diseases, and malignancies (Table 1). The mean disease duration was 4.49 (SD 6.3) years for SLE cases and 8.50 (SD 8.4) years for control subjects at the assessment visit. A total of 1727 (87%) out of 1980 SLE cases and 610 (78%) out of 784 control subjects were female. The ethnicities of the 1980 SLE cases were White (n = 1276, 64%), Black (n = 270, 14%), Chinese (n = 193, 10%), Filipino (n = 70, 3.5%), and First Nations (n = 17, 1%). Ethnicities among the 784 control subjects were White (n = 584, 74%), Black (n = 44, 6%), Chinese (n = 42, 5%), Filipino (n = 14, 1.8%), and First Nations (n = 3, 0.4%; Table 2). The rate of occurrence of the individual classification criteria attributes in the cases and control subjects are reported in Table 3.

Operating characteristics. The EULAR/ACR SLE classification criteria had a sensitivity of 98% (95% CI 97%-98%), a specificity of 99% (95% CI 98%-100%), a +LR of 95.5 (95% CI 48.0-190), and a –LR of 0.03 (95% CI 0.02-0.03). The ROC curve for the full cohort is illustrated in Figure 1, with an area under the curve (AUC) of 0.9964. In patients with early disease, defined as 5 years or less duration, the EULAR/ACR SLE classification criteria had a sensitivity of 97% (95% CI 96%-98%), a specificity of 99% (95% CI 98%-100%), a +LR of 174.40 (95% CI 43.78-694.64), and a –LR of 0.03 (95% CI 0.02-0.04). The ROC curve, with an AUC of 0.9974, for patients with early disease is illustrated in Supplementary Figure S1 (available from the authors upon request). In women, the EULAR/ACR SLE classification criteria had a sensitivity of 98% (95% CI 97%-98%), a specificity of 99% (95% CI 98%-100%), a +LR of 99.1 (95% CI 44.7-219.8), and a –LR of 0.03 (95% CI 0.02-0.03). The ROC curve for female patients, with an AUC of 0.9963, is illustrated in Supplementary Figure S2 (available from the authors upon request); the ROC curve for male patients, with an AUC of 0.9961, is illustrated in Supplementary Figure S3 (available from the authors upon request). The ROC curve for White patients, with an AUC of 0.9955, is illustrated in Supplementary Figure S4 (available from the authors upon request); the ROC curve for Black patients, with an AUC of 0.9997, is illustrated in Supplementary Figure S5 (available from the authors upon request). The sensitivity, specificity, +LR, and –LR results for the full cohort, for those with early disease, across sexes, and

Table 1. Summary of control subjects.

Diagnosis	Subjects, N = 784, n (%)
Rheumatic diseases	
Systemic sclerosis	271 (34.6)
Raynaud phenomenon	78 (9.9)
Rheumatoid arthritis	76 (9.7)
Undifferentiated connective tissue disease	35 (4.5)
Psoriatic arthritis	31 (4.0)
Eosinophilic fasciitis	18 (2.3)
Osteoarthritis	21 (2.7)
Gout	17 (2.1)
Fibromyalgia	12 (1.5)
Vasculitis[a]	20 (2.6)
Ankylosing spondylitis	11 (1.4)
Mixed connective tissue disease	9 (1.1)
Inflammatory myositis[b]	13 (1.7)
Sjogren syndrome	12 (1.5)
Psoriasis	5 (0.6)
Polymyalgia rheumatica	6 (0.8)
Sarcoidosis	3 (0.4)
Spondyloarthritis	4 (0.5)
Arthralgia	6 (0.8)
Juvenile idiopathic arthritis	6 (0.8)
Inflammatory bowel disease[c]	5 (0.6)
Dupuytren contractures	2 (0.3)
Frozen shoulder	1 (0.1)
Behçet disease	1 (0.1)
Osteoporosis	1 (0.1)
Infections	
Septic arthritis	5 (0.6)
Viral infection	1 (0.1)
Staphylococcus aureus skin infection	1 (0.1)
Tinea corporis	1 (0.1)
Metabolic disease	
Thyroid disease[d]	5 (0.6)
Diabetes[e]	6 (0.8)
Other conditions	
Interstitial lung disease	2 (0.3)
Psychosis	1 (0.1)
Sickle cell anemia	1 (0.1)
Multiple thrombosis	1 (0.1)
Rhabdomyolysis	1 (0.1)
Recurrent pericarditis	1 (0.1)
Inflammatory arthritis	3 (0.4)
Dysphagia	1 (0.1)
Glomerulonephritis	1 (0.1)
Cardiogenic shock	1 (0.1)
Vascular disease[f]	3 (0.4)
Preeclampsia	1 (0.1)
Liver disease[g]	2 (0.3)
Skin conditions[h]	49 (6.3)
Alopecia	1 (0.1)
Mechanical low-back pain	2 (0.3)
Fatigue	1 (0.1)
Telangiectasia	1 (0.1)
Headache	3 (0.4)
Serum sickness	1 (0.1)
Antiphospholipid antibody syndrome	1 (0.1)
Superficial thrombophlebitis	1 (0.1)
Lymphedema	1 (0.1)
Multiple sclerosis	1 (0.1)
Antisynthetase syndrome	1 (0.1)
Rotator cuff tendonitis	1 (0.1)
Relapsing polychondritis	1 (0.1)
IgG4-related disease	1 (0.1)
Malignancy[i]	12 (1.5)

