Use of Consensus Methodology to Determine Candidate Items for Systemic Lupus Erythematosus Classification Criteria

Sindhu R. Johnson, Dinesh Khanna, David Daikh, Ricard Cervera, Nathalie Costedoat-Chalumeau, Dafna D. Gladman, Bevra H. Hahn, Falk Hiepe, Jorge Sánchez-Guerrero, Elena Massarotti, Dimitrios T. Boumpas, Karen H. Costenbader, David Jayne, Thomas Dörner, Diane L. Kamen, Marta Mosca, Rosalind Ramsey-Goldman, Josef S. Smolen, David Wofsy, and Martin Aringer

ABSTRACT. ***Objective.*** Given the complexity and heterogeneity of systemic lupus erythematosus (SLE), high-performing classification criteria are critical to advancing research and clinical care. A collaborative effort by the European League Against Rheumatism and the American College of Rheumatology was undertaken to generate candidate criteria, and then to reduce them to a smaller set. The objective of the current study was to select a set of criteria that maximizes the likelihood of accurate classification of SLE, particularly early disease.

Methods. An independent panel of international SLE experts and the SLE classification criteria steering committee (conducting SLE research in Canada, Mexico, United States, Austria, Germany, Greece, France, Italy, and Spain) ranked 43 candidate criteria. A consensus meeting using nominal group technique (NGT) was conducted to reduce the list of criteria for consideration.

Results. The expert panel NGT exercise reduced the candidate criteria for SLE classification from 43 to 21. The panel distinguished potential "entry criteria," which would be required for classification, from potential "additive criteria." Potential entry criteria were antinuclear antibody (ANA) ≥ 1:80 (HEp-2 immunofluorescence), and low C3 and/or low C4. The use of low complement as an entry criterion was considered potentially useful in cases with negative ANA. Potential additive criteria included lupus nephritis by renal biopsy, autoantibodies, cytopenias, acute and chronic cutaneous lupus, alopecia, arthritis, serositis, oral mucosal lesions, central nervous system manifestations, and fever.

Conclusion. The NGT exercise resulted in 21 candidate SLE classification criteria. The next phases of SLE classification criteria development will require refinement of criteria definitions, evaluation of the ability to cluster criteria into domains, and evaluation of weighting of criteria. (First Release April 15 2019; J Rheumatol 2019;46:721–6; doi:10.3899/jrheum.180478)

Key Indexing Terms:
SYSTEMIC LUPUS ERYTHEMATOSUS
CLASSIFICATION CRITERIA
NOMINAL GROUP TECHNIQUE
CONSENSUS METHODS

From the Division of Rheumatology, Department of Medicine, Toronto Western Hospital, University of Toronto, Toronto; Division of Rheumatology, Department of Medicine, Mount Sinai Hospital, Toronto; University Health Network, and Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada; Division of Rheumatology, Department of Medicine, University of Michigan, Ann Arbor, Michigan; University of California, Los Angeles, Los Angeles; University of California, San Francisco, California; Brigham and Women's Hospital, Boston; Harvard Medical School, Boston, Massachusetts; Medical University of South Carolina, Charleston, South Carolina; Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA; Hospital Clínic, Barcelona, Spain; AP-HP, Cochin Hospital, Internal Medicine Department, Centre de référence maladies auto-immunes et systémiques rares d'île de France, Paris; Université Paris Descartes-Sorbonne Paris Cité, Paris; INSERM U 1153, Center for Epidemiology and Statistics Sorbonne Paris Cité (CRESS), Paris, France; Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin; Berlin Institute of Health, Department of Rheumatology and Clinical Immunology, Berlin; University Medical Center and Faculty of Medicine Carl Gustav Carus, TU Dresden, Dresden, Germany; Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico; National and Kapodestrian University of Athens, and Biomedical Research Foundation of the Athens Academy, Athens, Greece; Department of Medicine, University of Cambridge, Cambridge, UK; University of Pisa, Pisa, Italy; Medical University of Vienna, Vienna, Austria.

