# A Comparison of 2019 EULAR/ACR SLE Classification Criteria with Two Sets of Earlier SLE Classification Criteria

**Running Title:** Comparison of Lupus Classification Criteria

Michelle Petri MD MPH[1], Daniel W Goldman PhD MCS[1], Graciela S Alarcón MD MPH[2], Caroline Gordon MA MD FRCP[3], Joan T Merrill MD[4], Paul R Fortin MD MPH FRCPC[5], Ian N Bruce MD FRCP[6], David Isenberg MD FRCP FAMS[7], Daniel Wallace MD FACP[8], Ola Nived MD[9], Rosalind Ramsey-Goldman MD[10], Sang-Cheol Bae MD PhD MPH[11], John G Hanly MD FRCPC[12], Jorge Sanchez-Guerrero MD MSc[13], Ann E Clarke MD MSc [14], Cynthia Aranow MD[15], Susan Manzi MD MPH[16], Murray Urowitz MD[13], Dafna D. Gladman MD FRCPC [13], Ken Kalunian MD[17], Victoria P Werth MD[18], Asad Zoma MD[19], Sasha Bernatsky MD PhD[20], Munther Khamashta MD PhD FRCP[21], Søren Jacobsen MD DMSc[22], Jill P. Buyon MD[23], Mary Anne Dooley MD MPH[24], Ronald van Vollenhoven MD PhD[25], Ellen Ginzler MD MPH[26], Thomas Stoll MD[27], Christine Peschken MD MSc[28], Joseph L. Jorizzo MD[29], Jeffery P. Callen MD[30], Sam Lim MD MPH[31], Murat Inanç MD[32], Diane L Kamen MD[33], Anisur Rahman PhD FRCP[7], Kristjan Steinsson MD[34], Andrew G. Franks Jr MD[35], Laurence S. Magder MPH PhD [36]

[1]Department of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

[2]Department of Medicine, Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, Alabama, USA

[3]Rheumatology Research Group, Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

[4]Department of Clinical Pharmacology, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA

[5]Division of Rheumatology, CHU de Québec - Université Laval, Quebec City, Canada

[6]Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, The University of Manchester and NIHR Manchester Biomedical Research Centre, Manchester University Hospital NHS Foundation Trust, Manchester Academic Health Science Center, Manchester, UK

[7]Centre for Rheumatology Research, University College, London, UK

[8]Cedars-Sinai/David Geffen School of Medicine at UCLA, Los Angeles, California, USA

[9]Department of Clinical Sciences Lund, Section of Rheumatology, Lund University, Lund, Sweden

[10]Northwestern University and Feinberg School of Medicine, Chicago, Illinois, USA

[11]Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul, Korea

[12]Division of Rheumatology, Department of Medicine and Department of Pathology, Queen Elizabeth II Health Sciences Centre and Dalhousie University, Halifax, Nova Scotia, Canada

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

[14]Division of Rheumatology, Cumming School of Medicine University of Calgary, Calgary, Alberta, Canada

[15]Feinstein Institute for Medical Research, Manhasset, New York, USA

[16]Autoimmunity Institute, Allegheny Health Network, Pittsburgh, Pennsylvania, USA

[17]UCSD School of Medicine, La Jolla, California, USA

[18]Division of Dermatology, Hospital of the University of Pennsylvania and the Veteran's Administration Medical Center, Philadelphia, PA 19104

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

[20]Division of Rheumatology, Department of Medicine, McGill University Health Centre, Montreal, Quebec, Canada

[21]Lupus Research Unit, The Rayne Institute, St Thomas' Hospital, King's College London School of Medicine, London, UK

[22]Copenhagen Lupus and Vasculitis Clinic, Centre for Rheumatology and Spine Diseases, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

[23]Division of Rheumatology, Department of Medicine, New York University School of Medicine, New York, NY 10016.