[a] Vasculitis includes polyarteritis nodosa, cryoglobulinemic vasculitis, giant cell arteritis, and granulomatosis with polyangiitis. [b] Inflammatory myositis includes dermatomyositis, polymyositis, and checkpoint inhibitor myositis. [c] Inflammatory bowel disease includes ulcerative colitis. [d] Thyroid disease includes hypothyroidism, Hashimoto thyroiditis, and livedo reticularis. [e] Diabetes includes scleredema diabeticorum and diabetic cheiroarthropathy. [f] Vascular disease includes peripheral vascular disease, thromboangiitis obliterans, and chilblains. [g] Liver disease includes autoimmune hepatitis and hemochromatosis. [h] Skin conditions includes rash, lichen planus, eczema, rosacea, morphea, and scleromyxedema. [i] Malignancy includes lymphoma, leukemia, melanoma, neuroendocrine tumor, and oral, thyroid, and breast cancer.

across ethnicities confirmed the accurate performance of the EULAR/ACR SLE criteria and are reported in Table 4.

The LRs for threshold scores of 6 to 12 are reported in Table 5. By balancing both sensitivity and specificity through maximizing the +LR and minimizing the –LR, a threshold score of 10 maximizes the accurate classification of patients with SLE with early disease (+LR 174.40, 95% CI 43.78-694.64; –LR 0.03, 95% CI 0.02-0.04), women (+LR 99.13, 95% CI 44.71-219.80; –LR 0.03, 95% CI 0.02-0.03), and men (+LR 84.94, 95% CI 21.41-336.96; –LR 0.02, 95% CI 0.01-0.05). An increase in the threshold score from 10 to 11 for the full cohort does not result in a statistically significant improvement in the +LR (threshold score 10: +LR 95.57, 95% CI 47.96-190.44; vs threshold score 11: +LR 106.17, 95% CI 50.78-221.99). In contrast, an increase in the threshold score from 10 to 11 for the full cohort results in a statistically significant worsening in the –LR (threshold score 10: –LR 0.03, 95% CI 0.02-0.03; vs threshold score 11: –LR 0.05, 95% CI 0.04-0.06).

DISCUSSION

This study independently validates the operating characteristics of the EULAR/ACR classification criteria for SLE. We first confirmed the accuracy of the classification criteria in identifying patients with SLE compared to disease-based control subjects. Second, we demonstrated that these criteria work well in early disease, across sexes, and across ethnicities. Most importantly, we demonstrated that a threshold score of 10 is the optimal score for accurate classification, particularly in early disease. To increase the threshold score would increase the risk of misclassification.

LRs are considered superior methods for statistical modeling and making inferences about test performance. LRs are more informative than sensitivity or specificity because they consider

Table 2. Comparison of demographics across SLE cases and control subjects.

