



UNIVERSITY OF LEEDS

This is a repository copy of *2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/151919/>

Version: Accepted Version

Article:

Aringer, M, Costenbader, K, Daikh, D et al. (61 more authors) (2019) 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, 78 (9). pp. 1151-1159. ISSN 0003-4967

<https://doi.org/10.1136/annrheumdis-2018-214819>

(c) 2019, Author(s) (or their employer(s)). No commercial re-use. See rights and permissions. Published by BMJ. This is an author produced version of a paper published in the *Annals of the Rheumatic Diseases*. Uploaded in accordance with the publisher's self-archiving policy. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

2019 EULAR/ACR Classification Criteria for Systemic Lupus Erythematosus

Martin Aringer, Karen H. Costenbader, David I. Daikh, Ralph Brinks, Marta Mosca, Rosalind Ramsey-Goldman, Josef S. Smolen, David Wofsy, Dimitrios Boumpas, Diane L. Kamen, David Jayne, Ricard Cervera, Nathalie Costedoat-Chalumeau, Betty Diamond, Dafna D. Gladman, Bevra H. Hahn, Falk Hiepe, Søren Jacobsen, Dinesh Khanna, Kirsten Lerstrøm, Elena Massarotti, W. Joseph McCune, Guillermo Ruiz-Irastorza, Jorge Sanchez-Guerrero, Matthias Schneider, Murray B. Urowitz, George Bertias, Bimba F. Hoyer, Nicolai Leuchten, Chiara Tani, Sara K. Tedeschi, Zahi Touma, Gabriela Schmajuk, Branimir Anic, Florence Assan, Daniel Tak Mao Chan, Ann E. Clarke, Mary K. Crow, László Czirják, Andrea Doria, Winfried B. Graninger, Bernadett Halda-Kiss, Sarfaraz Hasni, Peter Izmirly, Michelle Jung, Gábor Kumánovics, Xavier Mariette, Ivan Padjen, José M. Pego-Reigosa, Juanita Romero-Díaz, Iñigo Rúa-Figueroa Fernández, Raphaële Seror, Georg Stummvoll, Yoshiya Tanaka, Maria G. Tektonidou, Carlos Vasconcelos, Edward M. Vital, Daniel Wallace, Sule Yavuz, Pier Luigi Meroni, Marvin J Fritzler, Ray P. Naden, Thomas Dörner, Sindhu R. Johnson.

Martin Aringer MD, University Medical Center and Faculty of Medicine Carl Gustav Carus, TU Dresden, Dresden, Germany

Karen H. Costenbader MD MPH, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

David I. Daikh MD, University of California at San Francisco and VA Medical Center, San Francisco, CA, USA

Ralph Brinks PhD, Policlinic and Hiller Research Unit for Rheumatology, Medical Faculty, Heinrich-Heine-University Duesseldorf, Duesseldorf, Germany

Marta Mosca MD PhD, Rheumatology Unit, Azienda Ospedaliero Universitaria Pisana, University of Pisa, Pisa, Italy

Rosalind Ramsey-Goldman MD DrPH, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Josef S. Smolen MD, Medical University of Vienna, Austria

David Wofsy MD, Russell/Engleman Rheumatology Research Center, University of California at San Francisco, San Francisco, USA

Dimitrios Boumpas MD, Medical School, National and Kapodestrian University of Athens, and Biomedical Research Foundation of the Athens Academy, Athens, Greece; Medical School, University of Cyprus, Nicosia, Cyprus.

Diane L. Kamen MD MSCR, Medical University of South Carolina, Charleston, SC, USA

David Jayne MD FRCP FRCPE FMedSci, Department of Medicine, University of Cambridge, United Kingdom

Ricard Cervera MD PhD FRCP, Department of Autoimmune Diseases, Hospital Clínic, University of Barcelona, Barcelona, Catalonia, Spain

Nathalie Costedoat-Chalumeau, MD PhD AP-HP, Cochin Hospital, Internal Medicine Department, Centre de référence maladies auto-immunes et systémiques rares d'île de France, Paris, France ; Université Paris Descartes-Sorbonne Paris Cité, Paris, France ; INSERM U 1153, Center for Epidemiology and Statistics Sorbonne Paris Cité (CRESS), Paris, France

Betty Diamond MD, Feinstein Institute, Manhasset, NY, United States

Dafna D. Gladman, MD FRCPC, Division of Rheumatology, Department of Medicine, Toronto Western Hospital, University of Toronto, Toronto, Ontario, Canada

Bevra H Hahn MD, University of California at Los Angeles, Los Angeles, CA, USA

Falk Hiepe MD, Charité – Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Department of Rheumatology and Clinical Immunology, Berlin, Germany

Søren Jacobsen MD DMSc, Copenhagen Lupus and Vasculitis Clinic, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Dinesh Khanna MD MS, University of Michigan, Ann Arbor, MI, USA

Kirsten Lerstrøm, Lupus Europe, co-opted trustee for research, Essex, UK

Elena Massarotti MD, Brigham and Women's Hospital, Boston MA; Harvard Medical School, Boston, USA

W. Joseph McCune MD, University of Michigan, Ann Arbor, MI, USA

Guillermo Ruiz-Irastorza MD, Autoimmune Diseases Research Unit, Department of Internal Medicine, Biocruces Bizkaia Health Research Institute, Hospital Universitario Cruces, UPV/EHU, Bizkaia, The Basque Country, Spain

Jorge Sanchez-Guerrero MD MSc, Division of Rheumatology, Department of Medicine Mount Sinai Hospital/University Health Network, University of Toronto, Toronto, Ontario, Canada; and Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico

Matthias Schneider MD, Policlinic and Hiller Research Unit for Rheumatology, Medical Faculty, Heinrich-Heine-University, Düsseldorf, Germany

Murray B. Urowitz MD, Division of Rheumatology, Department of Medicine, Toronto Western Hospital, University of Toronto, Toronto, Ontario, Canada

George Bertsias MD, Rheumatology, Clinical Immunology and Allergy, University of Crete Medical School, Heraklion, Greece

Bimba F. Hoyer MD, Charité – Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Department of Rheumatology and Clinical Immunology, Berlin, Germany and University of Schleswig-Holstein at Kiel, Kiel, Germany

Nicolai Leuchten MD, University Medical Center and Faculty of Medicine Carl Gustav Carus, TU Dresden, Dresden, Germany

Chiara Tani MD, Rheumatology Unit, Azienda Ospedaliero Universitaria Pisana, University of Pisa, Pisa, Italy

Sara K. Tedeschi MD MPH, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

Zahi Touma, MD PhD, Division of Rheumatology, Department of Medicine, Toronto Western Hospital, Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada

Gabriela Schmajuk MD MS, University of California at San Francisco and the VA Medical Center, San Francisco, USA

Branimir Anic MD, Division of Clinical Immunology and Rheumatology, University of Zagreb School of Medicine and University Hospital Centre Zagreb, Zagreb, Croatia

Florence Assan MD, Université Paris Sud, Hôpitaux Universitaires Paris-Sud, AP-HP, INSERM UMR 1184, Le Kremlin-Bicêtre, France

Tak Mao Chan MD, University of Hong Kong, Hong Kong

Ann E. Clarke MD MSc, Division of Rheumatology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

Mary K. Crow MD, Hospital for Special Surgery, New York, NY, USA

László Czirják MD, University of Pécs Medical School, Pécs, Hungary

Andrea Doria MD, Rheumatology Unit, Department of Medicine (DIMED), University of Padova, Padova, Italy

Winfried B. Graninger MD, Medical University of Graz, Graz, Austria

Bernadett Halda-Kiss MD, University of Pécs Medical School, Pécs, Hungary

Sarfraz Hasni MD, NIAMS, NIH, Bethesda, MD

Peter Izmirly MD, New York University School of Medicine, New York, New York, USA

