



Published in final edited form as:

Arthritis Care Res (Hoboken). 2021 September ; 73(9): 1231–1235. doi:10.1002/acr.24263.

Comparison of the 2019 European Alliance of Associations for Rheumatology/American College of Rheumatology Systemic Lupus Erythematosus Classification Criteria With Two Sets of Earlier Systemic Lupus Erythematosus Classification Criteria

Michelle Petri¹, Daniel W. Goldman¹, Graciela S. Alarcón², Caroline Gordon³, Joan T. Merrill⁴, Paul R. Fortin⁵, Ian N. Bruce⁶, David Isenberg⁷, Daniel Wallace⁸, Ola Nived⁹, Rosalind Ramsey-Goldman¹⁰, Sang-Cheol Bae¹¹, John G. Hanly¹², Jorge Sanchez-Guerrero¹³, Ann E. Clarke¹⁴, Cynthia Aranow¹⁵, Susan Manzi¹⁶, Murray Urowitz¹³, Dafna D. Gladman¹³, Ken Kalunian¹⁷, Victoria P. Werth¹⁸, Asad Zoma¹⁹, Sasha Bernatsky²⁰, Munther Khamashta²¹, Søren Jacobsen²², Jill P. Buyon²³, Mary Anne Dooley²⁴, Ronald van Vollenhoven²⁵, Ellen Ginzler²⁶, Thomas Stoll²⁷, Christine Peschken²⁸, Joseph L. Jorizzo²⁹, Jeffery P. Callen³⁰, Sam Lim³¹, Murat Inanç³², Diane L. Kamen³³, Anisur Rahman⁷, Kristjan Steinsson³⁴, Andrew G. Franks Jr²³, Laurence S. Magder³⁵

¹Johns Hopkins University, Baltimore, Maryland;

²University of Alabama at Birmingham;

³University of Birmingham, Birmingham, UK;

⁴Oklahoma Medical Research Foundation, Oklahoma City;

⁵CHU de Québec – Université Laval, Quebec City, Quebec, Canada;

⁶The University of Manchester and Manchester University Hospital NHS Foundation Trust, Manchester Academic Health Science Center, Manchester, UK;

⁷University College, London, UK;

⁸Cedars-Sinai and University of California, Los Angeles;

⁹Lund University, Lund, Sweden;

¹⁰Northwestern University, Chicago, Illinois;

¹¹Hanyang University Hospital for Rheumatic Diseases, Seoul, Republic of Korea;

Address correspondence to Michelle Petri, MD, MPH, Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, MD 21205. mpetri@jhmi.edu.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Petri had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Petri, Magder.

Acquisition of data. Petri, Goldman, Alarcón, Gordon, Merrill, Fortin, Bruce, Isenberg, Wallace, Nived, Ramsey-Goldman, Bae, Hanly, Sanchez-Guerrero, Clarke, Aranow, Manzi, Urowitz, Gladman, Kalunian, Werth, Zoma, Bernatsky, Khamashta, Jacobsen, Buyon, Dooley, van Vollenhoven, Ginzler, Stoll, Peschken, Jorizzo, Callen, Lim, Inanç, Kamen, Rahman, Steinsson, Franks, Magder.

Analysis and interpretation of data. Petri, Magder.

No potential conflicts of interest relevant to this article were reported.

¹²Queen Elizabeth II Health Sciences Centre and Dalhousie University, Halifax, Nova Scotia, Canada;

¹³Centre for Prognosis Studies in the Rheumatic Diseases, Toronto Western Hospital and University of Toronto, Toronto, Ontario, Canada;

¹⁴University of Calgary, Calgary, Alberta, Canada;

¹⁵Feinstein Institute for Medical Research, Manhasset, New York;

¹⁶Allegheny Health Network, Pittsburgh, Pennsylvania;

¹⁷University of California at San Diego, La Jolla;

¹⁸Hospital of the University of Pennsylvania and Department of Veterans Affairs Medical Center, Philadelphia;

