EXTENDED REPORT

Comparison of the American-European Consensus Group Sjögren’s syndrome classification criteria to newly proposed American College of Rheumatology criteria in a large, carefully characterised sicca cohort

Astrid Rasmussen,¹ John A Ice,³ He Li,¹,² Kiely Grundahl,¹ Jennifer A Kelly,¹ Lida Radfar,³ Donald U Stone,⁴ Kimberly S Hefner,⁵ Juan-Manuel Anaya,⁶ Michael Rohrer,⁷ Rajaram Gopalakrishnan,⁸ Glen D Houston,² David M Lewis,² James Chodosh,⁹ John B Harley,¹⁰,¹¹ Pamela Hughes,¹² Jacen S Maier-Moore,¹ Courtney G Montgomery,¹ Nelson L Rhodus,¹³ A Darise Farris,¹ Barbara M Segal,¹⁴ Roland Jonsson,¹⁵,¹⁶ Christopher J Lessard,¹,² R Hal Scofield,¹,¹⁷,¹⁸ Kathy L Moser Sivils¹

Handling editor Tore K Kvien

ABSTRACT

Objective To compare the performance of the American–European Consensus Group (AECG) and the newly proposed American College of Rheumatology (ACR) classification criteria for Sjögren’s Syndrome (SS) in a well-characterised sicca cohort, given ongoing efforts to resolve discrepancies and weaknesses in the systems.

Methods In a multidisciplinary clinic for the evaluation of sicca, we assessed features of salivary and lacrimal gland dysfunction and autoimmunity as defined by tests of both AECG and ACR criteria in 646 participants. Global gene expression profiles were compared in a subset of 180 participants.

Results Application of the AECG and ACR criteria resulted in classification of 279 and 268 participants with SS, respectively. Both criteria were met by 244 participants (81%). In 26 of the 35 AECG+/ACR participants, the minor salivary gland biopsy focal score was ≥1 (74%), while nine had positive anti-Ro/La (26%). There were 24 AECG−/ACR+ who met ACR criteria mainly due to differences in the scoring of corneal staining. All patients with SS, regardless of classification, had similar gene expression profiles, which were distinct from the healthy controls.

Conclusions The two sets of classification criteria yield concordant results in the majority of cases and gene expression profiling suggests that patients meeting either set of criteria are more similar to other SS participants than to healthy controls. Thus, there is no clear evidence for increased value of the new ACR criteria over the old AECG criteria from the clinical or biological perspective. It is our contention, supported by this report, that improvements in diagnostic acumen will require a more fundamental understanding of the pathogenic mechanisms than is at present available.

INTRODUCTION

Sjögren’s Syndrome (SS) is a chronic, systemic disease that may be second only to rheumatoid arthritis in prevalence among the rheumatic autoimmune diseases.¹,² The principal manifestations of the disease are dry eyes and dry mouth resulting from immune-mediated damage and dysfunction of the lacrimal and salivary glands³–⁴ which develop a characteristic lymphocytic infiltrate that can be objectively measured with a focus score.⁵ Approximately 67% of patients with SS have circulating autoantibodies to anti-Ro (SSA) and/or anti-La (SSB).⁶ Extraglandular manifestations, which include vasculitis, peripheral neuropathy, renal tubular acidosis, pulmonary involvement, lymphoproliferative disease and/or immunological abnormalities, are present in a subset of patients and found most commonly among those with high levels of anti-Ro and anti-La autoantibodies.⁷⁸ The diagnosis of SS commonly requires a multidisciplinary approach and may be difficult to establish. Sicca symptoms are common, non-specific, and there is no gold standard diagnostic test. For research purposes, 11 sets of classification criteria have been proposed since the mid-1960s.⁹–¹⁹ The last of these, the 2002 revised American–European Consensus Group (AECG) Classification Criteria, have had widespread acceptance and adoption in clinical and research studies of SS, having been cited >1500 times.²⁰ They consist of six criteria, two subjective and four objective (table 1).¹⁹ In 2012, the American College of Rheumatology (ACR) endorsed a new set of preliminary criteria proposed by the Sjögren’s International Collaborative Clinical Alliance (sicca).¹¹²² These criteria are centred around three objective features (table 1).

We undertook this study to compare the new ACR criteria to the revised AECG criteria in a cohort of participants with sicca symptoms that have been carefully evaluated for SS.

