


## ORIGINAL RESEARCH

## Validation of the 2019 EULAR/ACR classification criteria for systemic lupus erythematosus in an academic tertiary care centre

Irene Schmidtman,<sup>1</sup> Daniel Kraus ,<sup>2</sup> Arndt Weinmann,<sup>2</sup> Katharina Pütz,<sup>2</sup> Paul Claßen,<sup>2</sup> Eva Maria Schleicher,<sup>2</sup> Simone Cosima Boedecker-Lips,<sup>2</sup> Julia Weinmann-Menke<sup>2,3</sup>

**To cite:** Schmidtman I, Kraus D, Weinmann A, *et al.* Validation of the 2019 EULAR/ACR classification criteria for systemic lupus erythematosus in an academic tertiary care centre. *RMD Open* 2023;**9**:e003037. doi:10.1136/rmdopen-2023-003037

IS and DK contributed equally. SCB-L and JW-M contributed equally.

Received 27 January 2023  
Accepted 22 June 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Institute for Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center of the Johannes Gutenberg University, Mainz, Germany

<sup>2</sup>1st Department of Medicine, University Medical Center of the Johannes Gutenberg University, Mainz, Germany

<sup>3</sup>Research Center for Immunotherapy (FZI), University Medical Center of the Johannes Gutenberg University, Mainz, Germany

**Correspondence to**

Dr Julia Weinmann-Menke; Julia.Weinmann-Menke@unimedizin-mainz.de

**ABSTRACT**

**Objectives** To assess the sensitivity and specificity of the 2019 EULAR/American College of Rheumatology (ACR) classification criteria for systemic lupus erythematosus (SLE) in outpatients at an academic tertiary care centre and to compare them to the 1997 ACR and the 2012 Systemic Lupus International Collaborating Clinics criteria.

**Methods** Prospective and retrospective observational cohort study.

**Results** 3377 patients were included: 606 with SLE, 1015 with non-SLE autoimmune-mediated rheumatic diseases (ARD) and 1756 with non-ARD diseases (hepatocellular carcinoma, primary biliary cirrhosis, autoimmune hepatitis). The 2019 criteria were more sensitive than the 1997 criteria (87.0% vs 81.8%), but less specific (98.1% vs 99.5% in the entire cohort and 96.5% vs 98.8% in patients with non-SLE ARD), resulting in Youden Indexes for patients with SLE/non-SLE ARD of 0.835 and 0.806, respectively. The most sensitive items were history of antinuclear antibody (ANA) positivity and detection of anti-double-stranded deoxyribonucleic acid (dsDNA) antibodies. These were also the least specific items. The most specific items were class III/IV lupus nephritis and the combination of low C3 and low C4 complement levels, followed by class II/IV lupus nephritis, either low C3 or low C4 complement levels, delirium and psychosis, when these were not attributable to non-SLE causes.

**Conclusions** In this cohort from an independent academic medical centre, the sensitivity and specificity of the 2019 lupus classification criteria were confirmed. Overall agreement of the 1997 and the 2019 criteria was very good.

**INTRODUCTION**

Systemic lupus erythematosus (SLE) is an enigmatic disease. The name was most likely coined in the early Middle Ages, if not earlier, to denote the devouring of flesh like a wolf's bite in affected patients.<sup>1</sup> However, just like a wolf may prowl through the woods and never

**WHAT IS ALREADY KNOWN ON THIS TOPIC**

⇒ The 2019 EULAR/American College of Rheumatology (ACR) classification of systemic lupus erythematosus (SLE) have been designed and validated to ensure highly specific selection of patients with SLE for clinical studies. Data are scarce on the application of this classification in clinical practice and how it performs compared with the 1997 ACR classification.

**WHAT THIS STUDY ADDS**

⇒ This study applies the 2019 and the 1997 classification criteria to a large cohort of patients with SLE and non-SLE from an academic tertiary care centre. The sensitivity of the 2019 criteria was somewhat lower, and the specificity was comparable to the original validation cohort. There was good overall agreement of the 2019 classification with the 1997 classification.

**HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY**

⇒ This study underscores the applicability of the 2019 EULAR/ACR classification in clinical practice.

be seen, the disease notoriously eludes diagnosis and may be hard to grasp.