¹⁹Lanarkshire Centre for Rheumatology, Hairmyres Hospital, East Kilbride, Scotland, UK;

²⁰McGill University Health Centre, Montreal, Quebec, Canada;

²¹The Rayne Institute, St Thomas' Hospital, King's College London, London, UK;

²²Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark;

²³New York University, New York, New York;

²⁴University of North Carolina, Chapel Hill;

²⁵Amsterdam University Medical Centres, University of Amsterdam, Free University, and Amsterdam Rheumatology and Immunology Center, Amsterdam, The Netherlands;

²⁶State University of New York Downstate Medical Center, Brooklyn;

²⁷Kantonsspital, Schaffhausen, Switzerland;

²⁸University of Manitoba, Winnipeg, Manitoba, Canada;

²⁹Wake Forest University, Winston-Salem, North Carolina, and Weill Cornell Medicine, New York, New York;

³⁰University of Louisville, Louisville, Kentucky;

³¹Emory University, Atlanta, Georgia;

³²Istanbul University, Istanbul, Turkey;

³³Medical University of South Carolina, Charleston;

³⁴Fossvogur Landspítali University Hospital, Reykjavik, Iceland;

³⁵University of Maryland, Baltimore.

Abstract

Objective.—The Systemic Lupus International Collaborating Clinics (SLICC) 2012 systemic lupus erythematosus (SLE) classification criteria and the revised American College of Rheumatology (ACR) 1997 criteria are list based, counting each SLE manifestation equally. We derived a classification rule based on giving variable weights to the SLICC criteria and compared

its performance to the revised ACR 1997, the unweighted SLICC 2012, and the newly reported European Alliance of Associations for Rheumatology (EULAR)/ACR 2019 criteria sets.

Methods.—The physician-rated patient scenarios used to develop the SLICC 2012 classification criteria were reemployed to devise a new weighted classification rule using multiple linear regression. The performance of the rule was evaluated on an independent set of expert-diagnosed patient scenarios and compared to the performance of the previously reported classification rules.

Results.—The weighted SLICC criteria and the EULAR/ACR 2019 criteria had less sensitivity but better specificity compared to the list-based revised ACR 1997 and SLICC 2012 classification criteria. There were no statistically significant differences between any pair of rules with respect to overall agreement with the physician diagnosis.

Conclusion.—The 2 new weighted classification rules did not perform better than the existing list-based rules in terms of overall agreement on a data set originally generated to assess the SLICC criteria. Given the added complexity of summing weights, researchers may prefer the unweighted SLICC criteria. However, the performance of a classification rule will always depend on the populations from which the cases and non-cases are derived and whether the goal is to prioritize sensitivity or specificity.

INTRODUCTION

The Systemic Lupus International Collaborating Clinics (SLICC) classification criteria for SLE were derived from a set of 702 expert-rated patient scenarios (1). Recursive partitioning was used to derive an initial rule that was simplified and refined based on SLICC physician consensus. The SLICC group then validated the classification criteria on a new validation sample of 690 expert-rated patient scenarios (1). In previous validation work, the SLICC 2012 systemic lupus erythematosus (SLE) classification criteria (1) were more sensitive than the American College of Rheumatology (ACR) 1997 revised criteria (2,3) but less specific. Subsequent studies in other cohorts confirmed these conclusions (4–6).

The 1997 ACR revised criteria and the SLICC 2012 criteria counted each SLE manifestation equally with 1 exception: the SLICC criteria counted lupus nephritis by biopsy as a standalone sufficient for classification. However, when physicians evaluate a patient for SLE, they may give greater weight to some nonrenal criteria over other nonrenal criteria. Therefore, we hypothesized that a classification score that gave greater weight to some nonrenal criteria than to others might have greater agreement with physician diagnosis. Therefore, our objective was to derive and test a classification rule that differentially weighted the variables used in the SLICC classification rule. We then compared this rule to the European Alliance of Associations for Rheumatology (EULAR)/ACR 2019 classification rule (7) that used a weighted approach. We also compared the revised ACR 1997 and the original SLICC classification rule to the new EULAR/ACR 2019 classification rule.