METHODS

Participant recruitment

The participating individuals were evaluated in the Sjögren’s Research Clinic at Oklahoma Medical...
Table 1 Comparison of the Revised American–European Consensus Group (AECG) Classification criteria and the American College of Rheumatology (ACR) Classification criteria for Sjögren’s Syndrome (SS)

<table>
<thead>
<tr>
<th>AECG classification*</th>
<th>ACR classification†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inclusion criteria</strong></td>
<td><strong>Inclusion criteria</strong></td>
</tr>
<tr>
<td>I. Ocular symptoms: a positive response to at least one of the following questions:</td>
<td>None</td>
</tr>
<tr>
<td>1. Have you had daily, persistent, troublesome dry eyes for more than 3 months?</td>
<td>Keratoconjunctivitis sicca with ocular staining score ≥3 (assuming that individual is not currently using daily eye drops for glaucoma and has not had corneal surgery or cosmetic eyelid surgery in the last 5 years)</td>
</tr>
<tr>
<td>2. Do you have a recurrent sensation of sand or gravel in the eyes?</td>
<td></td>
</tr>
<tr>
<td>3. Do you use tear substitutes more than three times a day?</td>
<td></td>
</tr>
<tr>
<td>II. Oral symptoms: a positive response to at least one of the following questions:</td>
<td>None</td>
</tr>
<tr>
<td>1. Have you had a daily feeling of dry mouth for more than 3 months?</td>
<td></td>
</tr>
<tr>
<td>2. Have you had recurrently or persistently swollen salivary glands as an adult?</td>
<td></td>
</tr>
<tr>
<td>3. Do you frequently drink liquids to aid in swallowing dry food?</td>
<td></td>
</tr>
<tr>
<td>III. Ocular signs—that is, objective evidence of ocular involvement defined as a positive result for at least one of the following two tests:</td>
<td></td>
</tr>
<tr>
<td>1. Schirmer’s I test, performed without anaesthesia (≤5 mm in 5 min)</td>
<td></td>
</tr>
<tr>
<td>2. Rose Bengal score or other ocular dye score (≥4 according to van Bijsterveld’s scoring system)</td>
<td></td>
</tr>
<tr>
<td>IV. Histopathology: in minor salivary glands (obtained through normal appearing mucosa) focal lymphocytic sialectasias, evaluated by an expert histopathologist, with a focus score ≥1, defined as number of lymphocytic foci (which are adjacent to normal-appearing mucous acini and contain more than 50 lymphocytes) per 4 mm² of glandular tissue</td>
<td>Labial salivary gland biopsy exhibiting focal lymphocytic sialectasinitis with a focus score ≥1 focus/4 mm²</td>
</tr>
<tr>
<td>V. Salivary gland involvement: objective evidence of salivary gland involvement defined by a positive result for at least one of the following diagnostic tests:</td>
<td>None</td>
</tr>
<tr>
<td>1. Unstimulated whole salivary flow (≤1.5 mL in 15 min)</td>
<td></td>
</tr>
<tr>
<td>2. Parotid sialography showing the presence of diffuse sialectasias (punctate, cavity or destructive pattern), without evidence of obstruction in major ducts</td>
<td></td>
</tr>
<tr>
<td>3. Salivary scintigraphy showing delayed uptake, reduced concentration and/or delayed excretion of tracer</td>
<td></td>
</tr>
<tr>
<td>VI. Autoantibodies: presence in the serum of the following autoantibodies:</td>
<td></td>
</tr>
<tr>
<td>1. Antibodies to Ro (SSA) or La (SSB) antigens, or both</td>
<td></td>
</tr>
</tbody>
</table>

**Classification rules**

For primary SS:

In patients without any potentially associated disease, primary SS may be defined as follows:

A. The presence of any 4 of the 6 items is indicative of primary SS, as long as at least one item IV (histopathology) or VI (serology) is positive

B. The presence of any 3 of the 4 objective criteria items (ie, items III, IV, V, VI)

C. The classification tree procedure represents a valid alternative method for classification, although it should be more properly used in clinical-epidemiological survey

For Secondary SS:

In patients with a potentially associated disease (for instance, another well-defined connective tissue disease), the presence of item I or item II plus any 2 from among items III, IV and V may be considered as indicative of secondary SS

**Exclusion criteria**

1. Past head and neck radiation treatment
2. Hepatitis C infection
3. AIDS
4. Pre-existing lymphoma
5. Sarcoidosis
6. Graft versus host disease
7. Use of anticholinergic drugs (since a time shorter than 4-fold the half-life of the drug)

*Revised AECG classification criteria.19
†ACR criteria.21
§Sialography; this test was not performed in the current study.
‡Scintigraphy; this test was not performed in the current study.
ANA, antinuclear antibodies.