In order to facilitate the identification of patients with lupus for clinical trials, classification criteria have been devised and revised over the years. Early work to arrive at a useful classification began in 1962,<sup>2</sup> was published in 1971,<sup>3</sup> and subsequently revised. The 1997 classification that was adopted by the American College of Rheumatology (ACR)<sup>4</sup> was widely accepted and used for clinical trials (as well as for the diagnosis at the bedside) worldwide for some 20 years. This classification was a modification of an earlier version which had been published in 1982 and was also widely used for many years.<sup>5</sup> In 1997,

the immunological criterion was updated. However, the updated classification itself was never validated until the Systemic Lupus International Collaborating Clinics (SLICC) published an alternative classification in 2012 and compared its performance to the 1997 classification.<sup>6</sup> When applied against the validation cohort of the SLICC classification, the 2012 SLICC classification was more sensitive than the 1997 ACR classification (97% vs 83%), but less specific (84% vs 96%),<sup>6</sup> thereby increasing the risk of including patients without SLE in lupus trials, which would dilute any treatment effect. Of note, in the original 1982 publication, sensitivity and specificity both were 96%.<sup>5</sup>

In 2019, the EULAR and the ACR jointly published new classification criteria that incorporated state-of-the-art knowledge about the role of serological findings in SLE.<sup>7,8</sup> The working group employed sophisticated methodology. The new classification was meticulously derived from a carefully curated cohort of patients with lupus and non-lupus, and the initial publication included a carefully assembled validation cohort.<sup>7</sup> The non-lupus subjects in this cohort had ‘conditions mimicking SLE’, and prior to both derivation and validation, all subjects were centrally adjudicated for the absence or presence of SLE.

The sensitivity of the 2019 EULAR/ACR classification was improved over the 1997 version and comparable to the 2012 SLICC classification. The specificity of the 2019 classification was 93% in the validation cohort, which was equal to the 1997 ACR classification<sup>4</sup> when applied to the same cohort, and superior to the 2012 SLICC classification, which, interestingly, exhibited the same specificity as in the initial publication (84%).<sup>6</sup> Overall, the 2019 criteria had a larger combined sensitivity and specificity than the 1997 and the 2012 criteria (1.90 vs 1.76 and 1.80, respectively, in the validation cohort).<sup>7</sup>

Importantly, the 2019 classification declared a history of antinuclear antibody (ANA) positivity as an entry criterion. This excludes a small number of patients with lupus who have never had a positive ANA test from being classified as SLE, which the authors deemed acceptable for the purpose of recruiting for clinical trials.<sup>7</sup>

Derivation and validation of the 2019 EULAR/ACR were extremely well done, and the new classification was implemented immediately by rheumatologists worldwide. But how does the classification perform in daily clinical practice, where recruitment for studies takes place? Here we report the application of the criteria in a cohort of patients with SLE, patients with non-SLE ARD, and patients with hepatic diseases from a large outpatient clinic of a university medical centre in Germany.

## METHODS

### Subjects

This observational cohort study included patients with a diagnosis of SLE as well as patients with other autoimmune-mediated rheumatic diseases (ARD) and

patients from the hepatology clinic (non-ARD) of a single academic medical centre in Germany. The 1997 ACR and the 2019 EULAR/ACR classification systems were applied prospectively to all lupus subjects as they presented to the clinic. Lupus criteria are recorded if and only if the signs and symptoms cannot be attributed to an alternative diagnosis, as adjudicated by the patient’s rheumatologist. For patients without lupus, classification was performed retrospectively by chart review.

### Patient and public involvement

Patients were not involved in the design and execution of this study.

### Diagnosis of SLE

All cases of suspected SLE in the rheumatology clinic were discussed by two experienced physicians. The diagnosis was made or dismissed by consensus. If no agreement was reached, a third rheumatologist was consulted.

### Statistical analysis

Sensitivity and specificity of the classification systems as well as individual items were computed in the usual manner. The 1997 ACR and the 2019 EULAR/ACR classification systems were compared using McNemar’s test and by Cohen’s kappa. All analyses were carried out with SAS V.9.4.

## RESULTS

### Study population

The study cohort comprised 3377 subjects: 606 from the lupus clinic, 1015 patients without SLE from the rheumatology clinic and 1756 patients without ARD (table 1). Of these, 474 had a diagnosis of vasculitis, 421 had rheumatoid arthritis and 120 had undifferentiated connective tissue disease; 1172 had hepatocellular carcinoma, 253 had primary biliary cirrhosis, 331 had autoimmune hepatitis. In the SLE cohort, 88.3% were women as opposed to 50.1% in the non-SLE cohort (table 1). The subjects were predominantly Caucasian.