[24]Division of Rheumatology and Immunology, Department of Medicine, University of North Carolina, Chapel Hill, North Carolina, USA

[25]Amsterdam University Medical Centres, University of Amsterdam, Amsterdam, Netherlands; Free University (VU) Amsterdam, Amsterdam, Netherlands; Amsterdam Rheumatology and Immunology Center, Amsterdam, Netherlands

[26]Department of Medicine, SUNY Downstate Medical Center, Brooklyn, New York, USA

[27]Department of Rheumatology, Kantonsspital, Schaffhausen, Switzerland

[28]Department of Medicine and Community Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada

[29]Department of Dermatology, Wake Forest University School of Medicine and Weill Cornell Medicine, New York, New York, USA

[30]Department of Medicine, University of Louisville School of Medicine, Louisville, Kentucky USA

[31]Division of Rheumatology, Emory University School of Medicine, Atlanta, Georgia, USA

[32]Division of Rheumatology, Department of Internal Medicine, Istanbul Medical Faculty, Istanbul University, Istanbul, Turkey

[33]Division of Rheumatology, Medical University of South Carolina, Charleston, South Carolina, USA

[34]Department of Rheumatology, Center for Rheumatology Research Fossvogur Landspitali University Hospital, Reyjkavik, Iceland

[35]Division of Rheumatology, Department of Medicine, & The Department of Dermatology, New York University School of Medicine, New York, NY 10016

[36]Division of Biostatistics and Bioinformatics, Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore Maryland

**Corresponding Author:**

Michelle Petri M.D. M.P.H.

Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Email: mpetri@jhmi.edu

Telephone:  410-955-3823

Fax:  410-614-0498

**Keywords:** Systemic lupus erythematosus, Classification Criteria,

**Funding**: The Hopkins Lupus Cohort is funded by NIH R01 AR069572.

**Conflict of Interest**:  No relevant conflict of interest.

## Abstract

**Objective**: The Systemic Lupus International Collaborating Clinics (SLICC) 2012 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, unweighted SLICC 2012 and the newly reported European League Against Rheumatism (EULAR)/ACR 2019 criteria.

**Methods**: The physician-rated patient scenarios used to develop the SLICC 2012 classification criteria were re-employed 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**: 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 two new weighted classification rules did not perform better than the existing list-based rules in terms of overall agreement on a dataset 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.

## Significance and Innovations

- In an independent, multi-national cohort, the EULAR/ACR 2019 classification criteria did not perform significantly better than the ACR 1997 and SLICC 2012 classification criteria.
- The performance of classification rules depend on the populations from which the cases and non-cases are derived.

## Introduction

The Systemic Lupus International Collaborating Clinics (SLICC) classification criteria for SLE were derived from a set of 702 expert-rated patient scenarios. 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 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 one exception: the SLICC criteria counted lupus nephritis by biopsy as a "stand alone", sufficient for classification. However, when physicians evaluate a patient for SLE, they may give greater weight to some non-renal criteria over other non-renal criteria. Therefore, we hypothesized that a classification score that gave greater weight to some non-renal criteria than others might have greater agreement with physician diagnosis. Therefore, our objective was to derive and test a classification rule which differentially weighted the variables used in the SLICC classification rule. We then compared this rule to the European League Against Rheumatism (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 re-employed to devise a weighted classification rule (1). In brief, these were based on patients with a clinical diagnosis of SLE (n=293) or non-SLE (n=423: rheumatoid arthritis (119), myositis (55), chronic cutaneous lupus erythematosus (50), undifferentiated connective tissue disease (44), vasculitis (37), primary antiphospholipid syndrome (33), scleroderma (28), fibromyalgia (25), Sjögren's syndrome (15), rosacea (8), psoriasis (7), sarcoidosis (1) and juvenile inflammatory arthritis (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 (n=353: RA (118), undifferentiated connective tissue disease (89), primary antiphospholipid antibody syndrome (30), vasculitis (29), chronic cutaneous lupus (24), scleroderma (20), Sjögren's syndrome (15), myositis (14), psoriasis (8), fibromyalgia (4), alopecia areata (1), and sarcoidosis (1)) scenarios. 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 "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.

By using physicians' diagnosis as the gold standard, it was determined that sensitivity and specificity were optimized when patients were classified as SLE if they had lupus nephritis

and/or if on the new weighted criteria they achieved 56 points or more with at least one clinical component and one immunologic component.