	SLE Cases, N = 1980	Disease-Based Control Subjects, N = 784
Sex		
Female	1727 (87)	610 (78)
Male	253 (13)	174 (22)
Disease duration, yrs, mean (SD)	4.49 (6.3)	8.50 (8.4)
Ethnicity		
White	1276 (64)	584 (74)
Black	270 (14)	44 (6)
Chinese	193 (10)	42 (5)
Filipino	70 (3.5)	14 (1.8)
First Nations	17 (1)	3 (0.4)

Data are in n (%) unless otherwise indicated. SLE: systemic lupus erythematosus.

Table 3. Frequency of SLE classification criteria in the full cohort and patient subsets.

Criteria	Total cohort, N = 2764, n (%)		Women, n = 2337 n (%)		Men, n = 427 n (%)		Early Disease, ≤ 5 Years Duration, n = 1719, n (%)		White, n = 186 n (%)		Black, n = 314 n (%)	
	SLE Cases	Con.	SLE Cases	Con.	SLE Cases	Con.	SLE Cases	Con.	SLE Cases	Con.	SLE Cases	Con.
Fever	277 (14.0)	12 (1.5)	229 (13.3)	9 (1.5)	48 (19.0)	3 (1.7)	233 (17.1)	10 (2.8)	180 (14.1)	9 (1.5)	36 (13.3)	1 (2.3)
Leukopenia	550 (27.8)	0 (0)	477 (27.6)	0 (0)	73 (28.9)	0 (0)	394 (28.9)	0 (0)	306 (24.0)	0 (0)	103 (38.1)	0 (0)
Thrombocytopenia	365 (18.4)	4 (0.5)	312 (18.1)	2 (0.3)	53 (20.9)	2 (1.1)	236 (17.3)	2 (0.6)	209 (16.4)	3 (0.5)	41 (15.2)	0 (0)
Autoimmune hemolysis	62 (3.1)	0 (0)	55 (3.2)	0 (0)	7 (2.8)	0 (0)	46 (3.4)	0 (0)	36 (2.8)	0 (0)	10 (3.7)	0 (0)
Delirium	81 (4.1)	0 (0)	70 (4.1)	0 (0)	11 (4.3)	0 (0)	57 (4.2)	0 (0)	48 (3.8)	0 (0)	18 (6.7)	0 (0)
Psychosis	86 (4.3)	1 (0.1)	75 (4.3)	1 (0.2)	11 (4.3)	0 (0)	53 (3.9)	0 (0)	51 (4.0)	0 (0)	16 (5.9)	0 (0)
Seizure	92 (4.6)	1 (0.1)	82 (4.8)	1 (0.2)	10 (4.0)	0 (0)	64 (4.7)	1 (0.3)	63 (4.9)	1 (0.2)	12 (4.4)	0 (0)
Nonscarring alopecia	875 (44.2)	9 (1.1)	812 (47.0)	7 (1.1)	63 (24.9)	2 (1.1)	614 (45.1)	2 (0.6)	494 (38.7)	5 (0.9)	141 (52.2)	2 (4.5)
Oral ulcers	698 (35.3)	6 (0.8)	624 (36.2)	5 (0.8)	74 (29.2)	1 (0.6)	478 (35.1)	2 (0.6)	460 (36.1)	4 (0.7)	87 (32.2)	0 (0)
Subacute or discoid lupus	263 (13.3)	1 (0.1)	225 (13.0)	1 (0.2)	38 (15.0)	0 (0)	170 (12.5)	1 (0.3)	154 (12.1)	1 (0.2)	65 (24.1)	0 (0)
Acute cutaneous lupus	1037 (52.4)	4 (0.5)	923 (53.5)	4 (0.7)	113 (44.7)	0 (0)	696 (51.1)	1 (0.3)	675 (52.9)	4 (0.7)	109 (40.4)	0 (0)
Pleural or pericardial effusion	521 (26.3)	5 (0.6)	445 (25.8)	3 (0.5)	76 (30.0)	2 (1.1)	351 (25.8)	3 (0.8)	322 (25.2)	3 (0.5)	78 (28.9)	0 (0)
Acute pericarditis	307 (15.5)	6 (0.8)	259 (15.0)	6 (1.0)	47 (18.6)	0 (0)	210 (15.4)	3 (0.8)	189 (14.8)	3 (0.5)	52 (19.3)	2 (4.5)
Joint involvement	1278 (64.5)	189 (24.1)	1139 (66.0)	151 (24.8)	138 (54.5)	38 (21.8)	867 (63.7)	59 (16.5)	814 (63.8)	144 (24.7)	202 (74.8)	16 (36.4)
Proteinuria > 0.5 g/24h	713 (36.0)	5 (0.6)	595 (34.5)	5 (0.8)	118 (46.6)	0 (0)	453 (33.3)	1 (0.3)	381 (29.9)	4 (0.7)	133 (49.3)	1 (2.3)
Class II or V nephritis	194 (9.8)	1 (0.1)	156 (9.0)	0 (0)	38 (15.0)	1 (0.6)	127 (9.3)	1 (0.3)	111 (8.7)	0 (0)	37 (13.7)	0 (0)
Class III or IV nephritis	288 (14.5)	0 (0)	235 (13.6)	0 (0)	53 (20.9)	0 (0)	201 (14.8)	0 (0)	168 (13.2)	0 (0)	40 (14.8)	0 (0)
Antiphospholipid antibodies	567 (28.6)	6 (0.8)	503 (29.1)	6 (1.0)	64 (25.3)	0 (0)	402 (29.5)	2 (0.6)	418 (32.8)	6 (1.0)	50 (18.5)	0 (0)
Low C3 or low C4	1002 (50.6)	12 (1.5)	878 (50.9)	12 (2.0)	124 (49.0)	0 (0)	685 (50.3)	6 (1.7)	580 (45.5)	10 (1.7)	144 (53.3)	0 (0)
Low C3 and low C4	460 (23.2)	2 (0.3)	397 (23.0)	1 (0.2)	63 (24.9)	1 (0.6)	319 (23.4)	13 (3.6)	271 (21.2)	1 (0.2)	68 (25.2)	1 (2.3)
Anti-dsDNA or anti-Smith antibody	1424 (71.9)	25 (3.2)	1249 (72.4)	22 (3.6)	174 (68.8)	3 (1.7)	971 (71.3)	13 (3.6)	830 (65.0)	15 (2.6)	229 (84.8)	3 (6.8)