Michelle Jung, University of Calgary, Calgary, Alberta, Canada

Gábor Kumánovics MD, University of Pécs Medical School, Pécs, Hungary

Xavier Mariette MD PhD, Université Paris Sud, Hôpitaux Universitaires Paris-Sud, AP-HP, INSERM UMR 1184, Le Kremlin-Bicêtre, France

Ivan Padjen MD, Division of Clinical Immunology and Rheumatology, University of Zagreb School of Medicine and University Hospital Centre Zagreb, Zagreb, Croatia

José M. Pego-Reigosa MD PhD, Department of Rheumatology, University Hospital of Vigo, IRIDIS Group, Instituto de Investigación Sanitaria Galicia Sur (IISGS), Vigo, Spain

Juanita Romero-Díaz MD, , MSc, Department of Immunology and Rheumatology, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico

Iñigo Rúa-Figueroa Fernández MD, Hospital Dr Negrin, Las Palmas, Spain

Raphaèle Seror MD, Université Paris Sud, Hôpitaux Universitaires Paris-Sud, AP-HP, INSERM UMR 1184, Le Kremlin-Bicêtre, France

Georg Stummvoll MD, Medical University of Vienna, Vienna, Austria

Yoshiya Tanaka MD PhD, University of Occupational & Environmental Health, Kitakyushu, Japan

Maria G. Tektonidou MD PhD, Medical School, National and Kapodistrian University of Athens, Athens, Greece

Carlos Vasconcelos MD PhD, Centro Hospitalar do Porto, ICBAS, University of Porto, Porto, Portugal

Edward M Vital MRCP PhD, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds; NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom

Daniel J. Wallace MD, Cedars-Sinai, Los Angeles, CA, USA

Sule Yavuz, Istanbul Bilim University, Istanbul, Turkey

Pier Luigi Meroni MD, Clinical Immunology and Rheumatology Unit, IRCCS Istituto Auxologico Italiano, Milan, Italy

Marvin J Fritzer PhD MD, Faculty of Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Ray P. Naden MB ChB FRACP, Department of Medicine, McMaster University, Hamilton, Ontario, Canada

Thomas Dörner MD, Charité – Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Department of Rheumatology and Clinical Immunology, Berlin, Germany

Sindhu R. Johnson MD PhD FRCPC, Division of Rheumatology, Department of Medicine, Toronto Western Hospital, Mount Sinai Hospital; Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada

Corresponding Authors

Martin Aringer MD, Division of Rheumatology, Department of Medicine III, University Medical Center TU Dresden, Fetscherstrasse 74, 01307 Dresden, Germany. Phone +49 351 458 4422 Fax +49 351 458 5801 E-mail martin.aringer@uniklinikum-dresden.de

Sindhu R. Johnson MD PhD, Division of Rheumatology, Ground Floor, East Wing, Toronto Western Hospital, 399 Bathurst Street, Toronto, Ontario, Canada, M5T 2S8. Phone 1-416-603-6417 Fax.1-416-603-4348. Email: Sindhu.Johnson@uhn.ca

Word Count 4,474

Acknowledgements. This body of work was jointly supported by the European League Against Rheumatism and the American College of Rheumatology. One part of the derivation and validation cohort was supported by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health. The authors wish to acknowledge the diligent work of Banita Aggarwal and Keshini Devakandan in data entry, data cleaning, queries to submitting investigators, data cutting and maintenance of the derivation and validation cohorts; and of Corine Sinnette, MA, in the preparatory work for the multicriteria decision analysis exercise.

Key words. Systemic lupus erythematosus, lupus, classification criteria, consensus methods, multi-criteria decision analysis, validation

ABSTRACT

Objective. To develop new classification criteria for systemic lupus erythematosus (SLE) jointly supported by the European League Against Rheumatism (EULAR) and the American College of Rheumatology (ACR).

Methods. This international initiative had 4 phases: 1) Evaluation of anti-nuclear antibody (ANA) as an entry criterion through systematic review and meta-regression of the literature and criteria generation through an international Delphi exercise, an early patient cohort and a patient survey. 2) Criteria reduction by Delphi and nominal group technique (NGT) exercises. 3) Criteria definition and weighting based on criterion performance and on results of a multi-criteria decision analysis. 4) Refinement of weights and threshold scores in a new derivation cohort of 1001 subjects and validation compared to previous criteria in a new validation cohort of 1270 subjects.

Results. The 2019 EULAR/ACR classification criteria for SLE include positive ANA at least once as obligatory entry criterion; followed by additive weighted criteria grouped in 7 clinical (constitutional, hematologic, neuropsychiatric, mucocutaneous, serosal, musculoskeletal, renal) and 3 immunological (antiphospholipid antibodies, complement proteins, SLE-specific antibodies) domains, and weighted from 2 to 10. Patients accumulating ≥ 10 points are classified. In the validation cohort, the new criteria had a sensitivity of 96.1% and specificity of 93.4%, compared to 82.8% sensitivity and 93.4% specificity of the ACR 1997 and 96.7% sensitivity and 83.7% specificity of the Systemic Lupus International Collaborating Clinics (SLICC) 2012 criteria.

Conclusion. These new classification criteria were developed using rigorous methodology with multidisciplinary and international input, and have excellent sensitivity and specificity. Use of ANA entry criterion, hierarchically clustered and weighted criteria reflect current thinking about SLE and provide an improved foundation for SLE research.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a complex autoimmune disease with variable clinical features (1;2). SLE manifestations are associated with multiple autoantibodies, ensuing immune complex formation and deposition, and other immune processes (2;3). This complex clinical presentation and pathogenesis makes SLE a difficult disease to grasp and define. Classification criteria are essential for the identification of relatively homogeneous groups of patients for inclusion in research studies and trials (4;5). The 1982 revised American College of Rheumatology (ACR) SLE classification criteria (6) and their 1997 revision (7) have been used worldwide. Since then, our understanding of the disease has advanced. Additional specific skin manifestations were described, some clinical symptoms were better understood, and immunological tests, such as diminished levels of serum complement components C3 and C4 or testing for anti- β 2 glycoprotein I antibodies, entered routine clinical practice. Better understanding of organ system involvement, such as mucocutaneous abnormalities, led to questions about whether some of the independently counted criteria were in fact manifestations of the same phenomenon (8).

The 2012 Systemic Lupus International Collaborating Clinics (SLICC) classification criteria addressed many of these issues (9). Mucocutaneous and neuropsychiatric manifestations were added, as were hypocomplementemia and new anti-phospholipid antibody tests; and criteria definitions were refined. The SLICC criteria emphasized that SLE is primarily an autoantibody disease, requiring at least one immunological criterion to be present, and categorized histology-proven nephritis compatible with SLE as sufficient for classification, if anti-nuclear antibodies (ANA) or antibodies to double-stranded DNA (dsDNA) were present. While achieving their goal of increasing sensitivity, the SLICC criteria have lower specificity than the 1997 ACR criteria (9;10).

Existing SLE classification criteria perform better in patients with longstanding disease than in new-onset SLE (13), and there is an increasing recognition and demand that subjects with early SLE should be included in clinical studies and trials. We therefore attempted to enrich our sample populations for early SLE in several phases of the project.

In parallel with improved understanding of SLE, the field of classification criteria development has also seen advances (4;14-16). In order to minimize investigator bias, it is now recommended that the cohorts in which the criteria are tested are from independent centers (4). Other methodologic recommendations include a balanced use of both expert-based and data-driven methods, and inclusion of the patient perspective (15;16). The approach chosen for these 2019 EULAR/ACR SLE classification criteria was specifically designed to maintain this balance and to uphold rigorous methodology.