### Performance of the 2019 versus the 2012 and the 1997 classification systems

In the current SLE cohort, the 2019 EULAR/ACR criteria were more sensitive than both the 2012 SLICC and the 1997 ACR criteria (87% vs 85.3% and 81.8%, respectively; table 2). Specificity was somewhat lower than both the 2012 and the 1997 systems when the 2019 EULAR/ACR criteria were used, regardless of the cohort that was considered as a non-lupus reference: In patients with non-SLE ARD, specificities were 96.5%, 98.1% and 98.8% with the 2019, the 2012 and the 1997 criteria, respectively (table 2). In patients without ARD, specificities were 99.1%, 99.5% and 99.8%, respectively (table 2). The resulting Youden Indexes for the 2019 and the 1997 classifications were 0.835 and 0.806; the combined sensitivities and specificities<sup>7</sup> were 183.5% and 180.6%,

**Table 1** Characteristics of the study cohort

	SLE	Non-SLE
n	606	2771
Female, n (%)	535 (88.3)	1388 (50.1)
Age in years, median (IQR)	31.5 (22.8–43.6)	60.9 (50.3–69.6)
Patients with ARD		1015 (36.6%)
Rheumatoid arthritis, n (%)		421 (41.5)
Vasculitis, n (%)		474 (46.7)
Undifferentiated connective tissue disease, n (%)		120 (11.8)
Patients without ARD		1756 (63.4%)
Autoimmune hepatitis, n (%)		331 (18.8)
Hepatocellular carcinoma, n (%)		1172 (66.7)
Primary biliary cirrhosis, n (%)		253 (14.4)

ARD, autoimmune-mediated rheumatic diseases; SLE, systemic lupus erythematosus .

respectively, when using the non-SLE ARD cohort as the reference.

Overall, the agreement of the 1997 ACR and the 2019 EULAR/ACR classification systems was very good with a kappa coefficient of 0.88. However, the 2019 criteria are more likely to detect lupus than the 1997 criteria: In 2.6% of all cases from the combined cohorts, the 2019 criteria would detect lupus where the 1997 criteria would not (table 3). Conversely, in 0.6% of cases, the 1997 criteria would detect lupus where the 2019 criteria would not (p<0.0001 by McNemar’s test).

**Table 3** Cross-tabulation of the 1997 ACR and the 2019 EULAR/ACR classification in the entire cohort (n=3377)

	2019 EULAR/ACR negative	2019 EULAR/ACR positive	Row total
1997 ACR negative	2777 82.2%	89 2.6%	2866 84.9
1997 ACR positive	21 0.6%	490 14.5%	511 15.1
Column total	2798 82.9%	579 17.1%	3377 100.0%

P value<0.0001 by McNemar’s test. Kappa coefficient: 0.88 (95% CI, 0.86 to 0.90).  
ACR, American College of Rheumatology.

When the entry criterion of a history of ANA positivity was disregarded, sensitivity of the 2019 criteria increased from 87% to 89.9%, while specificity decreased from 98.1% to 97.2% (table 4).

**Items’ performance of the 2019 classification system**

When the individual items of the 2019 EULAR/ACR classification were analysed for the current cohort, the entry criterion of a history of ANA positivity had a sensitivity of 94.9% (95% CI, 92.8% to 96.5%; table 4), making ANA positivity the single most sensitive item. The next most sensitive item was detection of anti-double-stranded deoxyribonucleic acid (dsDNA) antibodies (78.9%; 95% CI, 75.4% to 82.1%; table 4).

Among the 31 (5.1%) of patients with SLE who did not have documented ANA positivity, 14 had dsDNA antibodies and 18 had malar rash. Three had lupus nephritis, and in 13 cases, complement (C3/C4) was depleted.