Table 2 shows the performance of the four 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 four 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 dataset. 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 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 revised ACR 1997, 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 dataset, most of the non-SLE cases were

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 cases had oral ulcers, so the importance for distinguishing SLE from non-SLE was increased.

Fourth, one 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 rheumatoid arthritis might not have complement checked, or SLE patients might not have IgA anticardiolipin or IgA anti-

beta 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 that 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 SLICC 2012 criteria.

Seventh, while the overall performance of the four 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 cut-off.

In summary, 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 four 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.

## References

1. Petri M, Orbai A-M, 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;64:2677–2686.

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;25:1271–1277.

3. Hochberg MC. Updating the American College of Rheumatology Revised criteria for the classification of Systemic Lupus Erythematosus. *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.

5. Sag E, Tartaglione A, Batu ED, Ravelli A, Khalil SMA, Marks SD, et al. Performance of the new SLICC classification criteria in childhood systemic lupus erythematosus: a multicentre study. *Clin Exp Rheumatol* 2014;32:440–444.

6. Fonseca AR, Gaspar-Elsas MIC, Land MGP, Oliveira SKF de. Comparison between three systems of classification criteria in juvenile systemic lupus erythematous. *Rheumatology* 2015;54:241–247.

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. *Ann Rheum Dis* 2019;78:1151–1159.

8. Choi MY, Clarke AE, Pierre Y St., 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.

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;77:911–913.

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;42:451–459.

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;71:1404–1409.

## Tables

Table 1: Weighting factors for manifestations scored in the 2012 SLICC[1] and the 2019 EULAR/ACR[2] criteria (7)

| Manifestation | 2012 SLICC Weighting Factors | EULAR/ACR Weighting Factors |
|---|---|---|
| acute cutaneous | 26 | Maximum score of either: |
| chronic cutaneous | 12 | 6 - acute cutaneous |
| oral ulcers | 16 | 4 - subacute cutaneous |
| alopecia | 9 | 4 - discoid lupus |
| | | 2 - oral ulcers or |
| | | 2 - alopecia |
| arthritis | 9 | 6 arthritis |
| serositis | 16 | Maximum score of either: |
| | | 6 - acute pericarditis or |
| | | 5 - effusion |
| renal without biopsy | 9 | Maximum score of either: |
| renal with biopsy | automatically | 10 - class III/IV nephritis |
| | classified | 8 - class II/V nephritis or |
| | | 4 - proteinuria ≥0.5g/day |
| neurologic | 9 | Maximum score of either: |
| | | 5 - seizures |

| Manifestation | 2012 SLICC Weighting Factors | EULAR/ACR Weighting Factors |
|---|---|---|
| | | 3 - psychosis or |
| | | 2 - delirium |
| hemolytic anemia | 1 | Maximum score of either: |
| leukopenia or lymphopenia | 14 | 4 - autoimmune hemolysis |
| thrombocytopenia | 15 | 4 - thrombocytopenia or |
| | | 3 - leukopenia |
| ANA | 17 | pre-requisite 1:80 |
| anti-dsDNA | 19 | 6 - anti-Sm or |
| anti-Sm | 16 | 6 - anti-dsDNA |
| 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 |

[1]Note, to satisfy the weighted SLICC criteria, the patient had to have either biopsy-proven lupus nephritis, or a score of 56 or higher with both clinical and immunologic manifestations.

[2]Note, to satisfy the EULAR/ACR criteria, the patient had to be positive for ANA, have a total score of 10 or more, and have at least one clinical manifestation.

Table 2: Sensitivity and specificity of four different SLE 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) |
|---|---|---|---|---|---|---|---|
| | N (%) | 95% CI | N (%) | 95% CI | N (%) | 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 |

Table 3: Sensitivity and specificity of four different SLE classification rules, omitting ANA negative SLE and disease controls from the SLICC dataset.

| Classification Rule | Sensitivity (n=341) | | Specificity (n=197) | | Overall Agreement (n=538) | | Kappa (Chance-adjusted Agreement) |
|---|---|---|---|---|---|---|---|
| | N (%) | 95% CI | N (%) | 95% CI | N (%) | 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 |