In order to be eligible for an appointment at the clinic, at least one ocular and one oral question had to be answered affirmatively. The exclusion criteria for evaluation at the clinic were also based on the recommendations of the AECG (table 1).19 Additionally, we excluded individuals who presented with known current pregnancy or inability to provide informed consent.
With very few exceptions, participants were evaluated in a single morning clinic visit using standardised protocols. Patients underwent an oral exam consisting of measurement of stimulated and timed whole unstimulated salivary flow (WUSF), a lip biopsy and collection and storage of saliva. Participant evaluation did not include sialography or scintigraphy. The ocular specialist performed ocular surface staining with lissamine green and fluorescein, an unanaesthetised Schirmer’s I test, and collection and storage of tears. The ocular vital dye score was determined using the quantitative dot-counting method rather than by descriptive features, and the score for each section was recorded independently before generating a final score for each eye. Blood samples were collected for general laboratory tests and extraction of DNA, RNA and serum. A physician completed a detailed history and physical examination, including general medical, rheumatological and neurological evaluations. If patients gave a history of a past diagnosis of rheumatoid arthritis, mixed connective tissue disease, systemic sclerosis, myositis, primary biliary cirrhosis, multiple sclerosis, or systemic lupus erythematosus, classification criteria for these illnesses were specifically ascertained by history, medical record review and testing for the corresponding autoantibodies.

All procedures were approved by the Oklahoma Medical Research Foundation and University of Minnesota Institutional Review Boards. Each participant provided written informed consent prior to entering the study.

Biopsy
The dentist performed lip biopsies to obtain minor salivary glands in all patients, unless slides from a previous biopsy were available and contained sufficient tissue for re-examination by our pathologists. A portion of each specimen was formalin-fixed and paraffin-embedded, sections were cut and stained with hematoxylin-eosin, while other fragments were cryologically preserved. Two dental pathologists reviewed the specimens independently; the results were compared and a consensus reading was generated. The lymphocytic infiltration of the glands was graded by focus score.

Clinical laboratory and serology
Anti-Ro/SSA and anti-La/SSB autoantibodies were determined by multiple methods. Additionally, all patients were tested for rheumatoid factor (RF), antinuclear antibodies (ANA), precipitins for autoantibodies associated with other connective tissue disorders, hepatitis C serology, complete blood count (CBC) with differential, immunoglobulin profile and urinalysis (see online supplementary text).

Classification
Each study participant was classified according to both the revised AECG, and to the newly proposed ACR criteria. We eliminated from analysis the participants that did not have results for all the features of both classification systems with the exception of sialography and scintigraphy (table 1).

Peripheral blood mRNA transcript measurements
Global gene expression profiles comprising transcript levels for >15 000 loci were compared in a subset of 180 participants (see online supplementary text).

Statistical analysis
Performance of the tests was assessed via sensitivity, specificity, positive predictive value and negative predictive value estimated by considering the AECG criteria as the ‘gold standard’, and summarising the results with exact binomial 95% CI. McNemar’s Test of paired samples was used to assess whether the two sets of criteria were significantly different with respect to dichotomous variables. The κ statistic was used to quantify the degree of agreement between the new classification criteria and the AECG criteria. Details of the statistical analyses for the gene expression data are available in the online supplementary text.

RESULTS
The initial cohort of participants evaluated at either the Sjögren’s Research Clinic at Oklahoma Medical Research Foundation or the Sjögren’s Clinic in the University of Minnesota comprised 837 individuals. Of these, 646 had all data points of both AECG and ACR classification criteria and, thus, constitute the study cohort. The demographic characteristics of both cohorts are comparable in makeup with respect to age, sex, race and ethnicity (see online supplementary table S1).

We tabulated the presence or absence of each of the six AECG classification criteria for SS and each of the three ACR criteria (summary in table 2; details in online supplementary table S2). Of the 646 study participants, 279 and 268 patients were classified as SS according to AECG and ACR criteria, respectively. Of the 303 participants classified by either system as SS, 244 (81%) individuals met both sets of criteria.