This study was jointly supported by the European League Against Rheumatism and the American College of Rheumatology. Dr. Johnson is supported by a Canadian Institutes of Health Research New Investigator Award. Dr. Khanna was supported by a grant from the US National Institutes of Health/ National Institute of Arthritis and Musculoskeletal and Skin Diseases K24 AR 063120.

S.R. Johnson, MD, PhD, FRCPC, Division of Rheumatology, Department of Medicine, Toronto Western Hospital, Mount Sinai Hospital, and Institute of Health Policy, Management and Evaluation, University of Toronto; D. Khanna, MD, MS, Division of Rheumatology, Department of Medicine, University of Michigan; R. Cervera, MD, PhD, FRCP, Hospital

Systemic lupus erythematosus (SLE) is a complex, systemic autoimmune disease characterized by heterogeneity in disease manifestations and prognosis. Classification criteria are used to identify more homogeneous groups of patients for inclusion in clinical trials and observational studies[1]. With the support of the European League Against Rheumatism (EULAR) and the American College of Rheumatology (ACR), development of new classification criteria for SLE is currently under way[2]. A secondary goal of this initiative is to classify individuals with SLE earlier in this disease course. In line with ACR and EULAR standards, the SLE classification criteria development process was designed to consist of 4 phases, with balanced use of expert-based and data-driven methods meeting the standards set by the ACR and EULAR[1,3,4,5].

In Phase 1 of criteria development, positive antinuclear antibodies (ANA) were evaluated as a potential entry criterion for SLE classification[6]. Through systematic review and metaregression of the literature, a minimum titer of 1:80 on the indirect immunofluorescence HEp-2 ANA test resulted in 97.8% sensitivity and acceptable specificity for SLE. This suggested that ANA at this titer may constitute a reasonable entry criterion for SLE classification, provided that patients who were historically positive would be counted as positive. However, given that a positive ANA at this titer has only a limited specificity, classification of SLE requires further disease characteristics to achieve a high degree of specificity[6].

Clínic, Barcelona; N. Costedoat-Chalumeau, MD, PhD, AP-HP, Cochin Hospital, Internal Medicine Department, Centre de référence maladies auto-immunes et systémiques rares d'île de France, and Université Paris Descartes-Sorbonne Paris Cité, and INSERM U 1153, CRESS; D.D. Gladman, MD, FRCPC, Division of Rheumatology, Department of Medicine, Toronto Western Hospital, University of Toronto; B.H. Hahn, MD, University of California, Los Angeles; F. 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; J. Sánchez-Guerrero, MD, MSc, Division of Rheumatology, Department of Medicine, Mount Sinai Hospital/University Health Network, University of Toronto, and Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán; E. Massarotti, MD, Brigham and Women's Hospital, and Harvard Medical School; D.T. Boumpas, MD, FACP, FACR, National and Kapodestrian University of Athens, and Biomedical Research Foundation of the Athens Academy; K.H. Costenbader, MD, MPH, Brigham and Women's Hospital, and Harvard Medical School; D. Daikh, MD, University of California, San Francisco; D. Jayne, MD FRCP, FRCPE, FMedSci, Department of Medicine, University of Cambridge; T. 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; D.L. Kamen, MD, MSCR, Medical University of South Carolina; M. Mosca, MD, PhD, University of Pisa; R. Ramsey-Goldman, MD, DrPH, Northwestern University Feinberg School of Medicine; J.S. Smolen, MD, Medical University of Vienna; D. Wofsy, MD, University of California, San Francisco; M. Aringer, MD, University Medical Center and Faculty of Medicine Carl Gustav Carus, TU Dresden.

Address correspondence to Dr. S.R. Johnson, Division of Rheumatology, Ground Floor, East Wing, Toronto Western Hospital, 399 Bathurst St., Toronto, Ontario M5T 2S8, Canada. E-mail: Sindhu.Johnson@uhn.ca

Accepted for publication November 30, 2018.