METHODS

Methodologic overview. Using a methodological approach based on measurement science the criteria were developed in four phases (10): 1) criteria generation, 2) criteria reduction, 3) criteria definition and weighting and 4) refinement and validation (Figure 1). The whole initiative was overseen by a 12-member steering committee (MA, KHC, DD, MM, RR-G, JSS, DW, DB, DK, DJ, TD and SRJ) nominated by EULAR and the ACR in equal numbers, based on SLE and/or methodological experience and previous involvement in international projects.

The current project, jointly supported by the European League Against Rheumatism (EULAR) and the ACR, was originally based on two key concepts. One, we hypothesized that the presence of ANA would be better employed as an entry criterion than as a classification criterion (10). Such an approach was thought to reflect underlying SLE pathogenesis, and take into account ANA test characteristics of high sensitivity and limited specificity. Two, we expected individual criteria would not be of equal utility (weight) for the classification of SLE (11), for example mucosal ulcers vs. biopsy-proven lupus nephritis. Accordingly, the validity of using positive ANA as an entry criterion was explicitly addressed in phase 1 of the current activity (12). Likewise, methodologic strategies to develop weighted criteria were used.

Phase 1. Criteria generation. The purpose of Phase 1 was to test ANA as a potential entry criterion and identify candidate criteria that should be considered for SLE classification using both data- and expert-based methods, including the patient perspective. *Phase 1a* comprised a systematic literature review of Medline, Embase and the Cochrane databases with meta-regression to evaluate the operating characteristics of ANA testing for consideration as an entry criterion (12). *Phase 1b* consisted of a Delphi exercise of international SLE experts from

the Americas, Europe and Asia (17). These experts included rheumatologists, dermatologists, nephrologists, pediatricians and non-clinical SLE researchers, providing a broad perspective. The Delphi participants were asked to nominate a broad set of items potentially useful in the classification of SLE (17). In round 2 and 3, participants rated the items from 1 (not at all appropriate) to 9 (completely appropriate) for classification of SLE. Criteria were retained if they reached a median rating of ≥ 6.5 ; i.e. at least 50% of the ratings in the high range (7, 8 or 9). Participants were also asked about the importance of ANA and histopathology for classification of SLE. *Phase 1c* established an international cohort of patients with early SLE or conditions mimicking SLE to identify criteria that may discriminate subjects with early (less than 12 months) disease (18). *Phase 1d* comprised a cross-sectional survey of SLE patients, administered via the quarterly journal of the German SLE patient organization, which asked about symptoms within one year before and after the patient's diagnosis of SLE (19). While at a risk of recall bias and not necessarily representative of other regions worldwide, this survey was done to explicitly take a patient standpoint into account.

For phase 2 and 3, additional renowned European and North American SLE experts were nominated by the steering committee and invited to participate.

Phase 2. Criteria reduction. *Phase 2a.* The objective of this phase was to select a set of criteria from Phase 1 that maximized the likelihood of accurate classification of SLE, particularly of early disease. An independent panel of 7 of the international SLE experts (RC, NC-C, DDG, BHH, FH, EM, JS-G) ranked the candidate criteria from phase 1. A consensus meeting of 19 international SLE experts (n=7 nominal group technique (NGT) experts + steering committee + DK [moderator]) using NGT was conducted to reduce the list of criteria (20). Data for each candidate criterion were reviewed and discussed until consensus was achieved. The NGT experts voted on items to be retained. *Phase 2b.* NGT participants pointed out that some criteria could be correlated. With the idea of potentially clustering criteria into domains, associations between candidate criteria were evaluated separately in two cohorts, the phase 1c early SLE and the Euro-lupus cohorts (21).

Phase 3. Criteria definition and weighting. *Phase 3a.* The operating characteristics of the retained candidate criteria were evaluated by literature review. Candidate criteria were

hierarchically organized into clinical and immunological domains, and definitions for the candidate criteria were iteratively refined. SLE patient advocates participated in the review of data and the steering committee discussions (22).

Phase 3b. 164 case vignettes reflecting broad SLE clinical presentation were sampled from SLE centers across several countries. A panel of 6 of the international experts not involved in earlier phases of the project (BD, SJ, WJMcc, GR-I, MS, MBU) and 11 members of the steering committee assessed and ranked a representative sample of the cases. Subsequently, at a face-to-face meeting, this panel of 17 international SLE experts iteratively compared pairs of criteria, using multi-criteria decision analysis facilitated by 1000minds software (23). The panel unanimously agreed to further reduce the list of criteria. Based on the results, provisional criteria weights were assigned and a provisional threshold score for classification was determined as the lowest score at which the expert panel had achieved consensus on classifying a case vignette as SLE (24).

Phase 4. Refinement and validation. International SLE experts not involved in phase 2 or phase 3 panels were asked to contribute cases diagnosed as SLE and controls with conditions mimicking SLE sampled from patients evaluated at their centers. Each center was asked to contribute up to 100 cases and an equal number of controls, preferentially sampling those with early disease, and regardless of their specific clinical or immunological manifestations. Pseudonymized data on the criteria were collected using a standardized data collection form. Ethics committee approval and informed consent were obtained as per local requirements. The status (“SLE” or not) of each case underwent independent adjudication by three of four SLE experts (GB, BFH, NL, CT) from different centers. Queries were sent back to the submitting investigator for clarification. Of this cohort, 501 SLE and 500 control subjects were randomly selected to comprise the derivation cohort, while the remaining 696 SLE and 574 control subjects formed the validation cohort.

Refinement. The performance of the draft criteria set was iteratively tested in the derivation cohort. A data-driven threshold for classification was determined by receiver operating characteristics (ROC) analysis and compared to the provisional expert-based consensus threshold. The data of SLE subjects below the threshold (misclassified) were reviewed for groups of patients with unequivocal SLE who still missed classification, and criteria weights adjusted slightly, while preserving the weighting hierarchy (Details below in Results Phase IV).

Sensitivity and specificity was tested against the ACR 1997 and the SLICC 2012 criteria. In addition, ANA as an entry criterion was tested against not having an entry criterion. Finally, the criteria weights were simplified to whole numbers. Refinements to the criteria set were presented to the steering committee and Phase III expert panel, and unanimously endorsed. *Validation.* The sensitivity and specificity of the final criteria were tested in the validation cohort and compared to previous SLE criteria sets.

Statistical analysis. Descriptive statistics were used to summarize the data. Confidence intervals were calculated using the bias-corrected and accelerated bootstrap method (BCa method) with $B = 2000$ bootstrap samples. The BCa method resamples the input data B times (with replacement) and calculates the required statistics (sensitivity, specificity, AUC). Based on the B bootstraps samples, the bias-correction is applied and the associated 95% confidence intervals for the statistics are estimated. The BCa method has proven to yield very accurate coverage of estimated confidence intervals (25). The number B of bootstrap resamples is recommended to be at least $B = 1000$. We have chosen $B = 2000$ and additionally checked if $B = 5000$ bootstraps changed the estimated confidence bounds, which was not the case. Statistical analyses were performed using R, version 3.4.0 (The R Foundation of Statistical Computing).

RESULTS

Phase 1. Criteria generation.

Phase 1a. ANA as an entry criterion. A systematic review of MEDLINE, EMBASE and the Cochrane database identified 13,080 patients from 64 studies reporting ANA by immunofluorescence on HEp-2 cells. Meta-regression of the operating characteristics of ANA found a sensitivity of 97.8% [95% confidence interval (CI) 96.8% – 98.5%] for ANA of $\geq 1:80$ supporting use of ANA as an entry criterion (12). Since some SLE centers do not have access to HEp-2 ANA, and in view of ongoing work on the standardization of serology and potential future advances in the field, the steering committee and additional autoantibody consultants (MJF, PLM) recommended the provision “or an equivalent positive ANA test. Testing by immunofluorescence on HEp-2 cells or a solid phase ANA screening immunoassay with at least equivalent performance is highly recommended.”