The comparison of the new ACR classification criteria with the AECG criteria (table 2) shows that they are not significantly different (McNemar’s test of paired samples: p=0.19) and there was a concordance rate of 0.81 (95% CI 0.77 to 0.86) based on the κ statistic. The analysis of the sensitivity, specificity, positive predictive value and negative predictive value of each set of criteria was done using the other criteria as the gold standard and

<table>
<thead>
<tr>
<th>Table 2 Comparison of classification by AECG criteria versus ACR criteria for Sjögren’s Syndrome (SS)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AECG criteria</strong></td>
</tr>
<tr>
<td>SS</td>
</tr>
<tr>
<td>ACR</td>
</tr>
<tr>
<td>SS</td>
</tr>
<tr>
<td>DNMC</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

p=0.19 by McNemar’s test; κ=0.81 (95% CI 0.77 to 0.86).
*Specificity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of the ACR criteria considering the AECG criteria as the ‘gold standard’. 
†Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of the AECG criteria considering the ACR criteria as the ‘gold standard’. 
ACR, American College of Rheumatology Criteria Classification Criteria for SS; AECG, American European Consensus Group Classification Criteria for Sjögren’s Syndrome; DNMC, did not meet criteria; SS, Sjögren’s Syndrome.
Clinical and epidemiological research

was similar for both classification systems. The sensitivity of the ACR criteria was 87.5 (95% CI 82.9 to 90.9) with a specificity of 93.4 (95% CI 90.3 to 95.7); the positive predictive value was 91.0 (95% CI 86.8 to 94.0) and the negative predictive value was 90.7 (95% CI 87.2 to 93.4). Thus, 12.5% (35 of 279) of participants classified as SS under the AECG criteria were not considered SS when evaluated by the ACR criteria; conversely, 8.9% (24 of 268) met only the ACR criteria.

The differences between how the two systems classified the sicca participants revolved around which objective measures of ocular and oral involvement were included in addition to the sicca syndrome. The top of the table shows patients classified as SS under the AECG criteria but not ACR criteria by having, in addition to either histopathological or serological criteria, subjective ocular and oral symptoms plus either an abnormal Schirmer’s test and/or an abnormal WUSF test (table 3). They did not meet the ACR criteria because they had only one of the histopathology or serology criteria but did not have an abnormal ocular staining examination.

Alternatively, there were 24 participants classified as SS by ACR criteria but not AECG criteria (ACR+/AECG− participants; table 3), and they met criteria mainly due to differences in the scoring of the ocular staining (n=17): the ACR criteria use the OSS score which is abnormal at ≥23 score out of 12 possible points, while the AECG criteria use the vBS score that is abnormal with a score ≥4 out of 9 possible points (figure 1). Seven of the ACR+/AECG− participants met the ACR criteria but not the AECG criteria by having positive ANA plus RF.

The performance of each individual test was assessed in three subsets of participants: (1) classified by ACR criteria, (2) classified by AECG criteria and (3) classified as having SS by either one or both sets of criteria (see online supplementary table S3). As expected, the tests performed consistently across all groups and in summary, the Schirmer’s I test had a low sensitivity (range 0.49–0.54) with higher specificity (0.71–0.73), while the WUSF had both low sensitivity (0.59–0.65) and specificity (0.52–0.57). On the other hand, the serology and histopathology performed well, with sensitivities of 0.63–0.64 and 0.84–0.86, respectively, and specificities of 0.94–0.96 and 0.89–0.95, respectively.

The most important difference in individual test performance was in the evaluation of keratoconjunctivitis sicca by ocular surface staining. The use of the AECG scoring system by vBS resulted in a sensitivity of 0.57–0.61 with a specificity of 0.70–0.71. The ACR OSS scoring very significantly improved the sensitivity (0.80–0.90) but at the expense of the specificity (0.45–0.51) (see online supplementary table S3). We compared the number of participants that had a positive score by one method versus the other and found highly significant differences (table 4). When assessing this difference in any patient that was classified as having SS (by either or both sets of criteria, n=303), 23% of those having a positive OSS did not have a positive vBS (p<1×10−6). This difference was 24% (p<1×10−6) if all participants were included, irrespective of whether they were classified as SS or not. An intermediate result was obtained if the OSS was considered abnormal at a cut-off of ≥4 rather than ≥3: 58 participants (13.5%) went from being OSS (+) to OSS (−), (p=0.001); 53 of these 58 were AECG (−).

When evaluating gene expression profiles for participants meeting only one set of criteria, those who met both sets of criteria, and healthy controls, we found that the participants who met criteria for SS by one or more sets of criteria tended to cluster together and were distinct from controls (figure 2). Furthermore, using low-stringency criteria designed to maximise determination of differences, we found no gene expression difference between the participants meeting both sets of criteria versus those meeting only one set of criteria.