PATIENTS AND METHODS

The physician-rated patient scenarios used to develop the SLICC 2012 classification criteria (the training set) were reemployed to devise a weighted classification rule (1). Briefly,

these were based on patients with a clinical diagnosis of SLE (n = 293) or non-SLE (n = 423: rheumatoid arthritis [RA; n = 119], myositis n = [55], chronic cutaneous lupus erythematosus [n = 50], undifferentiated connective tissue disease [n = 44], vasculitis [n = 37], primary antiphospholipid syndrome [n = 33], scleroderma [n = 28], fibromyalgia [n = 25], Sjögren's syndrome [n = 15], rosacea [n = 8], psoriasis [n = 7], sarcoidosis [n = 1], and juvenile inflammatory arthritis [n = 1]).

These patient scenarios were then classified as either SLE or non-SLE based on ratings by 32 SLICC rheumatologists. Based on these scenarios and ratings, the SLICC group developed the SLICC 2012 classification criteria.

To derive the new weighted SLICC classification rule, a multiple linear regression model was fit to these data using the SLICC 2012 criteria variables as predictors and the binary outcome (physician classification of SLE, the gold standard) as the outcome. To generate the weights for each criterion, we then multiplied each criterion's coefficient by 100 and rounded to the nearest integer. The direct Coombs criterion was not included in the weighted score because its weight was very small. The weights for the remaining SLICC 2012 manifestations and for the EULAR/ACR 2019 manifestations are shown in Table 1 (7).

A cutoff for classification was chosen as the score that maximized the sum of sensitivity and specificity of the new weighted criteria with physician diagnosis. We evaluated the performance of these weighted SLICC criteria on the independent validation set of patient scenarios collected by SLICC investigators to validate the SLICC 2012 classification rule. These patient scenarios were collected and rated in a similar manner to those used in the derivation step. As in the training set, there were SLE (n = 337) and non-SLE scenarios (n = 353: RA [n = 118], undifferentiated connective tissue disease [n = 89], primary antiphospholipid antibody syndrome [n = 30], vasculitis [n = 29], chronic cutaneous lupus [n = 24], scleroderma [n = 20], Sjögren's syndrome [n = 15], myositis [n = 14], psoriasis [n = 8], fibromyalgia [n = 4], alopecia areata [n = 1], and sarcoidosis [n = 1]). We then compared the performance of the newly derived weighted rule to the performance of the revised ACR 1997 criteria, the previous SLICC 2012 criteria, and the new EULAR/ACR 2019 criteria. In classifying patients based on the EULAR/ACR criteria, we did not include fever because that variable was not in our data set. In addition, because we did not have information about biopsy class, all those with lupus nephritis were given the maximum number of lupus nephritis points in computing the EULAR/ACR 2019 score. To address the issue of the omission of antinuclear antibody (ANA)-negative lupus from the EULAR/ACR 2019 criteria, we did an additional subset analysis omitting the ANA-negative SLE and disease controls.

All patients gave informed written consent to participate in the study. The study was approved by the Johns Hopkins University School of Medicine Institutional Review Board and complied with the Helsinki Declaration.

RESULTS

A new modification of the SLICC 2012 criteria, assigning a weight to each criterion, was developed. The weights derived for each criterion are shown in Table 1, and are juxtaposed to the weights in the EULAR/ACR 2019 criteria. The weights for the SLICC criteria were derived by statistical modeling and did not reflect physician judgment.