Aimed at maximizing the range of potential disease-specific criteria, Phase 1 also comprised 3 independent studies to generate a list of candidate items. First, a large international Delphi exercise of SLE experts nominated 145 candidate criteria[7]. The experts rated the criteria on a 1–9 scale for their acceptability for the classification of SLE. Items with a mean acceptability score of 6.5 were retained if at least 50% of participants rated the item acceptability at ≥ 7. None of the individual neuropsychiatric SLE (NPSLE) items made the inclusion thresholds of the expert Delphi exercise. However, based on comments during the Delphi exercise and a motion by patient representatives from Lupus Europe, the steering committee reached consensus that NPSLE was an important organ manifestation that needed further consideration. A provisional composite central nervous system (CNS) dysfunction criterion was formed. Using this expert-based approach, 40 items were retained for further consideration.

Second, a data-driven exercise evaluated features of patients in the first few years of their disease (early) and compared those who subsequently were diagnosed with SLE to those who were diagnosed with a mimicking disease[8,9]. A multicenter "early disease" cohort was established with data from Europe and North America. The results led to the addition of 3 criteria occurring with increased frequency in early SLE: arthralgias, fever, and fatigue. Third, as per EULAR recommendations, the patient perspective was specifically addressed in a cross-sectional survey of 339 German patients with SLE, focusing on manifestations experienced early in their disease[10]. Again, fatigue (89%), fever (54%), and arthralgias (87%) were supported as criteria for consideration in early disease.

Thus, a total of 43 candidate criteria were proposed for consideration in the next phase of criteria development. The 43 criteria needed to be reduced to a more manageable number and further refined. The primary objectives of our study, comprising Phase 2 of the SLE classification criteria development process, were to reduce the number of candidate criteria and to identify criteria that should be retained for the next phase, with the aim of selecting a set of items that maximizes the likelihood of accurate classification of SLE, particularly early disease. The results of our study informed Phase 3, where the relative contribution of each criterion to the classification of SLE and threshold for classification of SLE were assessed. In Phase 4, the draft criteria set will be refined in a derivation cohort, and then comparatively evaluated against previous criteria sets in a validation cohort.

MATERIALS AND METHODS

Candidate criteria. The 43 candidate criteria nominated from the Phase 1 studies were the following: ANA on HEp-2 cells with a pattern compatible with SLE, titer > 1:160; ANA-positive (any pattern) > 1:160; low C3 and C4; ANA-positive by HEp-2; low C3; lupus nephritis by renal biopsy with immune deposits; anti-dsDNA antibody; anti-Sm antibody glomeru-

lonephritis [dysmorphic urinary red blood cell (RBC) or urinary RBC casts (≥ 1 cast/high power field)]; acute, subacute, or chronic SLE rash (can include malar, discoid, subacute cutaneous lupus erythematosus); rash with dermoepidermal interface changes and immunoglobulin and/or complement deposition on immunofluorescence; persistent proteinuria (> 0.5 g/day); malar rash; active urine sediment (without urinary tract infection); serositis; arthritis; presence of multiple autoantibodies; CNS dysfunction (seizures, psychosis, chorea, myelitis, optic neuritis, stroke or acute confusional state); oral mucosal lesions on the hard palate; thrombocytopenia; leukopenia (< 4000/mm^3 on 2 or more occasions); antiphospholipid antibodies (lupus anticoagulant, anticardiolipin, anti-β_2–glycoprotein 1 antibody, or prolonged Russell's viper venom time); thrombocytopenia (severe); autoimmune hemolytic anemia; photosensitive rash; antiphospholipid antibody syndrome (clinical signs/history and antibodies); urine cellular casts; discoid rash; lymphopenia (< 1500/mm^3 on 2 or more occasions); positive lupus anticoagulant; pleural effusion; pleuritis; subacute cutaneous lupus erythematosus; alopecia with associated scalp inflammation; pericardial effusion; photosensitivity; Raynaud phenomenon; fever; lupus profundus; lymphopenia (< 1000/mm^3 on 2 or more occasions); arthralgia; and fatigue (Table 1).