DISCUSSION

The pathophysiological mechanisms underlying SS are still poorly understood, and accurately determining who does and does not have SS is difficult.23 In the clinical setting, the diagnosis of SS relies on interpreting and integrating all aspects of the patient’s history, test results and the expert opinion of the clinician. For research purposes, many classification systems have been proposed in the last few decades, and the coexistence of more than one system may lead to heterogeneity and confusion.

Table 3 Summary of patients classified as Sjögren’s Syndrome by only one set of criteria and not the other

<table>
<thead>
<tr>
<th>Histology</th>
<th>vBS score</th>
<th>OSS</th>
<th>Ro/La</th>
<th>ANA+RF</th>
<th>Schirmer’s</th>
<th>WUSF</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>AECG criteria only</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>5</td>
</tr>
<tr>
<td>26</td>
<td>0</td>
<td>9</td>
<td>14</td>
<td>27</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACR criteria only</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>22</td>
<td>4</td>
<td>7</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The differences between the participants meeting both sets of criteria (figure 1). Seven of the ACR+/AECG− participants met the ACR criteria but not the AECG criteria by having positive ANA plus RF.

The most important difference in individual test performance was in the evaluation of keratoconjunctivitis sicca by ocular surface staining. The use of the AECG scoring system by vBS resulted in a sensitivity of 0.57–0.61 with a specificity of 0.70–0.71. The ACR OSS scoring very significantly improved the sensitivity (0.80–0.90) but at the expense of the specificity (0.45–0.51) (see online supplementary table S3). We compared the number of participants that had a positive score by one method versus the other and found highly significant differences (table 4). When assessing this difference in any patient that was classified as having SS (by either or both sets of criteria, n=303), 23% of those having a positive OSS did not have a positive vBS (p<1×10−6). This difference was 24% (p<1×10−6) if all participants were included, irrespective of whether they were classified as SS or not. An intermediate result was obtained if the OSS was considered abnormal at a cut-off of ≥4 rather than ≥3: 58 participants (13.5%) went from being OSS (+) to OSS (−), (p=0.001); 53 of these 58 were AECG (−).

When evaluating gene expression profiles for participants meeting only one set of criteria, those who met both sets of criteria, and healthy controls, we found that the participants who met criteria for SS by one or more sets of criteria tended to cluster together and were distinct from controls (figure 2). Furthermore, using low-stringency criteria designed to maximise determination of differences, we found no gene expression difference between the participants meeting both sets of criteria versus those meeting only one set of criteria.

DISCUSSION

The pathophysiological mechanisms underlying SS are still poorly understood, and accurately determining who does and does not have SS is difficult.23 In the clinical setting, the diagnosis of SS relies on interpreting and integrating all aspects of the patient’s history, test results and the expert opinion of the clinician. For research purposes, many classification systems have been proposed in the last few decades, and the coexistence of more than one system may lead to heterogeneity and confusion.
in the interpretation of research studies. As has recently been highlighted by Vitali et al., the SS community should be striving for the common goal of reaching a final agreement on classification criteria for the disease. The first steps in this direction are to evaluate the performance of the new criteria and compare them to the currently used AECG criteria in external cohorts of patients and controls. Such comparison in our uniformly evaluated cohort of patients presenting with sicca helps serve this purpose.

This cohort has been evaluated in a homogeneous and standardised manner at two research clinics in Oklahoma and Minnesota by a multidisciplinary team of experts. The evaluation includes all the exams and laboratory procedures detailed in the AECG and ACR classification criteria, including the subjective components of the AECG criteria, and both ocular staining scoring systems. We did not assess the revised Japanese Ministry of Health criteria, because they are intended as an aid for clinical diagnosis and not for research classification which is the aim of our SS clinics. Additionally, they have not been tested in a non-Japanese population, and include additional invasive procedures which we felt were not justified for our participants.

It is relevant to note that there are differences in the enrolment strategy of our cohort in comparison to the sicca cohort. The most important difference is that while in the sicca cohort only 79% of the participants had both subjective dry eyes and dry mouth; the totality of our participants responded affirmatively to at least one ocular and one oral dryness question of the AECG criteria as evidence of symptoms of oral and ocular dryness.

The two clinics have so far evaluated 837 individuals, but only 646 for whom we had all the data points for both sets of criteria were included in the current analysis. Of these, 303 participants were classified as having SS by either one or both sets of criteria but almost 20% of the participants met only one set of criteria for SS, and not the other. This level of disagreement between the two classification systems is similar to that reported by Shiboski et al. These patients would have been excluded from any study based on only one of the classification methods; thus, knowing their characteristics becomes relevant to future research.