Con.: controls; SLE: systemic lupus erythematosus.

both sensitivity and specificity simultaneously. Unlike predictive values, LRs are not affected by disease prevalence. The magnitude of the LR gives intuitive meaning as to how strongly a test result will raise (rule in) or lower (rule out) the likelihood of a disease. Cui and colleagues[17] inquired about the probability of illness in patients who have signs or symptoms suggestive of, but not diagnostic of, SLE and a classification criteria score of less than 10. Others may suggest a higher threshold score to increase specificity.[16] Our reporting of LRs across a range of thresholds is informative and may address these concerns. The pretest odds of a diagnosis, multiplied by the LR, determines the posttest odds based on the Bayes theorem.[24] When the pretest probability lies between 30% and 70%, test results with very high +LRs (eg, above 10) rule in disease. A very low –LR (eg, below 0.1) rules out that the patient has the disease. Our findings of small –LRs indicate that failure to fulfill the classification criteria at a threshold of 10 makes the posttest probability of having SLE very small, in the case of a patient with signs or symptoms

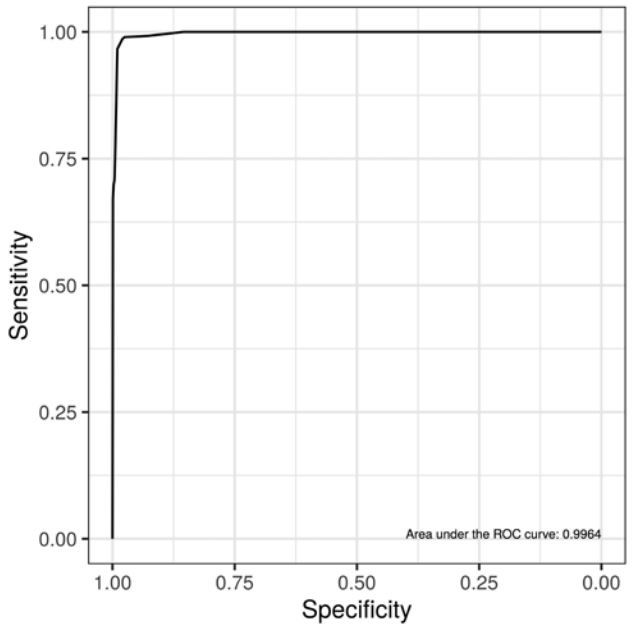

Figure 1. Receiver-operating characteristic (ROC) curve for the full cohort.