Consensus method. Nominal group technique (NGT) is a structured consensus method for group decision making that facilitates contribution from all the participants rather than an individual expert, in a formalized manner[11,12]. This methodology allows for the incorporation of a spectrum of experience and knowledge. It stimulates constructive debate while reducing the potential bias of an influential opinion, and is best suited for topics where there is insufficient evidence[12,13]. This approach has been successfully applied in the development of other rheumatologic classification criteria (systemic sclerosis) and outcome measures[14,15,16]. The NGT includes assembly of an expert panel, premeeting individual rankings, and a consensus meeting.

Expert panel. Internationally recognized SLE experts for the NGT panel were purposively sampled from the international SLE community, endorsed by the SLE classification criteria steering committee and consecutively invited. Inclusion criteria were recognized expertise in SLE based on research and patient care, and representation of Europe and North America. Dr. Dinesh Khanna (DK) served as independent moderator of the exercise.

Premeeting ranking. The 43 criteria, with their mean and median appropriateness scores, and the proportional endorsement were sent to the NGT expert panel. The experts were asked to review the criteria and rank them in order of importance (1 = most important). Their task was phrased as, "We are developing criteria for classification of SLE for clinical trials and other research studies. One aim is to increase the inclusion of patients with early SLE, who are less likely to have manifestations related to longterm SLE and organ damage. We are not developing diagnostic criteria. The primary objectives of this exercise are to identify criteria that should be retained for the next phase and to reduce the number of candidate criteria.

"In a person with an uncertain diagnosis, which criteria most increase the likelihood that the patient has SLE? In making this judgment, you should ask yourself: If there are two patients identical in every other respect, and one has this extra feature and one doesn't, is the patient with the feature more likely to have SLE than the other? For example, if both patients have a history of deep vein thrombosis and one has had an upper-limb deep vein thrombosis, does this latter feature really increase the likelihood of the patient having SLE? If not, it is not helpful in classifying the patient as having SLE."

The experts were then asked to submit their rankings and comments. The data were anonymized and median (range) rankings were calculated for each criterion.

Consensus meeting. The expert panel and steering committee met face-to-face in a room with rectangular tables arranged in an open U shape with a flip chart and large computer screen at the open end of the tables[12]. The data evaluating the performance characteristics of ANA testing for consideration as an entry criterion[6], the Delphi exercise data[7], the early SLE and mimicker

Table 1. Premeeting ranking of 43 candidate criteria.