Phase 1b. Delphi exercise. One hundred and forty-seven international SLE experts nominated 145 candidate criteria (17). By rating the appropriateness for SLE classification, the

participants in the second and third Delphi round reduced the list to 40 candidate criteria (Supplementary Table 1).

Phase 1c. International early SLE cohort. The cohort comprised 616 subjects who had been referred for possible SLE with a disease duration of less than one year (n=389 early SLE and n=227 mimicking diseases) from North America, Europe, Asia and South America (18). In addition to supporting many of the 40 candidate criteria derived from the Delphi exercise, the comparison between early SLE and non-SLE patients showed that fever occurred more frequently (34.5% versus 13.7%, $p < 0.001$) in SLE, while SLE patients less commonly suffered from arthralgias (20.3% versus 42.7%, $p = 0.001$) and fatigue (28.3% versus 37%, $p = 0.02$).

Phase 1d. Patient survey. 339 SLE patients (>99% Caucasian, 93% female) responded to the survey (19). More than half of these patients reported mucocutaneous findings in the first year of their disease (Supplementary Table 1), but also fatigue (89%), joint pain (87%) and fever (54%)(19). Given that these items were highlighted both in the early SLE cohort and the patient survey, fever, fatigue and arthralgias were forwarded to the next phase in addition to the 40 Delphi items. Accordingly, phases 1a-1d resulted in a total of 43 candidate criteria for consideration (Supplementary Table 1).

Phase 2. Criteria reduction.

Phase 2a. The expert panel NGT exercise reduced the candidate criteria from 43 to 21 (26). The panel distinguished potential “entry criteria”, which would be required for classification, from potential “additive criteria”. They endorsed “positive ANA ($\geq 1:80$ by HEp-2 immunofluorescence)” as an entry criterion. The 20 remaining additive criteria included: lupus nephritis by renal biopsy, autoantibodies, cytopenias, fever, arthritis, serositis, mucocutaneous and neuropsychiatric manifestations (Supplementary Table 1).

Phase 2b. Associations between the candidate criteria were evaluated in 389 subjects in the early SLE cohort and the 1000 SLE subjects of the Euro-lupus cohort. Modest statistically significant correlations were limited to the mucocutaneous ($r = 0.22$ to 0.30), neurologic ($r = 0.22$) and immunological ($r = 0.33$) domains in the early SLE cohort, and this modest correlation was replicated in the Euro-lupus cohort (21). Given these associations, criteria were clustered within domains, so that only one criterion within each domain would be counted.

Phase 3. Criteria definition and weighting

Phase 3a. Based on the literature, definitions of the 20 candidate additive criteria were refined, using a data-driven evaluation of operating characteristics (22), retaining only feasible items with a prevalence of at least 1% according to literature. Literature-review led to the consensus decision to evaluate five different candidate criteria within the neuropsychiatric domain (delirium, psychosis, seizure, mononeuropathy, cranial neuropathy) and potential separation of acute pericarditis from pleural or pericardial effusions; and between diminished C3 or C4 versus diminished C3 and C4 (Supplementary Table 1). The resulting 23 candidate criteria (Supplementary Table 1) were organized into seven clinical and three immunologic domains, with hierarchical clustering (22). Only the highest-ranking item in each domain was to be counted. Instead of devising exclusion definitions for each criterion, the decision was made to attribute any item to SLE only if no more likely explanation was present. For leukopenia and joint involvement, it was decided to formally test alternative definitions in the derivation cohort. Given the importance of testing for antibodies, particularly for anti-dsDNA, for which tests of relatively low specificity are in use, great care was taken to precisely define testing (Table 1).

Phase 3b. The 1.5 day in-person consensus meeting using multicriteria decision analysis involved 74 decisions between pairs of criteria. Criteria weights were calculated by the 1000minds™ software based on these decisions (Table 2). International Society of Nephrology/Renal Pathology Society class III or IV nephritis consistently attained higher weight than class II or V nephritis, so lupus nephritis by histology was separated into two different criteria. Class VI lupus nephritis as an end stage manifestation was unanimously eliminated. Likewise, the experts unanimously voted to not retain mononeuropathy and cranial neuropathy, which had been included into the set of potential neuropsychiatric items in phase 3a but turned out to add little to SLE classification. The use of weighted criteria led to a sum score that is a measure of the relative probability of a subject having SLE, with higher scores indicating higher likelihood. Experts reached full consensus on a classification of SLE at a provisional threshold score of >83 of a theoretical maximum of 305 (24).

Phase 4. Refinement and validation

Twenty-one centers from the United States, Canada, Mexico, Austria, Croatia, France, Germany, Greece, Hungary, Italy, Portugal, Spain, the United Kingdom, Turkey, Hong Kong and

Japan submitted a total of 2,339 cases from their cohorts. 1,197 SLE and 1,074 non-SLE diagnoses (Table 3) were verified by three adjudicators blinded to the proposed classification criteria system. Due to lack of consensus during adjudication, 68 subjects (2.9%) were excluded from the analysis.

Derivation cohort. Of the 2,271 triple-adjudicated cases, 501 SLE and 500 non SLE cases were randomly assigned to the derivation cohort. The provisional weighting system derived from phase 3 was tested in the derivation cohort. ROC analysis suggested a data-driven threshold of ≥ 70 (of a maximum of 305), with a sensitivity of 95.4% and a specificity of 95.2%, which was superior to the consensus-derived provisional threshold of >83 that had high specificity (98.8%), but lower sensitivity (81.6%). Review of subjects below the threshold of 70 identified a subgroup of SLE subjects with joint involvement and/or leukopenia. Thus, weights for leukopenia and joint involvement were each adjusted (Table 2) to reduce misclassification. When alternative definitions for leukopenia and joint involvement were tested, leukopenia defined as a white blood cell count (WBC) $<4000/\text{mm}^3$ at least once (9) also had a slightly higher sensitivity + specificity (1.944 vs. 1.942) than leukopenia defined as WBC $<4000/\text{mm}^3$ on 2 or more occasions (6;26). Joint involvement defined as EITHER “synovitis involving 2 or more joints, characterized by swelling or effusion”, OR “tenderness in 2 or more joints and at least 30 minutes of morning stiffness” (9) had a higher combined sensitivity and specificity than arthritis defined simply as synovitis of 2 or more joints (1.944 vs. 1.900). When re-tested, the revised criteria had increased sensitivity, and maintained sensitivity + specificity. Evaluating ANA as an entry criterion, the criteria with the ANA entry criterion had better performance than without (sensitivity + specificity 1.944 vs. 1.930). Next, the weights were simplified by division to whole numbers to achieve a threshold of 10 (Table 2). In the derivation cohort, the sensitivity and specificity of the final criteria set (Figure 2) were reaching the performance benchmarks set for this project (Table 4).

Validation. The validation cohort, i.e. the full cohort minus the derivation cohort, comprised 1,270 triple adjudicated subjects (n=696 SLE, n= 574 controls). The criteria, with positive ANA as an entry criterion, weighted criteria in seven clinical domains (constitutional, hematologic, neuropsychiatric, mucocutaneous, serosal, musculoskeletal, renal) and three immunological domains (anti-phospholipid antibodies, low complements, anti-Sm and anti-dsDNA as SLE-specific antibodies) and a classification threshold score of ≥ 10 (out of a theoretical maximum

of 51)(Figure 2), had a sensitivity of 96.1% and a specificity of 93.4% (Table 4). It demonstrated improved performance compared to the ACR 1997 and SLICC 2012 criteria.