Two items of the 2019 classification were highly specific for SLE: class III/IV lupus nephritis and the combination of low C3 and low C4 complement levels were each 100%

**Table 2** Sensitivity and specificity of the 2019 EULAR/ACR, the 2012 SLICC criteria and the 1997 ACR criteria when applied to the SLE cohort, the non-SLE ARD cohort and the non-ARD cohort: 95% CIs are given in parentheses

Cohort	2019 EULAR/ACR	2012 SLICC	1997 ACR
Sensitivity	87.0% (84.0% to 89.5%)	85.3% (82.2% to 88.0%)	81.8% (78.5% to 84.8%)
Specificity (non-SLE ARD)	96.5% (95.1% to 97.5%)	98.1% (97.1% to 98.9%)	98.8% (97.9% to 99.4%)
Vasculitis	97.9% (96.2% to 99.0%)	98.3% (96.7% to 99.3%)	99.6% (98.5% to 99.9%)
Rheumatoid arthritis	99.3% (97.9% to 99.9%)	100% (99.1% to 100%)	100% (99.1% to 100%)
Undifferentiated connective disease	89.3% (80.6% to 95.0%)	92.9% (85.1% to 97.3%)	94.0% (86.7% to 98.0%)
Specificity (non-ARD)	99.1% (98.5% to 99.5%)	99.5% (99.0% to 99.8%)	99.8% (99.5% to 100%)
Autoimmune hepatitis	98.5% (96.5% to 99.5%)	99.1% (97.4% to 99.8%)	99.7% (98.3% to 100%)
Primary biliary cirrhosis	99.6% (97.8% to 100%)	100% (98.6% to 100%)	100% (98.6% to 100%)
Hepatocellular carcinoma	99.1% (98.4% to 99.6%)	99.5% (98.9% to 99.8%)	99.8% (99.4% to 100%)

ACR, American College of Rheumatology; ARD, autoimmune-mediated rheumatic diseases; SLE, systemic lupus erythematosus; SLICC, Systemic Lupus International Collaborating Clinics.

**Table 4** Performance of individual criteria

Criterion	SLE		Non-SLE		Sensitivity	Specificity
	Pos.	Neg.	Pos.	Neg.	Sensitivity with 95% CI	Specificity with 95% CI
Criteria sets						
1997 ACR classification	496	110	15	2756	81.8% (78.5% to 84.8%)	99.5% (99.1% to 99.7%)
2019 EULAR/ACR classification	527	79	52	2719	87.0% (84.0% to 89.5%)	98.1% (97.5% to 98.6%)
2019 EULAR/ACR classification without entry criterion	545	61	77	2694	89.9% (87.3% to 92.2%)	97.2% (96.5% to 97.8%)
Entry criterion						
History of ANA positivity	575	31	938	1833	94.9% (92.8% to 96.5%)	66.1% (64.4% to 67.9%)
Immunology domains						
Anti-dsDNA antibodies	478	128	246	2525	78.9% (75.4% to 82.1%)	91.1% (90.0% to 92.2%)
Anti-Sm antibodies	70	536	18	2753	11.6% (9.1% to 14.4%)	99.4% (99.0% to 99.6%)
Low C3 and C4	267	339	0	2771	44.1% (40.1% to 48.1%)	100.0% (99.9% to 100.0%)
Low C3 or C4	463	143	2	2769	76.4% (72.8% to 79.7%)	99.9% (99.7% to 100.0%)
Antiphospholipid antibodies	147	459	25	2746	24.3% (20.9% to 27.9%)	99.1% (98.7% to 99.4%)
Clinical domains						
Proteinuria	175	431	109	2662	28.9% (25.3% to 32.7%)	96.1% (95.3% to 96.8%)
Class III/IV nephritis	78	528	0	2771	12.9% (10.3% to 15.8%)	100.0% (99.9% to 100.0%)
Class II/V nephritis	44	562	3	2768	7.3% (5.3% to 9.6%)	99.9% (99.7% to 100.0%)
Autoimmune haemolysis	67	539	4	2767	11.1% (8.7% to 13.8%)	99.9% (99.6% to 100.0%)
Leucopenia	102	504	150	2621	16.8% (13.9% to 20.1%)	94.6% (93.7% to 95.4%)
Thrombocytopenia	78	528	354	2417	12.9% (10.3% to 15.8%)	87.2% (85.9% to 88.4%)
Acute lupus erythematosus	242	364	11	2760	39.9% (36.0% to 44.0%)	99.6% (99.3% to 99.8%)
Subacute lupus erythematosus	10	596	6	2765	1.7% (0.8% to 3.0%)	99.8% (99.5% to 99.9%)
Discoid lupus erythematosus	166	440	14	2757	27.4% (23.9% to 31.1%)	99.5% (99.2% to 99.7%)
Non-scarring alopecia	65	541	25	2746	10.7% (8.4% to 13.5%)	99.1% (98.7% to 99.4%)
Oral ulcers	116	490	17	2754	19.1% (16.1% to 22.5%)	99.4% (99.0% to 99.6%)
Pleural/pericardial effusion	117	489	109	2662	19.3% (16.2% to 22.7%)	96.1% (95.3% to 96.8%)
Acute pericarditis	63	543	21	2750	10.4% (8.1% to 13.1%)	99.2% (98.8% to 99.5%)
Joint involvement	262	344	52	2719	43.2% (39.2% to 47.3%)	98.1% (97.5% to 98.6%)
Fever	46	560	48	2723	7.6% (5.6% to 10.0%)	98.3% (97.7% to 98.7%)
Seizure	32	574	22	2749	5.3% (3.6% to 7.4%)	99.2% (98.8% to 99.5%)
Delirium	0	606	3	2768	0.0% (0.0% to 0.6%)	99.9% (99.7% to 100.0%)
Psychosis	15	591	2	2769	2.5% (1.4% to 4.0%)	99.9% (99.7% to 100.0%)