Item	Ranking, Median (Range)
Premeeting ranking of ANA and complement criteria	
ANA on HEp-2 cells with a pattern compatible with SLE, titer ≥ 1:160	1.5 (1–5)
ANA-positive (any pattern) ≥ 1:160	2 (1–5)
Low C3 and C4	3 (1–5)
ANA-positive by HEp-2	3 (1–5)
Low C3	3.5 (2–5)
Premeeting ranking of 38 candidate additive criteria	
Lupus nephritis by renal biopsy with immune deposits	1 (1–2)
Anti-dsDNA antibody	2 (1–8)
Anti-Sm antibody	3 (1–18)
Glomerulonephritis (dysmorphic urinary red blood cells or urinary red blood cell casts (≥ 1 cast/high-powered field)	6 (1–35)
Acute, subacute, or chronic lupus rash (can include malar, discoid, subacute cutaneous lupus)	6 (1–32)
Rash with dermoepidermal interface changes and Ig and/or complement deposition on immunofluorescence	8 (1–36)
Persistent proteinuria (> 0.5 g/day)	9 (2–36)
Malar rash	11 (1–25)
Active urine sediment (without urinary tract infection)	12 (1–38)
Serositis (clinical signs, or pleural or pericardial effusion by imaging)	12 (2–23)
Arthritis	13 (2–33)
Presence of multiple autoantibodies	14 (1–37)
CNS dysfunction (seizures, psychosis, chorea, myelitis, optic neuritis, stroke, or acute confusional state)	14 (3–31)
Urinary red blood cell casts (≥ 1 cast/hpf)	16 (2–36)
Oral mucosal lesions on the hard palate	17 (2–30)
Thrombocytopenia	17 (2–36)
Leukopenia (< 4000/mm^3 on 2 or more occasions)	17 (3–37)
Antiphospholipid antibodies (lupus anticoagulant, anticardiolipin, anti-β_2–glycoprotein 1, or prolonged Russell's viper venom time)	17.5 (2–33)
Thrombocytopenia (severe)	18 (1–32)
Autoimmune hemolytic anemia	18 (1–34)
Photosensitive rash	19 (2–30)
Antiphospholipid antibody syndrome (clinical signs/history + antibodies)	19 (2–38)
Urine cellular casts	20 (2–37)
Discoid rash	20 (3–26)
Lymphopenia (< 1500/mm^3 on 2 or more occasions)	20 (3–36)
Positive lupus anticoagulant panel	21 (2–34)
Pleural effusion	23 (2–36)
Pleuritis	23 (2–33)
Subacute cutaneous lupus erythematosus	24 (3–33)
Alopecia with associated scalp inflammation	24.5 (3–38)
Pericardial effusion	25 (2–37)
Photosensitivity	27 (3–34)
Raynaud phenomenon	27 (3–35)
Fever	28 (2–37)
Lupus profundus	28 (3–38)
Lymphopenia (< 1000/mm^3 on 2 or more occasions)	29 (3–37)
Arthralgia	32 (3–37)
Fatigue	35 (4–38)

The ANA and complement criteria comprise 5 of 43 candidate criteria. The candidate additive criteria comprise 38 of 43 candidate criteria. ANA: antinuclear antibody; SLE: systemic lupus erythematosus; CNS: central nervous system.

disease cohort data[8,9], and the premeeting rankings were presented. The NGT facilitator (DK) presented an overview of the NGT process[12].

In a round-robin fashion, panelists were asked to comment upon the candidate criteria presented one at a time to the entire group. No interactive discussion was conducted at this time. After each panelist had an opportunity to speak, a serial brief discussion was led by the moderator with the goal of clarification of points made. Deliberations, including the steering committee, ensued until consensus was achieved on the inclusion, exclusion, or revision of each criterion. For each round of discussion, the process required that the first person to speak was different from that in the previous round. In this way, all panelists had the opportunity to speak first and avoid the effect of strong personalities[12]. The process ensured that all participants had an opportunity to contribute.

Institutional ethics approval (17-5926) and consent were obtained for the conduct of this study.

RESULTS

Expert panel. The expert panel and steering committee comprised 19 members (47% female, 53% male) conducting SLE research in Canada, Mexico, the United States, Austria, Germany, Greece, France, Italy, and Spain, for 43% European and 57% North America representation.

Premeeting rankings. The premeeting rankings for potential entry criteria (ANA and complements) and potential additive criteria are presented in Table 1. Lupus nephritis by renal biopsy with immune deposits, dsDNA, and anti-Sm antibodies were ranked highest.

Consensus meeting. The NGT exercise reduced the candidate criteria for SLE classification from 43 to 21. The panel distinguished potential "entry criteria," which would be required for classification, from other potential "additive criteria," summarized in Table 2.

Entry criteria. The panel agreed that ANA should be an entry criterion, and based on the Phase 1 systematic review and metaregression data[6], have a threshold of ≥ 1:80 (by HEp-2 immunofluorescence). Accordingly, "ANA on HEp-2 cells with a pattern compatible with SLE," ANA at a titer of > 1:160, ANA-positive by HEp-2, and low C3 were excluded. It was recognized that perhaps up to 2% of patients with SLE have a negative ANA at some time. Excluding all patients with negative ANA would exclude some of the population of patients with SLE. The use of low complement levels (and low C3 and/or low C4) as an entry criterion was considered potentially useful in cases with negative ANA. However, the inclusion of low complement levels was controversial. It was felt that complement was important but should not be an entry criterion. Main arguments against low complements as an entry criterion were that many patients would not have low complements in the early phase of disease if they did not already have renal involvement, and that low C4 was often genetically determined.