DISCUSSION

New SLE classification criteria were developed with support by both the ACR and EULAR. Through a 4-phase, iterative process, we have defined an additive, weighted multi-criteria system that produces a measure of the relative probability that an individual can be classified as SLE. The system defines a threshold above which experts would classify cases as SLE for the purpose of research studies. We have carefully defined the criteria to improve reliability and precision; and have grouped the criteria into ten hierarchical domains. We have validated the criteria against a large number of cases, including many patients with manifestations that resemble SLE but who do not have SLE. This approach, as well as the resulting criteria system, represents a paradigm shift for the classification of SLE.

We have defined positive ANA at any time as required entry criterion. There were three possible ways to deal with ANA testing. The previous criteria sets have treated ANA the same as the much more specific antibodies against Sm and dsDNA, which we considered suboptimal given important differences in sensitivity and specificity. We could have excluded ANA completely in classifying lupus, but we still consider ANA a useful test and concept. We therefore decided to test ANA as an entry criterion, which reflects the use of ANA as a highly sensitive screening test.

Criteria using ANA as entry criterion had better performance. During the phase 1 Delphi exercise, 58% of SLE experts did not feel comfortable and an additional 19% were uncertain about classifying a patient with SLE in the absence of ever having a positive ANA (17). The systematic literature review and meta-regression of data on 13,080 subjects demonstrated ANA $\geq 1:80$ have a sensitivity of 98% with a lower limit of the 95% confidence interval at 97% (12). In the phase 1 early SLE cohort, 99.5% of the 389 SLE patients were ANA positive (18). The frequencies of ANA positive SLE patients in the derivation and validation cohorts (99.6% and 99.3%, respectively) were in the same range. Since both in the early SLE cohort and in the derivation and validation cohorts, patients were included in many centers worldwide

independent of ANA positivity, the latter data provide additional support for ANA as an entry criterion.

Using ANA as entry criterion means the new criteria cannot classify SLE among patients who are persistently ANA negative. While possibly also distinguished by lower cytokine levels (27) and lower efficacy of immunomodulatory treatment (28), such a subgroup of patients exists. Although small, it may vary in size in different populations (12). This patient subset needs to be put high on the scientific agenda for further investigation. Additional characterization of this phenomenon may lead to an alternative entry criterion for this small group of patients. For the moment, we still think it is acceptable to exclude ANA negative patients from clinical trials.

Molecular classification criteria were also considered during the development of these criteria (29). Many novel biomarkers were nominated, such as increased circulating B lymphocyte stimulator (BLyS), IFN γ induced protein 10 kD (IP-10), monocyte chemoattractant protein-1 (MCP-1), TNF- α , type I interferon signature, or increased Th17 and plasma cell populations. They were all voted out in the expert Delphi exercise, largely because of limited availability in the clinical setting and/or insufficient evidence (5). However, inclusion of novel biomarkers, beyond autoantibodies, may ultimately further improve the specificity of SLE classification, increase alignment of classification with underlying disease pathogenesis and improve the performance and information content of clinical trials. Thus, testing of biomarkers against these criteria is an important area for future research.

A new clinical criterion, unexplained fever, turned out to be common and remarkably characteristic for SLE. However, since infections are a major cause of death in SLE, it is of utmost importance to stress that fever, like all other criteria manifestations, should only be counted if no better explanation exists, and that infections have to be suspected first in any patient with (potential) SLE, particularly when CRP is elevated (30). The concept that all criteria are only to be counted if SLE is thought to be the most likely cause of the manifestation (i.e. no other more likely cause exists) is central to these new EULAR/ACR criteria, and is explicitly stated as an overarching principle. Some criteria, such as delirium, psychosis and acute pericarditis, were in part re-defined based on existing scientific

definitions (22). Where alternative definitions were used, the performance of the alternative definitions was comparatively evaluated in the derivation cohort.

The differential weighting of criteria better represents their relative contribution to an individual's classification of SLE. For SLE, renal biopsy with Class III or IV lupus nephritis carries the most weight and in the presence of a positive ANA is enough to classify a patient as SLE. This further develops a concept of the SLICC criteria (9) and reflects the current thinking of SLE experts; in the Delphi exercise, 85% would classify SLE on renal pathology alone (17). Renal biopsy with class II or V lupus nephritis still carries a large weight (8 points) but is not by itself sufficient for the classification of SLE.

The numerical goal of this project was to keep the specificity similar to the specificity of the ACR 1997 criteria, but increase the sensitivity to the high sensitivity level of the SLICC criteria, if possible. The validation cohort data suggest that this goal has been achieved. From our data, it appears that the SLICC criteria increase in sensitivity was to a significant degree founded in accepting renal histology and adding subacute cutaneous lupus and low complement levels. These three advances are mirrored in the current criteria. Many of the other additional symptoms of the SLICC criteria were of very low frequency. Specificity was increased by weighting of criteria, by the NGT expert panel decision to not allow lymphopenia to go forward, and, importantly, by the decision that no criterion be counted if better explained by another condition.

The new criteria provide a simple, directed and highly accurate method for classifying SLE. An electronic 'app' is in preparation, which will assist in the use of these criteria. However, it is important to stress that classification criteria are not designed for diagnosis or treatment decisions (5). They should never be used to exclude patients who do not fully meet these criteria from receiving appropriate therapies. This is also pertinent to patients with ANA-negative SLE discussed above. Diagnosis of SLE remains the purview of an appropriately trained physician evaluating an individual patient (5).

The new SLE classification system also provides new research opportunities. With much interest in early or latent SLE (31;32), the additive point system and the relative probability of

classification it produces, allows for systematic study of individuals who fall below the classification threshold. This will facilitate studies of disease evolution and early intervention. Furthermore, the use of an additive scoring system will allow for studying the idea of 'ominosity', i.e. the potential implications of having very high scores on disease severity and subsequent prognosis. This work would need to reconsider the relative contribution of individual criteria (weights) and consider additional criteria that potentially contribute to ominosity.

It is anticipated that other groups will test these criteria, which will constitute important external validation. This will be particularly important for pediatric SLE and those with organ dominant, e.g. skin dominant disease, since it is a limitation of this criteria project that the patient cohorts do not represent these subgroups. Similar limitations also pertain to several racial/ethnic groups (for example, African American/Black, Hispanic and Asian patients) and to men with SLE, each only included in lower numbers (Table 3). It is important to independently test the EULAR/ACR criteria in these subgroups. Leukocyte counts, for example, are more frequently below $<4000/\text{mm}^3$ in African Americans (33), which may have influence on criteria performance. It is also possible that the academic center patient populations included differ from patients in community practice clinics. Investigators testing the new criteria in different populations are reminded about the critical importance of the correct attribution of each criterion. Criteria can only be counted when not better explained by another condition. The attribution process requires diligence and clinical experience.

In summary, our multiphase methodologic approach and ensuing classification system using ANA as an entry criterion and weighted, hierarchically clustered criteria, constitute a paradigm shift in the classification of SLE. These criteria have excellent performance characteristics and face validity, as the structure and weighting were designed to reflect current thinking about SLE. The inclusion of fever assists with the classification of early SLE. The separation of renal biopsy findings reflects their differential impact on the probability of SLE classification. These criteria have strong operating characteristics, with excellent sensitivity and specificity. This classification system was built using rigorous methodology that was both data-driven and expert-based. With the inclusion of over 200 SLE experts from multiple countries and medical

disciplines, methodologists, patient advocates and over 4,000 subjects, this work is the largest international, collaborative SLE classification effort to date.