Sensitivities and specificities are given with 95% CI. LL, lower limit of the 95% CI. UL, upper limit of the 95% CI. aPL, presence of antiphospholipid antibodies (ie, detection of  $\beta$ 2-glycoprotein I antibodies and/or cardiolipin antibodies and/or positive lupus anticoagulant assay).

ACR, American College of Rheumatology; ANA, anti-nuclear antibody; dsDNA, double-stranded deoxyribonucleic acid; SLE, systemic lupus erythematosus.

specific with narrow CIs (table 4), followed by class II/V lupus nephritis, either low C3 or low C4 complement levels, delirium, psychosis and autoimmune haemolysis, that each had 99.9% specificity.

Considering the individual non-SLE diagnoses, specificities of low C3 and low C4 were 100% for each diagnosis. Specificities of either low C3 or low C4 were 100% for all diagnoses but two: vasculitis and undifferentiated

connective tissue disease, where one patient from both groups met this criterion.

As expected, the least specific item was a history of ANA positivity (66.6%), followed by thrombocytopenia (87.2%; table 4). Interestingly, even anti-dsDNA antibodies were only 91.1% specific (table 4). Leucopenia was 94.6% specific. All other items had specificities of at least 96.1% (table 4).

## DISCUSSION

The 2019 EULAR/ACR criteria for the classification of SLE are the most recent and up-to-date effort to provide a clear framework for the identification of patients with SLE that can be included in clinical trials. SLE is a heterogeneous disease, more a chameleon than a wolf, and this makes it particularly challenging to design and carry out trials of novel medications.<sup>9</sup> Lupus trials are notoriously prone to fail to demonstrate benefit.<sup>10</sup> Being able to clearly define the study cohort is one important prerequisite to pave the way to success.<sup>11</sup>

For the purpose of generating homogeneity, the 2019 EULAR/ACR classification includes a history of ANA positivity as an entry criterion.<sup>7</sup> This provided for maximum specificity, which was one goal of the EULAR/ACR working group, at the expense of decreased sensitivity.<sup>7</sup> Importantly, the classification criteria were not designed to make the diagnosis of SLE, which remains a clinical decision.<sup>8</sup> Nonetheless, all of the items in the classification are typical features of SLE, and as a matter of fact, SLE classification criteria have been useful as an adjunctive tool to make the clinical diagnosis ever since the first inception.

Here we applied the new criteria to a patient cohort from a large academic centre in Germany with a wide range of rheumatological as well as a number of important non-rheumatological diagnoses that are managed in our outpatient clinic. The specificity of the 2019 criteria, when applied to our non-SLE ARD cohort, was higher than the specificity in the original 2019 validation cohort (96.5% vs 93%, respectively)<sup>7</sup> as well as the original derivation and validation cohorts of the 2012 SLICC classification (92% and 84%, respectively).<sup>6</sup> When applied to the non-ARD cohort, specificity was even higher (99.1%); however, these patients without ARD are not typically screened for study recruitment. In our patients with SLE, sensitivity was 87%, which is considerably lower than the 96% sensitivity in the original validation cohort.<sup>7</sup> Thirteen per cent of patients who were diagnosed with SLE at our institution do not meet the 2019 EULAR/ACR criteria. One possible explanation could be that local experts tended to make the diagnosis in patients which might not have been regarded as patients with SLE by the adjudication committee of the EULAR/ACR working group. If so, this would mean that fewer patients, but more likely those with a clearly non-equivocal diagnosis of SLE would be identified when screening the already diagnosed patients with lupus for clinical trials—which was one of the aims of developing novel classification criteria.<sup>8</sup> Systemic lupus is a heterogenous disease. If stringent classification criteria were confused with clinical decision-making, many patients with chronic and debilitating systemic disease might not be managed in a way that soothes their ill-feeling and improves their long-term prognosis.