Additive criteria. The panelists achieved consensus on criteria that would be excluded. Arthralgias, fatigue, and Raynaud phenomenon were not considered sufficiently specific.

The panelists queried whether criteria could be clustered into "buckets" that are clinically or physiologically related.

Table 2. Listing of entry criteria and additive criteria after the nominal group technique exercise.

Criteria
Entry criteria
Antinuclear antibody by HEp-2 immunofluorescence ≥ 1:80
Low C3 and/or low C4
Additive criteria
Lupus nephritis by renal biopsy with immune deposits
Rash with dermoepidermal interface changes and/or Ig and/or complement deposition on immunofluorescence
Anti-dsDNA antibody
Anti-Sm antibody
Presence of multiple autoantibodies*
aPL (LAC, aCL, anti-β_2–glycoprotein 1, or prolonged Russell's viper venom time)
Leukopenia (< 4000/mm^3 on 2 or more occasions)
Thrombocytopenia < 100,000 on 2 or more occasions
Autoimmune hemolytic anemia
Active urine sediment (without urinary tract infection)
Persistent proteinuria (≥ 0.5 g/day)
Acute cutaneous lupus: SLICC definition (includes subacute cutaneous lupus)
Chronic cutaneous lupus: SLICC definition
Alopecia with associated scalp inflammation
Arthritis*
Serositis (pleural, pericardial effusion, pleurisy, pericarditis, peritonitis)
Oral mucosal lesions on the hard palate
CNS manifestations (seizures, psychosis, chorea, myelitis, optic neuritis, stroke or acute confusional state)
Fever*

* To be defined. SLICC: Systemic Lupus International Collaborating Clinics; CNS: central nervous system; aPL: antiphospholipid antibodies; LAC: lupus anticoagulant; aCL: anticardiolipin.

For example, urinary RBC casts and urine cellular casts were seen as redundant with proteinuria. Following similar arguments, pericardial effusion, pleuritis, and pleural effusion were clustered into 1 criterion of serositis.

For lymphopenia and thrombocytopenia, the panelists agreed to remove "severe," and replace with thresholds as outlined in the Systemic Lupus International Collaborating Clinics (SLICC) criteria[11].

The panelists also agreed with clustering the CNS manifestations into 1 domain. The panel recommended to use CNS manifestations instead of CNS dysfunction, given that CNS manifestations commonly reflect inflammatory activity in the CNS.

For skin manifestations, the expert panel suggested reduction to 2 criteria according to the SLICC definition, that is, acute or subacute cutaneous lupus and chronic cutaneous lupus[17]. Accordingly, malar rash, discoid rash, photosensitive rash, subacute cutaneous lupus erythematosus, photosensitivity, and lupus profundus were removed from the candidate list. Several experts pointed out that some signs are important, but may lose specificity if used by non-experts, such as malar rash wrongly diagnosed in patients with rosacea. The panel agreed that several criteria needed

stringent definitions, particularly "presence of multiple autoantibodies," "arthritis," and "fever."

Additional discussion points. The expert panel discussed the differing level of importance of some criteria. The panel discussed differential weights for each criterion to indicate its importance. However, concern was raised about a system that was too computationally difficult to use in clinic. It was preferred to use a system of weighting that had computational ease.

DISCUSSION

In this NGT exercise, part of the second phase of the SLE classification criteria project supported by EULAR/ACR, candidate criteria for the classification of SLE were refined. Starting with 43 candidate criteria for SLE classification, the exercise resulted in 21 criteria, a more manageable number for creating a system of classification. However, 3 important issues were raised. First, there was concern regarding the lack of precise definitions for the candidate criteria. This would result in inconsistent application of the criteria, affecting the validity and reliability of the final classification system. While it was not within the scope of the NGT panel to devise these definitions, this defined a further important step.