References

- (1) Bertsias GK, Pamfil C, Fanouriakis A, Boumpas DT. Diagnostic criteria for systemic lupus erythematosus: has the time come? *Nat Rev Rheumatol* 2013; 9: 687-94.
- (2) Rahman A, Isenberg DA. Systemic lupus erythematosus. *N Engl J Med* 2008; 358: 929-39.
- (3) Tsokos GC. Systemic lupus erythematosus. *N Engl J Med* 2011; 365: 2110-21.
- (4) Johnson SR, Goek ON, Singh-Grewal D, Vlad SC, Feldman BM, Felson DT et al. Classification criteria in rheumatic diseases: a review of methodologic properties. *Arthritis Rheum* 2007; 57: 1119-33.
- (5) Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A et al. Distinctions between diagnostic and classification criteria? *Arthritis Care Res (Hoboken)* 2015; 67: 891-7.
- (6) 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.
- (7) Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997; 40: 1725.
- (8) Albrecht J, Berlin JA, Braverman IM, Callen JP, Connolly MK, Costner MI et al. Dermatology position paper on the revision of the 1982 ACR criteria for systemic lupus erythematosus. *Lupus* 2004; 13: 839-49.
- (9) Petri M, Orbai AM, Alarcon GS, Gordon C, Merrill JT, Fortin PR et al. Derivation and validation of systemic lupus international collaborating clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum* 2012; 2677-86.
- (10) Aringer M, Dorner T, Leuchten N, Johnson SR. Toward new criteria for systemic lupus erythematosus-a standpoint. *Lupus* 2016; 25: 805-11.
- (11) Costenbader KH, Karlson EW, Mandl LA. Defining lupus cases for clinical studies: the Boston weighted criteria for the classification of systemic lupus erythematosus. *J Rheumatol* 2002; 29: 2545-50.
- (12) Leuchten N, Hoyer A, Brinks R, Schoels M, Schneider M, Smolen J et al. Performance of Anti-nuclear Antibodies for Classifying Systemic Lupus Erythematosus: a Systematic Literature Review and Meta-regression of Diagnostic Data. *Arthritis Care Res (Hoboken)* 2018; 70: 428-38.
- (13) Ines L, Silva C, Galindo M, Lopez-Longo FJ, Terroso G, Romao VC et al. Classification of Systemic Lupus Erythematosus: Systemic Lupus International Collaborating Clinics Versus American College of Rheumatology Criteria. A Comparative Study of 2,055 Patients From a Real-Life, International Systemic Lupus Erythematosus Cohort. *Arthritis Care Res (Hoboken)* 2015; 67: 1180-5.
- (14) Felson DT, Anderson JJ. Methodological and statistical approaches to criteria development in rheumatic diseases. *Baillieres Clin Rheumatol* 1995; 9: 253-66.
- (15) Singh JA, Solomon DH, Dougados M, Felson D, Hawker G, Katz P et al. Development of classification and response criteria for rheumatic diseases. *Arthritis Rheum* 2006; 55: 348-52.

- (16) Dougados M, Gossec L. Classification criteria for rheumatic diseases: why and how? *Arthritis Rheum* 2007; 57: 1112-5.
- (17) Schmajuk G, Hoyer BF, Aringer M, Johnson SR, Daikh DI, Dorner T. Multi-center Delphi Exercise Reveals Important Key Items for Classifying Systemic Lupus Erythematosus. *Arthritis Care Res (Hoboken)* 2018; 70: 1488-94.
- (18) Mosca M, Costenbader KH, Johnson SR, Lorenzoni V, Sebastiani GD, Hoyer BF et al. 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.
- (19) Leuchten N, Milke B, Winkler-Rohlfing B, Daikh D, Dorner T, Johnson SR et al. Early symptoms of systemic lupus erythematosus (SLE) recalled by 339 SLE patients. *Lupus* 2018; 27: 1431-6.
- (20) Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. *Semin Arthritis Rheum* 2011; 41: 95-105.
- (21) Touma Z, Cervera R, Brinks R, Tani C, Hoyer BF, Costenbader KH et al. Associations among classification criteria items within systemic lupus erythematosus. *Arthritis Rheumatol* 2017; 69 (S10): #1622.(Abstract)
- (22) Tedeschi SK, Johnson SR, Boumpas D, Daikh D, Dorner T, Jayne D et al. Developing and Refining New Candidate Criteria for Systemic Lupus Erythematosus Classification: An International Collaboration. *Arthritis Care Res (Hoboken)* 2018; 70: 571-81.
- (23) Johnson SR, Naden RP, Fransen J, van den Hoogen F, Pope JE, Baron M et al. Multicriteria decision analysis methods with 1000Minds for developing systemic sclerosis classification criteria. *J Clin Epidemiol* 2014; 67: 706-14.
- (24) Tedeschi SK, Johnson SR, Boumpas DT, Daikh D, Dörner T Diamond B, et al. Multicriteria decision analysis process to develop new classification criteria for systemic lupus erythematosus. *Ann Rheum Dis* 2019; 78: 634-40.
- (25) Efron B, Tibshirani RJ. *An introduction to the bootstrap*. Boca Raton: Chapman & Hall / CRC; 1994.
- (26) Johnson S, Khanna D, Cervera R, Costedoat-Chalumeau N, Gladman DD, et al. Use of consensus methodology to determine candidate items for systemic lupus erythematosus classification criteria. *J Rheumatol* 2018; Dec 15 Epub. doi: 10.3899/jrheum.180478
- (27) Torell F, Eketjall S, Idborg H, Jakobsson PJ, Gunnarsson I, Svenungsson E et al. Cytokine Profiles in Autoantibody Defined Subgroups of Systemic Lupus Erythematosus. *J Proteome Res* 2019; 18: 1208-17.
- (28) Wallace DJ, Stohl W, Furie RA, Lisse JR, McKay JD, Merrill JT et al. A phase II, randomized, double-blind, placebo-controlled, dose-ranging study of belimumab in patients with active systemic lupus erythematosus. *Arthritis Rheum* 2009; 61: 1168-78.
- (29) Johnson SR, Hinchcliff M, Asano Y. Controversies: molecular vs. clinical systemic sclerosis classification. *J Scleroderma Relat Disord* 2016; 1: 277-85.

- (30) Littlejohn E, Marder W, Lewis E, Francis S, Jackish J, McCune WJ et al. The ratio of erythrocyte sedimentation rate to C-reactive protein is useful in distinguishing infection from flare in systemic lupus erythematosus patients presenting with fever. *Lupus* 2018; 27: 1123-9.
- (31) Ganczarczyk L, Urowitz MB, Gladman DD. "Latent lupus". *J Rheumatol* 1989; 16: 475-8.
- (32) Costenbader KH, Schur PH. We need better classification and terminology for "people at high risk of or in the process of developing lupus". *Arthritis Care Res (Hoboken)* 2015; 67: 593-6.
- (33) Hsieh MM, Everhart JE, Byrd-Holt DD, Tisdale JF, Rodgers GP. Prevalence of neutropenia in the U.S. population: age, sex, smoking status, and ethnic differences. *Ann Intern Med* 2007; 146: 486-92.