suggestive of, but not diagnostic of, SLE. The high +LR indicates that the classification criteria perform well for a population. The lower estimates of the EULAR/ACR SLE criteria specificity may have been underestimated because of incomplete use of the attribution rule of the criteria, under which items should only be counted for SLE if there were no more likely alternative explanations.[25,26] Readers, however, are cautioned: classification criteria are not designed for diagnosis and should not be used as justification to withhold appropriate treatment.[3]

A strength of this study is the large number of cases and controls. This allowed for good precision around our estimates. The criterion of "arthritis" has been noted to be problematic across all iterations of SLE classification criteria, as patients with SLE can have an erosive arthritis.[27] Our control group contained several forms of arthritis, including rheumatoid arthritis, and the operating characteristics of the classification criteria remained strong. Similarly, our control group contained several dermatologic conditions, single dominant manifestations that can be seen in SLE cases and in nonrheumatic disease controls. A limitation of our study is the number of ethnicities represented. External validation is needed across other ethnicities, including South Asian, First Nations/Indigenous, and others. Another limitation to consider is the lack of a gold standard for SLE. Ideally, classification criteria should be compared against a gold standard. However, SLE is a disease without a single diagnostic test or gold standard. All SLE cases in this study fulfilled either the ACR or SLICC criteria. However, comparing one system of classification to another can be a limitation. For example, if a patient is classified as having SLE based on the EULAR/ACR criteria but not on the modified 1997 ACR classification, it remains uncertain whether the EULAR/ACR criteria is a false positive or whether the 1997 ACR criteria missed a true SLE case. In the absence of a true gold standard, alternative strategies could include physician diagnosis or consensus methodology as the gold standard. Physician diagnosis can be a problematic gold standard, as the construct underlying SLE varies across individuals and centers.[4,28] Using consensus methods for case ascertainment is labor intensive and is biased toward established, clear-cut disease. Indicating that all SLE cases in this study fulfilled either the ACR or SLICC criteria facilitates comparison of our findings to other studies that use the same criteria.[2] Comparison of the new criteria against the ACR and SLICC criteria was required by both the ACR and EULAR for consideration of endorsement.[4,5]

Operating characteristics are a critical feature of classification criteria. Our reporting of LRs demonstrates the value of the criteria across important patient subgroups and further validates the threshold score of 10. The operating characteristics of the EULAR/ACR classification criteria are robust, reducing the risk of misclassification.

REFERENCES

1. Barber MRW, Johnson SR, Gladman DD, Clarke AE, Bruce IN. Evolving concepts in systemic lupus erythematosus damage assessment. Nat Rev Rheumatol 2021;17:307-8.

Table 4. Operating characteristics of the EULAR/ACR SLE classification criteria.

	Sensitivity, %, (95% CI)	Specificity, %, (95% CI)	Positive Likelihood Ratio (95% CI)	Negative Likelihood Ratio (95% CI)
Total cohort	98 (97-98)	99 (98-100)	95.5 (48.0-190)	0.03 (0.02-0.03)
Sex				
Female	98 (97-98)	99 (98-100)	99.1 (44.7-219.8)	0.03 (0.02-0.03)
Male	98 (95-99)	99 (96-100)	84.9 (21.4-337)	0.02 (0.01-0.05)
Ethnicity				
White	96 (95-97)	99 (98-100)	112.7 (47.1-269.7)	0.04 (0.03-0.05)
Black	99 (97-100)	100 (92-100)	Not able to estimate	0.01 (0.00-0.03)
Chinese	100 (98-100)	98 (87-100)	42 (6.1-291.2)	Not able to estimate
Filipino	100 (95-100)	93 (66-100)	14 (2.1-92.6)	Not able to estimate
Disease duration				
≤ 5 yrs	97 (96-98)	99 (98-100)	174.40 (43.78-694.64)	0.03 (0.02-0.04)

ACR: American College of Rheumatology; EULAR: European Alliance of Associations for Rheumatology; SLE: systemic lupus erythematosus

Table 5. Likelihood of accurate classification of SLE in the full cohort and patient subsets.