Using physicians' diagnosis as the gold standard, we determined that sensitivity and specificity were optimized when patients were classified as having SLE if they had lupus nephritis and/or if on the new weighted criteria they achieved 56 points with at least 1 clinical component and 1 immunologic component

Table 2 shows the performance of the 4 classification rules. As can be seen, the highest sensitivity was achieved by the SLICC 2012 criteria, whereas the revised ACR 1997 criteria had the highest specificity. The new weighted SLICC criteria and the EULAR/ACR 2019 classification criteria had intermediate specificity and sensitivity. All 4 sets of criteria had similar overall agreement with the physician diagnosis without statistically significant differences (at the 0.05 level) between any pair of rules. Table 3 shows the additional analysis omitting ANA-negative SLE and disease controls from the SLICC data set. In general, the sensitivities were similar to those in Table 2, but the specificities were better. There was significantly better performance of the SLICC criteria than the weighted SLICC criteria ($P = 0.0065$) or the revised ACR criteria ($P = 0.035$).

DISCUSSION

The existence of the SLICC patient scenarios gave us a unique opportunity to study the performance of the new EULAR/ACR 2019 classification criteria and determine whether weighting criteria made a difference. First, the EULAR/ACR 2019 criteria were reported to be more sensitive/specific than the SLICC 2012 criteria in the EULAR/ACR validation phase (7). There was no difference in our current study in overall agreement between the revised ACR 1997, the SLICC 2012, or the EULAR/ACR 2019 criteria. The performance obviously depends on the controls or non-SLE comparison cases. In the case of the SLICC data set, most of the non-SLE cases were patients with autoimmune diseases in which a positive ANA 1:80 (the entry criterion for EULAR/ACR 2019 criteria) would be common.

Second, weighting did not improve the performance of the SLICC 2012 criteria. The weighted SLICC classification rules did not perform better than the existing list-based SLICC rules in terms of overall agreement. In particular, the original SLICC classification criteria already heavily weighted lupus nephritis as a stand-alone criterion. Therefore, we do not recommend use of the weighted SLICC criteria.

Third, for some manifestations, the relative weights used in the weighted SLICC classification rule differed strikingly from the weights in the EULAR/ACR 2019 criteria. For example, in computing the weighted SLICC rule, oral ulcers were assigned a substantially higher weight than arthritis, whereas in the EULAR/ACR 2019 criteria this was not the case. This highlights the different approaches used to generate the weights. The SLICC weights were generated by a statistical analysis based on a set of SLE and non-SLE patient

scenarios, whereas the weights for EULAR/ ACR 2019 criteria were largely derived based on expert opinion. This also highlights the importance of the populations used to generate the validation data sets. The control group in the SLICC validation set consisted of patients with other rheumatic diseases. Many of these control patients also had arthritis, so the importance of arthritis for distinguishing SLE patients was attenuated. However, fewer of the non-SLE patients had oral ulcers, so the importance for distinguishing SLE from non-SLE was increased.

Fourth, 1 essential difference in the SLICC patient scenarios was that all cases and controls had the same autoantibody assays done in a central laboratory, such that there were almost no missing data. For example, in clinical practice, patients with RA might not have complement checked, or SLE patients might not have IgA anticardiolipin or IgA anti- β -2 glycoprotein checked, if the local laboratory was unable to perform the assay. If missing data were counted as negative, the EULAR/ACR criteria might appear to have better discrimination.

Fifth, the original SLICC 2012 criteria and the weighted SLICC 2012 criteria allowed for ANA-negative SLE. In our analysis that omitted ANA-negative SLE and disease controls, the SLICC 2012 criteria had a higher kappa than the other criteria. Given assay variability and the existence of true ANA-negative SLE (8), questions have been raised over whether ANA positivity should be employed to determine eligibility for clinical trials (9). In particular, ANA-negative lupus can include lupus nephritis (biopsy proven), which would always be classified as SLE using the SLICC 2012 criteria (10,11).

Sixth, we compared the various classification rules to the set of patient scenarios originally generated to validate the SLICC rule. One caveat is that this set was gathered and rated by the same investigators who originally generated the SLICC classification criteria. Thus, while statistically independent of the data used to generate the SLICC rule, some of the same opinions may have been driving the ratings of both the training and validation set. This is both a limitation (in comparing original SLICC to weighted SLICC) but also a potential strength in comparing the EULAR/ACR 2019 criteria to the SLICC 2012 criteria.