Figure 1. Development and validation of SLE classification criteria

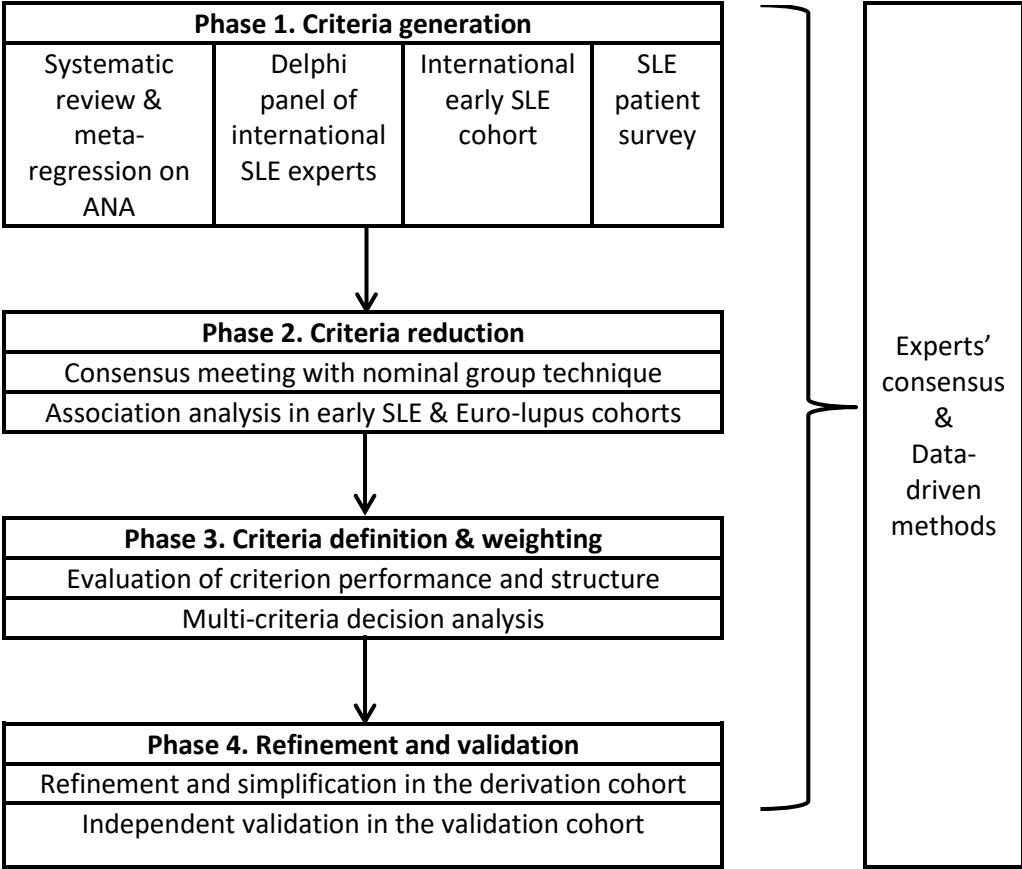


Table 1. Definitions of SLE classification criteria.

Criteria	Definition
Antinuclear antibodies (ANA)	Antinuclear antibodies (ANA) at a titer of $\geq 1:80$ on HEp-2 cells or an equivalent positive test at least once. Testing by immunofluorescence on HEp-2 cells or a solid phase ANA screening immunoassay with at least equivalent performance is highly recommended.
Fever	Temperature $>38.3^{\circ}$ Celsius.
Leukopenia	White blood cell count $<4,000/\text{mm}^3$.
Thrombocytopenia	Platelet count $<100,000/\text{mm}^3$.
Autoimmune hemolysis	Evidence of hemolysis, such as reticulocytosis, low haptoglobin, elevated indirect bilirubin, elevated LDH AND positive Coomb's (direct antiglobulin) test.
Delirium	Characterized by (1) change in consciousness or level of arousal with reduced ability to focus, and (2) symptom development over hours to <2 days, and (3) symptom fluctuation throughout the day, and (4) either (4a) acute/subacute change in cognition (e.g. memory deficit or disorientation), or (4b) change in behavior, mood, or affect (e.g. restlessness, reversal of sleep/wake cycle).
Psychosis	Characterized by (1) delusions and/or hallucinations without insight and (2) absence of delirium.
Seizure	Primary generalized seizure or partial/focal seizure.
Non-scarring alopecia	Non-scarring alopecia observed by a clinician*.
Oral ulcers	Oral ulcers observed by a clinician*.
Subacute cutaneous or discoid lupus	Subacute cutaneous lupus erythematosus observed by a clinician*: Annular or papulosquamous (psoriasiform) cutaneous eruption, usually photodistributed. Discoid lupus erythematosus observed by a clinician*: Erythematous-violaceous cutaneous lesions with secondary changes of atrophic scarring, dyspigmentation, often follicular hyperkeratosis/ plugging (scalp), leading to scarring alopecia on the scalp. If skin biopsy is performed, typical changes must be present. Subacute cutaneous lupus: interface vacuolar dermatitis consisting of a perivascular lymphohistiocytic infiltrate, often with dermal mucin noted. Discoid lupus: interface vacuolar dermatitis consisting of a perivascular and/or periappendageal lymphohistiocytic infiltrate. In the scalp, follicular keratin plugs may be seen. In longstanding lesions, mucin deposition and basement membrane thickening may be noted.
Acute cutaneous lupus	Malar rash or generalized maculopapular rash observed by a clinician*. If skin biopsy is performed, typical changes must be present (Acute cutaneous lupus: interface vacuolar dermatitis consisting of a perivascular lymphohistiocytic infiltrate, often

	with dermal mucin noted. Perivascular neutrophilic infiltrate may be present early in the course.
Pleural or pericardial effusion	Imaging evidence (such as ultrasound, x-ray, CT scan, MRI) of pleural or pericardial effusion, or both.
Acute pericarditis	≥2 of (1) pericardial chest pain (typically sharp, worse with inspiration, improved by leaning forward), (2) pericardial rub, (3) EKG with new widespread ST-elevation or PR depression, (4) new or worsened pericardial effusion on imaging (such as ultrasound, x-ray, CT scan, MRI).
Joint involvement	EITHER (1) synovitis involving 2 or more joints characterized by swelling or effusion OR (2) tenderness in 2 or more joints and at least 30 minutes of morning stiffness.
Proteinuria >0.5g/24h	Proteinuria >0.5g/24h by 24 hour urine or equivalent spot urine protein-to-creatinine ratio.
Class II or V lupus nephritis on renal biopsy according to ISN/RPS 2003 classification.	Class II: Mesangial proliferative lupus nephritis: Purely mesangial hypercellularity of any degree or mesangial matrix expansion by light microscopy, with mesangial immune deposit. A few isolated subepithelial or subendothelial deposits may be visible by immune-fluorescence or electron microscopy, but not by light microscopy. Class V: Membranous lupus nephritis: Global or segmental subepithelial immune deposits or their morphologic sequelae by light microscopy and by immunofluorescence or electron microscopy, with or without mesangial alterations.
Class III or IV lupus nephritis on renal biopsy according to ISN/RPS 2003.	Class III: Focal lupus nephritis: Active or inactive focal, segmental or global endo- or extracapillary glomerulonephritis involving <50% of all glomeruli, typically with focal subendothelial immune deposits, with or without mesangial alterations. Class IV: Diffuse lupus nephritis: Active or inactive diffuse, segmental or global endo- or extracapillary glomerulonephritis involving ≥50% of all glomeruli, typically with diffuse subendothelial immune deposits, with or without mesangial alterations. This class includes cases with diffuse wire loop deposits but with little or no glomerular proliferation.
Positive anti-phospholipid antibodies	Anti-Cardiolipin antibodies (IgA, IgG, or IgM) at medium or high titer (>40 APL, GPL or MPL, or >the 99th percentile) or positive anti-β2GP1 antibodies (IgA, IgG, or IgM) or positive lupus anticoagulant.
Low C3 OR low C4	C3 OR C4 below the lower limit of normal.
Low C3 AND low C4	Both C3 AND C4 below their lower limits of normal.
Anti-dsDNA antibodies OR Anti-Smith (Sm) antibodies.	Anti-dsDNA antibodies in an immunoassay with demonstrated ≥ 90% specificity for SLE against relevant disease controls OR Anti-Smith (Sm) antibodies.

*This may include physical examination or review of a photograph.