Score[a]	Total Cohort, N = 2764		Women, n = 2337		Men, n = 427		Early Disease, ≤ 5 Years Duration, n = 1719		White, n = 186		Black, n = 314	
	+LR[b] (95% CI)	–LR[b] (95% CI)	+LR (95% CI)	–LR (95% CI)	+LR (95% CI)	–LR (95% CI)	+LR (95% CI)	–LR (95% CI)	+LR (95% CI)	–LR (95% CI)	+LR (95% CI)	–LR (95% CI)
6	3.48 (3.12-3.89)	0.01 (0.00-0.01)	3.32 (2.94-3.75)	0.01 (0.00-0.01)	4.23 (3.23-5.52)	0.01 (0.00-0.04)	4.58 (3.76-5.57)	0.00 (0.00-0.01)	3.48 (3.06-3.95)	0.01 (0.00-0.02)	2.10 (1.54-2.85)	NA
7	29.85 (20.45-43.56)	0.01 (0.01-0.02)	27.44 (18.20-41.36)	0.01 (0.01-0.02)	43.16 (16.38-113.70)	0.01 (0.00-0.03)	25.33 (15.16-42.33)	0.01 (0.01-0.02)	35.96 (22.18-58.30)	0.02 (0.01-0.02)	22.00 (5.68-85.21)	NA
8	32.29 (21.77-47.88)	0.01 (0.01-0.02)	30.16 (19.60-46.42)	0.01 (0.01-0.02)	42.81 (16.28-112.80)	0.02 (0.01-0.04)	29.50 (16.92-51.46)	0.01 (0.01-0.02)	38.29 (23.24-63.11)	0.02 (0.01-0.03)	21.92 (5.66-84.89)	0.00 (0.00-0.03)
9	69.83 (38.83-125.58)	0.02 (0.02-0.03)	66.44 (34.74-127.09)	0.02 (0.01-0.03)	84.94 (21.41-336.96)	0.02 (0.01-0.05)	116.97 (37.90-360.95)	0.02 (0.01-0.03)	81.07 (38.82-169.32)	0.03 (0.02-0.04)	NA	0.01 (0.00-0.03)
10	95.57 (47.96-190.44)	0.03 (0.02-0.03)	99.13 (44.71-219.80)	0.03 (0.02-0.03)	84.94 (21.41-336.96)	0.02 (0.01-0.05)	174.40 (43.78-694.64)	0.03 (0.02-0.04)	112.68 (47.07-269.72)	0.04 (0.03-0.05)	NA	0.01 (0.00-0.03)
11	106.17 (50.78-221.99)	0.05 (0.04-0.06)	96.31 (43.43-213.54)	0.05 (0.04-0.07)	165.75 (23.48-1170.26)	0.05 (0.03-0.08)	170.06 (42.69-677.37)	0.05 (0.04-0.06)	136.27 (51.31-361.91)	0.07 (0.05-0.08)	NA	0.03 (0.02-0.06)
12	120.90 (54.48-268.31)	0.08 (0.06-0.09)	112.95 (47.18-270.42)	0.07 (0.06-0.09)	160.25 (22.69-1131.57)	0.08 (0.05-0.12)	165.45 (41.54-659.05)	0.08 (0.06-0.09)	175.75 (56.84-543.42)	0.10 (0.08-0.12)	NA	0.04 (0.03-0.08)

[a] EULAR/ACR SLE classification criteria threshold score; scores of 6 to 12 were evaluated in subjects with early disease (disease duration of ≤ 5 years). [b] LRs > 10 and < 0.1 are considered to provide strong evidence to rule in or rule out diagnoses, respectively, in most circumstances. +LR: positive likelihood ratio; –LR: negative likelihood ratio; ACR: American College of Rheumatology; EULAR: European Alliance of Associations for Rheumatology; NA: not applicable, because of either perfect sensitivity or perfect specificity; SLE: systemic lupus erythematosus.