Second, it was important to understand the validity of each of the candidate criteria, notably their individual sensitivity and specificity. While the expert panel largely agreed on the approximate sensitivity and specificity of items, it appeared evident that more work on this aspect was needed.

Third, the expert panel raised the issue of interdependency of items, proposing that some criteria might cluster into "buckets." This question had not been previously addressed in any SLE classification criteria set before. However, it became obvious in the discussion that this question will need to be addressed.

The next phase of work will therefore require identification of precise definitions for any criteria that have ambiguity. Potential solutions included development of an online, freely available reference guide with definitions and photographs. The group's recommendation was to look for established and widely accepted definitions for criteria items such as the American Rheumatism Association Glossary Committee Dictionary of the Rheumatic Diseases[18], other criteria sets, or other medical disciplines. In the absence of established definitions, the definitions for items should be explicitly stated in the new criteria system. The validity of each of the 21 retained criteria also needs to be evaluated, because a system of classification is only as strong as its weakest criterion[1,19]. The operating characteristics (sensitivity, specificity) of the items in both SLE cases and mimicking conditions are needed. Criteria with poor discrimination should be discarded[19].

Throughout the NGT exercise, novel concepts for SLE classification emerged. First was the notion of clustering criteria into "buckets." Historically, this has been done clinically using a body systems approach. However, the question arose: Is this methodologically appropriate? In a system of classification, items should be independent. The expert panel proposed the concept of hierarchical clustering of items into domains and subdomains. Prior to doing this, however, the relationship (correlation and interaction) of clinically related criteria will need to be evaluated to ascertain independence.

The expert panel expressed the general opinion that there are differences in the relative importance of some criteria over others for classification. Lupus nephritis by renal biopsy with immune deposits, and to a lesser degree, antibodies to dsDNA and Sm were ranked highly. Fever was considered a potential criterion to distinguish cases and controls in early disease but attributed with comparatively lower importance. The next phase should evaluate the relative weight of each criterion for classification while maintaining computational ease.

In the steering committee deliberations following the NGT exercise, there was some discussion on whether it was disappointing that the set of candidate items contains no surprises, maybe with the exception of fever. However, item generation has been as broad as possible, and the reduction in items has rigorously followed methods that have been scientifically established. Whether alternative unbiased genetic or mRNA approaches will lead to different insights remains to be seen[20].

Finally, consideration will need to be given to the effect of ANA as an entry criterion. In the Phase 1 systematic review of the literature including 13,080 subjects diagnosed with SLE, 95.9% were ANA-positive by indirect immunofluorescence on HEp-2 cells[6]. In the Phase 1 early SLE cohort, 99.5% of the 389 SLE patients were ANA-positive[8]. The Phase 1 Delphi study of international SLE experts found 58% do not feel comfortable and an additional 19% were uncertain about classifying SLE in the absence of ever having a positive ANA[7]. Together, these data support the decision to use ANA as an entry criterion. However, research is needed to evaluate the numbers of patients with a diagnosis of SLE who are ANA-negative, particularly those with hypocomplementemia. Subsequent work may require considerations to appropriately classify this subset of patients with SLE.

The NGT exercise identified a core set of candidate criteria with the intended goal of maximizing the likelihood of accurate classification of SLE, with the added motivation of discriminating early disease. The next phases of SLE classification criteria development will refine definitions, consider hierarchical clustering of items into domains and subdomains, evaluate their independence and relationships within domains, ascertain item weights, and consider different thresholds for established SLE versus disease earlier in its course. The performance of the final SLE criteria set will then be comparatively evaluated in a multiethnic, international cohort against previous SLE classification criteria.