Seventh, while the overall performance of the 4 classification rules did not differ significantly, components of agreement (sensitivity and specificity) differed by as much as 10–14 percentage points in some comparisons. Thus, the choice of a classification rule might depend on whether the researcher wanted to cast a wider net or to reduce the risk of false positives. It should also be noted that for any score, the tradeoff between sensitivity and specificity can be altered by choosing a different cutoff.

In conclusion, we modified the SLICC 2012 classification criteria for SLE by including weighting and compared these modified criteria with the established criteria and with the (weighted) EULAR/ACR 2019 criteria. All 4 sets of criteria performed well when using the physicians' diagnosis as the reference without statistically significant or clinically convincing differences. However, the performance of these rules may vary in different populations or with different choices of control patients. We therefore recommend any of these criteria for SLE classification purposes.

Funding Support:

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant R01-AR-069572).

REFERENCES

1. Petri M, Orbai AM, Alarcón GS, Gordon C, Merrill JT, Fortin PR, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum* 2012;64:2677–86. [PubMed: 22553077]
2. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982;25:1271–7. [PubMed: 7138600]
3. Hochberg MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology Revised criteria for the classification of Systemic Lupus Erythematosus [letter]. *Arthritis Rheum* 1997;40:1725.
4. Ighe A, Dahlström Ö, Skogh T, Sjöwall C. Application of the 2012 Systemic Lupus International Collaborating Clinics classification criteria to patients in a regional Swedish systemic lupus erythematosus register. *Arthritis Res Ther* 2015;17:3. [PubMed: 25575961]
5. Sag E, Tartaglione A, Batu ED, Ravelli A, Khalil SM, Marks SD, et al. Performance of the new SLICC classification criteria in childhood systemic lupus erythematosus: a multicentre study. *Clin Exp Rheumatol* 2014;32:440–4. [PubMed: 24642380]
6. Fonseca AR, Gaspar-Elsas MI, Land MG, de Oliveira SK. Comparison between three systems of classification criteria in juvenile systemic lupus erythematosus. *Rheumatology (Oxford)* 2015;54:241–7. [PubMed: 25125590]
7. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. *Arthritis Rheumatol* 2019;71: 1400–12. [PubMed: 31385462]
8. Choi MY, Clarke AE, St. Pierre Y, Hanly JG, Urowitz MB, Romero-Diaz J, et al. Antinuclear antibody–negative systemic lupus erythematosus in an international inception cohort. *Arthritis Care Res (Hoboken)* 2019;71:893–902. [PubMed: 30044551]
9. Pisetsky DS, Spencer DM, Lipsky PE, Rovin BH. Assay variation in the detection of antinuclear antibodies in the sera of patients with established SLE. *Ann Rheum Dis* 2018;77:911–3. [PubMed: 29440000]
10. Simmons SC, Smith ML, Chang-Miller A, Keddis MT. Antinuclear antibody-negative lupus nephritis with full house nephropathy: a case report and review of the literature. *Am J Nephrol* 2015;42:451–9. [PubMed: 26812129]
11. Tarazi M, Gaffney RG, Kushner CJ, Chakka S, Werth VP. Cutaneous lupus erythematosus patients with a negative antinuclear antibody meeting the American College of Rheumatology and/or Systemic Lupus International Collaborating Clinics criteria for systemic lupus erythematosus. *Arthritis Care Res (Hoboken)* 2019;71:1404–9. [PubMed: 31058462]

SIGNIFICANCE & INNOVATIONS

- In an independent, multinational cohort, the European Alliance of Associations for Rheumatology/ American College of Rheumatology (ACR) 2019 classification criteria did not perform significantly better than the ACR 1997 and Systemic Lupus International Collaborating Clinics 2012 classification criteria.
- The performance of classification rules depends on the populations from which the cases and non-cases are derived.

Table 1.