The combined sensitivities and specificities of the 2019 criteria and the 1997 criteria were 183.5% and 180.6%,

respectively. In the original validation cohort, these numbers were 190% and 176%.<sup>7</sup> Thus, the overall performance of the 2019 classification was lower, whereas the performance of the 1997 classification was better in our cohort than in the validation cohort.

The current study is not the first to externally validate the 2019 EULAR/ACR criteria.<sup>12</sup> However, to our knowledge, with 606 patients with SLE and 2771 patients without SLE, our cohort is by far the largest that has been published to date. In a recent analysis of the Australian Lupus Registry and Biobank that included 394 patients with lupus and 123 patients with non-SLE ARD, sensitivity and specificity of the 2019 criteria were 94.9% and 87.8%, respectively,<sup>13</sup> which is the inverse of the sensitivity and specificity in our cohort. The numbers were similar in a subset of patients with early disease ( $\leq 15$  months).<sup>13</sup> In a Spanish study of 79 patients with long-standing disease (mean duration, 15 years), sensitivity was 86.1%, similar to the sensitivity in our cohort.

When the diagnostic performances of the individual items were analysed, low C3 and/or C4 complement levels, lupus nephritis and delirium and psychosis had very high specificities (100% or 99.9%). This contrasts with the items' performance of the combined original derivation and validation cohorts, especially with regard to complement levels that were only highly specific when both C3 and C4 were reduced (specificity, 96%), but not when only one level was reduced (specificity, 83%).<sup>14</sup> We cannot tell with certainty why only two patients without SLE had either low C3 or low C4, resulting in this striking difference to the original derivation and validation cohorts, where more patients without SLE exhibited some kind of complement depletion.

When the attribution rule was not applied, joint involvement had very low specificity (82.9%); when properly attributed, specificity rose to 98.1%. This underscores the importance of correct attribution of signs and symptoms to SLE; signs and symptoms must not be taken into account if there is an alternative explanation that is more plausible than lupus.<sup>14</sup>

As already pointed out, classification criteria are not intended to be used to make a diagnosis. Notwithstanding, if the 2019 criteria were indeed used to diagnose patients with SLE at our centre, 13% of patients would not be diagnosed. With the 1997 ACR criteria, the sensitivity of 81.4% means that almost one in five patients who were already diagnosed and treated for SLE would not be considered to have SLE, had the classification criteria been applied as diagnostic criteria. This highlights the importance of clinical judgement in the management of SLE.

Still, clinical judgement may not always be accurate, especially in cases of early SLE that may be mistaken for other conditions, including 'undifferentiated' connective tissue disease. This may result in specificities that are lower than expected, especially against the background of the published literature.

The 2019 classification introduced a history of ANA detection as an entry criterion for the classification as SLE. This increases specificity from 97.2% to 98.1%, although at the cost of a loss of sensitivity, which decreases from 92.4% to 87% when the entry criterion is applied. As remarked on by the authors of the original publication,<sup>7</sup> maximum specificity is desired in order for the criteria to serve their purpose, which is to identify a homogeneous patient population for the inclusion in clinical trials. The decreased sensitivity must be kept in mind when using the criteria as a support tool for clinical decision-making. The classification criteria are not diagnostic criteria, and to rely solely on these criteria will leave some patients with SLE undiagnosed and, consequentially, untreated.