REFERENCES

1. 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.
2. Aringer M, Dorner T, Leuchten N, Johnson SR. Toward new criteria for systemic lupus erythematosus-a standpoint. Lupus 2016; 25:805-11.
3. Felson DT, Anderson JJ. Methodological and statistical approaches to criteria development in rheumatic diseases. Baillieres Clin Rheumatol 1995;9:253-66.
4. Dougados M, Betteridge N, Burmester GR, Euller-Ziegler L, Guillemin F, Hirvonen J, et al; EULAR. EULAR standardised operating procedures for the elaboration, evaluation, dissemination, and implementation of recommendations endorsed by the EULAR standing committees. Ann Rheum Dis 2004;63:1172-6.
5. Dougados M, Gossec L. Classification criteria for rheumatic diseases: why and how? Arthritis Rheum 2007;57:1112-5.
6. Leuchten N, Hoyer A, Brinks R, Schoels M, Schneider M, Smolen J, et al; Systemic Lupus Erythematosus Classification Criteria Steering Committee. Performance of antinuclear antibodies for classifying systemic lupus erythematosus: a systematic literature review and meta-regression of diagnostic data. Arthritis Care Res 2018;70:428-38.
7. Schmajuk G, Hoyer BF, Aringer M, Johnson SR, Daikh DI, Dorner T; SLE classification criteria steering committee and the international SLE expert panel of the initiative. Multicenter Delphi exercise to identify important key items for classifying systemic lupus erythematosus. Arthritis Care Res 2018;70:1488-94.
8. Mosca M, Costenbader KH, Johnson SR, Lorenzoni V, Sebastiani GD, Hoyer BF, et al. Brief report: how do patients with newly diagnosed systemic lupus erythematosus present? a multicenter cohort of early systemic lupus erythematosus to inform the development of new classification criteria. Arthritis Rheumatol 2019;71:91-8.
9. Touma Z, Costenbader KH, Johnson SR, Hoyer BF, Navara S, Bonfa E, et al. Do patients with SLE at onset differ from mimickers? A comparison of clinical and serological manifestations in a multicenter cohort to inform the development of new classification criteria for SLE [abstract]. Ann Rheum Dis 2016;75 Suppl 2:558.
10. Leuchten N, Milke B, Winker-Rohlfing B, Daikh D, Dorner T, Johnson SR, et al ; on behalf of the SLE Classification Criteria Steering Committee. Early symptoms of systemic lupus erythematosus (SLE) recalled by 339 SLE patients. Lupus 2018;279:1431-6.
11. Delbecq AL, Van de Ven AH, Gustafson DH. Group techniques for program planning, a guide to nominal group and Delphi processes. First ed. Glenview: Scott, Foresman and Company; 1975.
12. Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. Semin Arthritis Rheum 2011;41:95-105.
13. Johnson SR, O'Brien KK. Qualitative methods in systemic sclerosis research. J Rheumatol 2016;43:1265-7.
14. Fransen J, Johnson SR, van den Hoogen F, Baron M, Allanore Y, Carreira PE, et al. Items for developing revised classification criteria in systemic sclerosis: Results of a consensus exercise. Arthritis Care Res 2012;64:351-7.
15. Goldsmith CH, Boers M, Bombardier C, Tugwell P. Criteria for clinically important changes in outcomes: development, scoring and evaluation of rheumatoid arthritis patient and trial profiles. OMERACT Committee. J Rheumatol 1993;20:561-5.
16. Isenberg DA, Rahman A, Allen E, Farewell V, Akil M, Bruce IN, et al. BILAG 2004. Development and initial validation of an updated version of the British Isles Lupus Assessment Group's disease activity index for patients with systemic lupus erythematosus. Rheumatology 2005;44:902-6.
17. Petri M, Orbai AM, Alarcon 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.
18. American Rheumatism Association Glossary Committee. Dictionary of the rheumatic diseases. New York: Contact Associates International Ltd.; 1982.
19. Johnson SR, Fransen J, Khanna D, Baron M, van den Hoogen F, Medsger TA Jr, et al. Validation of potential classification criteria for systemic sclerosis. Arthritis Care Res 2012;64:358-67.
20. Johnson SR, Hinchcliff M, Asano Y. Controversies: molecular vs. clinical systemic sclerosis classification. J Scleroderma Relat Disord 2016;1:277-85.