Weighting factors for manifestations scored in the 2012 Systemic Lupus International Collaborating Clinics (SLICC) criteria and the 2019 European Alliance of Associations for Rheumatology (EULAR)/American College of Rheumatology (ACR) criteria (ref. 7) *

Manifestation	2012 SLICC weighting factors	EULAR/ACR weighting factors
Acute cutaneous	26	Maximum score of either: 6 (acute cutaneous); 4 (subacute cutaneous); 4 (discoid lupus); 2 (oral ulcers); or 2 (alopecia)
Chronic cutaneous	12	
Oral ulcers	16	
Alopecia	9	
Arthritis	9	6 (arthritis)
Serositis	16	Maximum score of either 6 (acute pericarditis) or 5 (effusion)
Renal without biopsy	9	Maximum score of either: 10 (class III/IV nephritis); 8 (class II/V nephritis); or 4 (proteinuria 0.5 gm/ day)
Renal with biopsy	Automatically classified	
Neurologic	9	Maximum score of either: 5 (seizures); 3 (psychosis); or 2 (delirium)
Hemolytic anemia	1	Maximum score of either: 4 (autoimmune hemolysis); 4 (thrombocytopenia); or 3 (leukopenia)
Leukopenia or lymphopenia	14	
Thrombocytopenia	15	
ANA	17	Prerequisite 1:80
Anti-dsDNA	19	6 (anti-Sm) or 6 (anti- dsDNA)
Anti-Sm	16	
Antiphospholipid antibodies	8	2 (antiphospholipid antibodies)
Low complement	11	Maximum score of either 4 (low C3 and C4) or 3 (low C3 or C4)
Fever		2 (fever)

* To satisfy the weighted SLICC criteria, the patient had to have either biopsy-proven lupus nephritis, or a score of 56 with both clinical and immunologic manifestations. To satisfy the EULAR/ACR criteria, the patient had to be positive for antinuclear antibody (ANA), have a total score of 10, and have at least 1 clinical manifestation. Anti- dsDNA = anti-double-stranded DNA.

Table 2.

Sensitivity and specificity of 4 different systemic lupus erythematosus classification rules based on physician diagnoses of patient scenarios*

Classification rule	Sensitivity (n = 349)		Specificity (n = 341)		Overall agreement (n = 690)		Kappa (chance-adjusted agreement)
	No. (%)	95% CI	No. (%)	95% CI	No. (%)	95% CI	
Revised ACR 1997	290 (83)	79–87	326 (96)	93–98	616 (89)	87–92	0.79
SLICC 2012	340 (97)	96–99	288 (84)	81–88	628 (91)	89–93	0.82
EULAR/ACR 2019	317 (91)	88–94	302 (89)	85–92	619 (90)	87–92	0.79
Weighted SLICC 2012	310 (89)	86–92	304 (89)	86–92	614 (89)	87–91	0.78

* 95% CI = 95% confidence interval; ACR = American College of Rheumatology; EULAR = European Alliance of Associations for Rheumatology; SLICC = Systemic Lupus International Collaborating Clinics.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Sensitivity and specificity of 4 different systemic lupus erythematosus (SLE) classification rules, omitting antinuclear antibody–negative SLE and disease controls from the Systemic Lupus International Collaborating Clinics (SLICC) data set *

Classification rule	Sensitivity (n = 341)		Specificity (n = 197)		Overall agreement (n = 538)		Kappa (chance-adjusted agreement)
	No. (%)	95% CI	No. (%)	95% CI	No. (%)	95% CI	
Revised ACR 1997	284 (83)	79–87	184 (93)	90–97	468 (87)	84–90	0.73
SLICC 2012	334 (98)	96–99	154 (78)	72–84	488 (91)	88–93	0.79
EULAR/ACR 2019	317 (93)	90–96	158 (80)	75–86	475 (88)	86–91	0.74
Weighted SLICC 2012	306 (90)	86–93	162 (82)	77–88	468 (87)	84–91	0.72

* 95% CI = 95% confidence interval; ACR = American College of Rheumatology; EULAR = European Alliance of Associations for Rheumatology.