To our knowledge, this is the largest study of the performance of the 2019 classification in an unselected cohort of patients from an academic tertiary care centre, which regularly participates in clinical trials of patients with lupus. As such, it provides important information about the actual implementation of this new classification system, which is highly relevant for the selection of a homogeneous group of patients for these trials. We deliberately chose to also include patients from the hepatology clinic at our centre in order to better understand the performance of the classification. Autoimmune hepatitis and primary biliary cirrhosis are two autoinflammatory diseases whose signs and systems might also be mistaken for lupus if these patients are referred to a rheumatologist by a primary care provider. The subjects with hepatocellular cancer are included in our continuous scientific data collection, and we did not exclude them from the data set for the current study in order to be able to better judge how the lupus classification discriminates against clearly non-inflammatory disease.

Our study has the limitation that it was a single-centre study. At the same time, this may be regarded as a strength of this analysis: The data presented herein stem from the application of SLE classification criteria in everyday practice, and as such provide a good estimate of the performance of the criteria when actually recruiting study subjects.

Taken together, the current, large study confirms the high specificity of the 2019 EULAR/ACR criteria, while at the same time demonstrating good overall agreement with the 1997 ACR criteria.

**Twitter** Daniel Kraus @bovender\_de

**Contributors** IS analysed the data. DK wrote and revised the manuscript. AW programmed the clinical database. KP collected data and assisted with the analysis. PC collected data and assisted with the analysis. EMS collected data and assisted with the analysis. JW-M contrived the project, collected and analysed the data and revised the manuscript. JW-M is responsible for the overall content as guarantor.

**Funding** This work was supported by the Deutsche Forschungsgemeinschaft (DFG), grant no. WE5779/2-3, to JW-M. EMS is supported by the Clinician Scientist Fellowship 'Else Kröner Research College: 2018\_Kolleg.05'. SCB-L is supported by the Clinician Scientist Fellowship 'TransMed Jumpstart Program: 2019\_A72' which is funded by the Else Kröner Fresenius Foundation.

**Competing interests** JW-M received honoraria from GlaxoSmithKline, AstraZeneca, Otsuka and Novartis unrelated to this work. DK received honoraria from AstraZeneca and Novartis unrelated to this work. The other authors declare nothing to disclose.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by Independent Ethics Committee of the state of Rhineland-Palatinate, nos. 837.467.13 (9152-F) and 2019-14695. Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** No data are available. Not applicable.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Daniel Kraus <http://orcid.org/0000-0002-6351-3706>

#### REFERENCES

- Virchow R. Historische Notizen Über lupus. *Archiv f Pathol Anat* 1865;32:139–43.
- Cohen AS, Canoso JJ. Criteria for the classification of systemic lupus erythematosus--status 1972. *Arthritis Rheum* 1972;15:540–3.
- Cohen AS, Reynolds WE, Franklin EC, et al. Preliminary criteria for the classification of systemic lupus erythematosus. *Bull Rheum Dis* 1971;21:643–8.
- Hochberg MC. Updating the American college of rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40:1725.
- Tan EM, Cohen AS, Fries JF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982;25:1271–7.
- Petri M, Orbai A-M, Alarcón GS, et al. Derivation and validation of the systemic lupus International collaborating clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum* 2012;64:2677–86.
- Aringer M, Costenbader K, Daikh D, et al. European League against rheumatism/American college of rheumatology classification criteria for systemic lupus erythematosus. *Arthritis Rheum* 2019;71:1400–12.
- Aringer M. EULAR/ACR classification criteria for SLE. *Semin Arthritis Rheum* 2019;49:S14–7.
- Murphy G, Isenberg DA. Author correction: new therapies for systemic lupus erythematosus - past imperfect, future tense. *Nat Rev Rheumatol* 2019;15:509.
- Thurman JM. BLISS in the treatment of lupus nephritis. *Clin J Am Soc Nephrol* 2021;16:969–71.
- Dall'Era M, Bruce IN, Gordon C, et al. Current challenges in the development of new treatments for lupus. *Ann Rheum Dis* 2019;78:729–35.
- Aringer M, Costenbader K, Dörner T, et al. Advances in SLE classification criteria. *J Autoimmun* 2022;132:102845.
- Tan BCH, Tang I, Bonin J, et al. The performance of different classification criteria for systemic lupus erythematosus in a real-world rheumatology department. *Rheumatology (Oxford)* 2022;61:4509–13.
- Aringer M, Brinks R, Dörner T, et al. European League against rheumatism (EULAR)/American college of rheumatology (ACR) SLE classification criteria item performance. *Ann Rheum Dis* 2021;80:775–81.