2. Johnson SR, Goek ON, Singh-Grewal D, et al. Classification criteria in rheumatic diseases: a review of methodologic properties. Arthritis Rheum 2007;57:1119-33.
3. Aggarwal R, Ringold S, Khanna D, et al. Distinctions between diagnostic and classification criteria? Arthritis Care Res 2015;67:891-7.
4. 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.
5. 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.
6. Leuchten N, Hoyer A, Brinks R, et al. Performance of antinuclear antibodies for classifying systemic lupus erythematosus: a systematic literature review and meta-regression of diagnostic data. Arthritis Care Res 2018;70:428-38.
7. Leuchten N, Milke B, Winkler-Rohlfing B, et al. Early symptoms of systemic lupus erythematosus (SLE) recalled by 339 SLE patients. Lupus 2018;27:1431-6.
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.
9. Tedeschi SK, Johnson SR, Boumpas DT, et al. Multicriteria decision analysis process to develop new classification criteria for systemic lupus erythematosus. Ann Rheum Dis 2019;78:634-40.
10. 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.
11. Aringer M, Brinks R, Dörner T, et al. European League Against Rheumatism (EULAR)/American College of Rheumatology (ACR) SLE classification criteria item performance. Ann Rheum Dis 2021; 80:775-81.
12. Johnson SR, Brinks R, Costenbader KH, et al. Performance of the 2019 EULAR/ACR classification criteria for systemic lupus erythematosus in early disease, across sexes and ethnicities. Ann Rheum Dis 2020;79:1333-9.
13. Guavita-Navarro D, Gallego-Cardona L, Arredondo AM, et al. Comparison of the sensitivity of the EULAR / ACR 2019 and SLICC 2012 classification criteria in a Colombian population with systemic lupus erythematosus. J Transl Autoimmun 2021;4:100133.
14. Magallares B, Lobo-Prat D, Castellví I, et al. Assessment of EULAR/ACR-2019, SLICC-2012 and ACR-1997 classification criteria in SLE with longstanding disease. J Clin Med 2021;10:2377.
15. Tao JJ, Hiraki LT, Levy DM, Silverman ED. Comparison of sensitivities of American College of Rheumatology and Systemic Lupus International Collaborating Clinics classification criteria in childhood-onset systemic lupus erythematosus. J Rheumatol 2019;46:731-8.
16. Rodrigues Fonseca A, Felix Rodrigues MC, Sztajnbok FR, Gerardin Poirot Land M, Knupp Feitosa de Oliveira S. Comparison among ACR1997, SLICC and the new EULAR/ACR classification criteria in childhood-onset systemic lupus erythematosus. Adv Rheumatol 2019;59:20.
17. Cui R, Wang Q, Zhang H, Wu S, Wan XJ, Dai SM. Correspondence on '2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus' by Aringer et al. Ann Rheum Dis 2020;81:e165.
18. Adamichou C, Nikolopoulos D, Genitsaridi I, et al. In an early SLE cohort the ACR-1997, SLICC-2012 and EULAR/ACR-2019 criteria classify non-overlapping groups of patients: use of all three criteria ensures optimal capture for clinical studies while their modification earlier classification and treatment. Ann Rheum Dis 2020;79:232-41.
19. Whittall Garcia LP, Gladman DD, Urowitz M, Touma Z, Su J, Johnson SR. New EULAR/ACR 2019 SLE classification criteria: defining ominosity in SLE. Ann Rheum Dis 2021;80:767-74.
20. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40:1725.
21. Petri M, Orbai AM, 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.
22. van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737-47.
23. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. J Clin Epidemiol 1991;44:763-70.
24. Johnson SR, Tomlinson GA, Granton JT, Hawker GA, Feldman BM. Applied Bayesian methods in the rheumatic diseases. Rheum Dis Clin North Am 2018;44:361-70.
25. Aringer M, Costenbader K, Johnson SR. Assessing the EULAR/ACR classification criteria for patients with systemic lupus erythematosus. Expert Rev Clin Immunol 2022;18:135-44.
26. Aringer M, Johnson SR. New lupus criteria: a critical view. Curr Opin Rheumatol 2021;33:205-10.
27. Aringer M, Petri M. New classification criteria for systemic lupus erythematosus. Curr Opin Rheumatol 2020;32:590-6.
28. Johnson SR, Gladman DD, Brunner HI, et al. Evaluating the construct of damage in systemic lupus erythematosus. Arthritis Care Res 2021 Dec 28 (Epub ahead of print).