Table 2. Relative weights of the additive classification criteria items

Domain	Item	Original	Modification	Revised	Simplified
Constitutional	Fever	13		13	2
Hematological	Leukopenia	12	+7	19	3
	Thrombocytopenia	26		26	4
	Autoimmune hemolysis	28		28	4
Neuropsychiatric	Delirium	12		12	2
	Psychosis	20		20	3
	Seizure	34		34	5
Mucocutaneous	Alopecia	13		13	2
	Oral ulcers	14		14	2
	SCLE/DLE	29		29	4
	ACLE	38		38	6
Serosal	Effusion	34		34	5
	Acute pericarditis	38		38	6
Musculoskeletal	Joint involvement	34	+4	38	6
Renal	Proteinuria	27		27	4
	Class II/V	55		55	8
	Class III/IV	74		74	10
APL antibodies	Anti-phospholipid	13		13	2
Complements	C3 or C4 low	19		19	3
	C3 and C4 low	27		27	4
SLE-specific antibodies	Anti-Sm	40		40	6
	Anti-dsDNA	38		38	6

Weights derived from the phase III consensus meeting with multicriteria decisions analysis (Original), added points for leukopenia and joint involvement (Modification), the resulting weights (Revised) and the final simplified weights (Simplified).

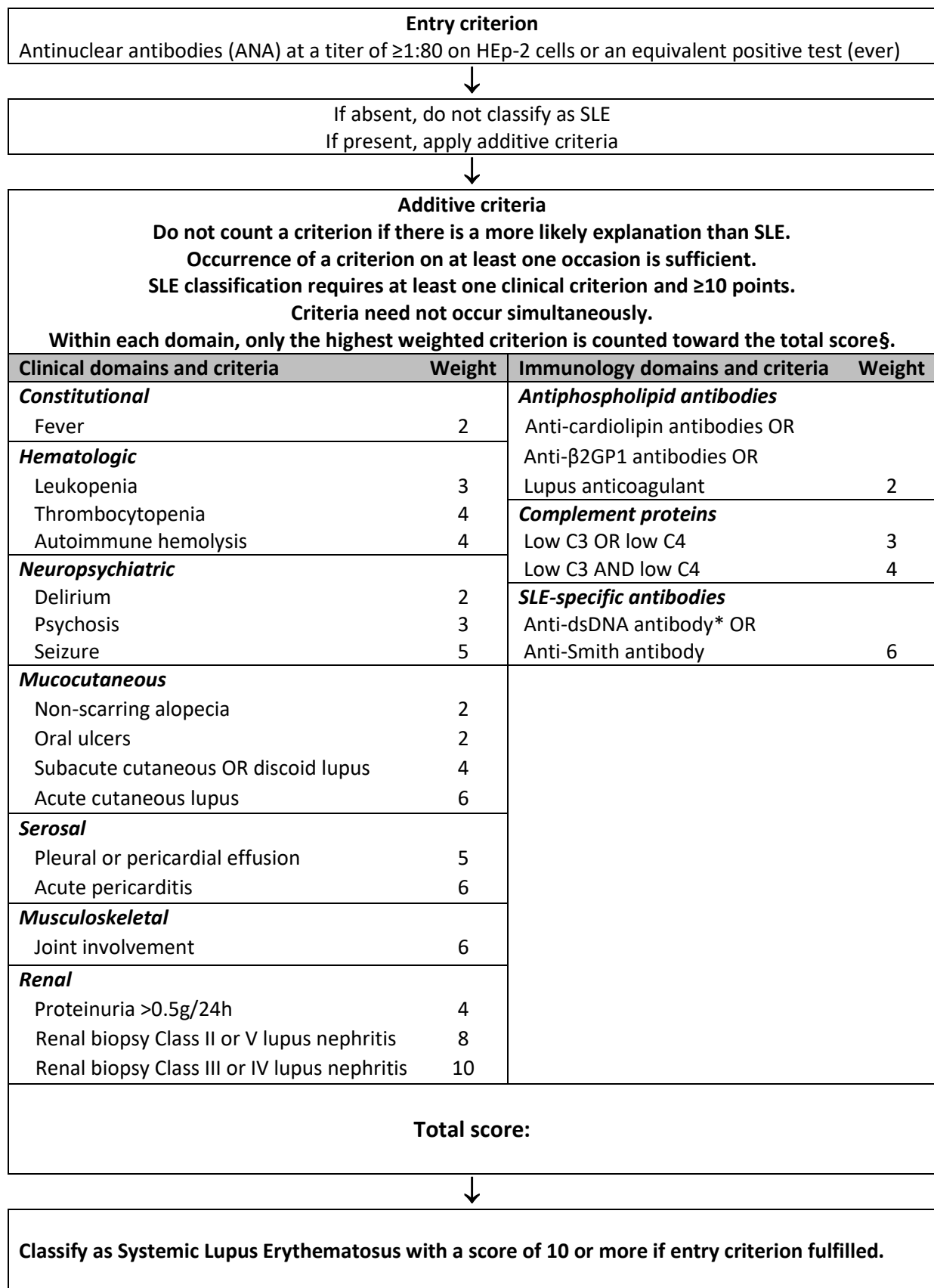
Table 3. Demographic characteristics of the derivation and validation cohorts

	Derivation cohort		Validation cohort	
	SLE	Non-SLE	SLE	Non-SLE
n	501	500	696	574
Female/male	447/54	421/79	608/88	490/84
Age (mean±SD) years	45±14	54±16	45±14	56±16
Disease duration (mean±SD) years	11±8	9±8	11±8	9±8
Ethnicity				
Black	29	10	56	12
East Asian	36	29	53	34
Hispanic	59	48	73	51
South/South East Asian	16	6	21	11
White	355	404	480	461
Other	6	3	13	5
SLE	501		696	
Non-SLE		500		574
Adult Onset Still's disease		2		11
Autoimmune thyroiditis		6		5
Behcet's disease		7		9
Cancer		2		3
Inflammatory myositis		37		27
Fibromyalgia		6		3
Membranous nephritis		11		14
Mixed connective tissue disease		9		15
Osteoarthritis		2		0
Primary antiphospholipid antibody Syndrome		45		48
Psoriatic arthritis		12		9
Rheumatoid arthritis		94		110
Sarcoidosis		2		2
Sjögren's syndrome		112		124
Spondyloarthritis		5		5
Systemic sclerosis		99		120
Tuberculosis		0		2
Undifferentiated connective tissue disease		16		20
Vasculitis		9		13
Viral infection		5		5
Other		19		29

Inflammatory myositis includes dermatomyositis, polymyositis, and juvenile dermatomyositis

SD Standard deviation

Figure 2. Classification criteria for systemic lupus erythematosus



Note: * In an assay with $\geq 90\%$ specificity against relevant disease controls

§ Additional criteria items within the same domain will not be counted.

Table 4. Operating characteristics of the new classification criteria compared to the ACR 1997 and SLICC 2012 classification criteria in the derivation and the validation cohorts.

	ACR 1997 criteria	SLICC 2012 criteria	EULAR/ACR 2019 criteria
Derivation			
Sensitivity [95% CI]	0.85 [0.81-0.88]	0.97 [0.95-0.98]	0.98 [0.97-0.99]
Specificity [95% CI]	0.95 [0.93-0.97]	0.90 [0.87-0.92]	0.96 [0.95-0.98]
Combined [95% CI]	1.80 [1.76-1.83]	1.87 [1.84-1.90]	1.94 [1.92-1.96]
Validation			
Sensitivity [95% CI]	0.83 [0.80-0.85]	0.97 [0.95-0.98]	0.96 [0.95-0.98]
Specificity [95% CI]	0.93 [0.91-0.95]	0.84 [0.80-0.87]	0.93 [0.91-0.95]
Combined [95% CI]	1.76 [1.73-1.80]	1.80 [1.77-1.84]	1.90 [1.87-1.92]

ACR American College of Rheumatology, SLICC Systemic Lupus International Collaborating Clinics, CI Confidence Intervals