In the case of patients classified as having SS by the AECG criteria only (n=35), two-thirds of them had a minor salivary gland biopsy consistent with SS and the remaining one-third had positive Ro/La serology. They did not meet ACR criteria because they did not have keratoconjunctivitis sicca, and their objective measures of dryness were confined to either the biopsy or the serology but not both (table 3). While we have not used formal expert consensus methodology, we believe that most experts would agree that SS is present in a person presenting with subjective dry eyes, subjective dry mouth, a confirmatory minor salivary gland biopsy or positive Ro/La serology plus at least one additional objective measure of dryness, be it a positive Schirmer’s I test or an abnormal WUSF test.

Conversely, the individuals who met only the ACR criteria (n=24) did so because of alternative positive serology status or differences in the evaluation of keratoconjunctivitis sicca. Seven of them (29%) met criteria by having positive ANA/RF but negative anti-Ro/La as one of two criteria. Again, we have not
Figure 2. Assessment of the American–European Consensus Group (AECG) and ACR criteria using whole-blood gene expression profiling in Sjögren’s Syndrome (SS). Heat maps are displayed using fold changes for differentially expressed (DE) transcripts; overexpressed transcripts (FC>0) are bright yellow, while underexpressed transcripts (FC<0) are light blue. Data is displayed in rows and columns, with rows defined by transcripts and columns defined by individual samples. Colours above each column header denote healthy controls (blue; n=73) and cases (red, yellow, or green). For cases, the colour represents the criteria used to define SS: red denotes cases meeting only ACR criteria (n=4); yellow denotes cases meeting only AECG criteria (n=25); and green denotes cases meeting both ACR and AECG criteria (n=127). A Sjögren’s-specific set of DE transcripts was defined by comparing all SS cases regardless of classification criteria to healthy controls. Hierarchical clustering was performed with respect to transcripts and samples, and dendrograms generated to visualise sample clustering. In panel A, all samples are displayed, with cases and controls generally segregating. In panel B, hierarchical clustering was performed using the DE transcripts defined in panel A, but removing cases meeting both AECG and ACR criteria. Panel C further limits the clustering to those individuals meeting only ACR or AECG. Distinct clustering of patients meeting only ACR or AECG criteria was not observed, suggesting molecular similarity between cases defined by either ACR or AECG criteria.

done formal expert testing, but it is unlikely that these indi-
viduals would be considered to have SS based on expert opinion,
especially without information about sicca symptoms. The
remaining 71% met criteria for keratoconjunctivitis sicca by
ACR but not by AECG criteria.

It has been proposed that it would be useful to know if the
OSS developed for the ACR criteria can be substituted by the
AECG vBS.21 Few already established cohorts that we are aware of,
if any, are currently able to directly compare the perform-
ance of the vBS with the OSS. In cohorts that were evaluated
before the publication of the OSS in 2010, determining the OSS
would require access to the breakdown of the scoring of each
eye: individual scores for medial and lateral bulbar conjunctiva
and cornea plus the description of patches of confluent staining,
staining in the papillary area and presence of filaments
(figure 1). The vBS does not take into consideration these last
three features,24 which add three possible points to the score of
each eye in the case of the OSS. The vBS is considered abnormal
with a score of ≥4 out of 9 possible points24 while the OSS is
positive at ≥3 out of 12 points.23

We are in the unique position of having recorded separately
each of the 12 possible scoring points for each eye in all our
cohort participants. Thus, we were able to determine both their
vBS and OSS scores and compare the performance of each
system. To reduce interobserver and intraobserver variability
inherent to the vBS scoring system,24 the vital dye score for
each section of the ocular surface was determined using the
sicca dot counting method. While there are no studies validating
the conversion of this scoring method with the traditional vBS
technique, the two are similar; we felt that an objective scoring
method would be more meaningful and reproducible in the
context of multiple observers. As expected, participants were
more likely to have an abnormal OSS score than vBS, resulting in
~25% of patients having a positive OSS but negative vBS.
This difference was highly significant in all cohort participants
and in patients who were classified as having SS by one or both
sets of criteria (p<1×10⁻³⁸). The OSS is superior in including
true positive cases but has a poor performance ruling out those
who do not have SS (ie, it is very sensitive but has poor speci-
city); the opposite is the case for the vBS. It is pertinent to note
that only a minor proportion of cases of keratoconjunctivitis
sicca are due to SS.21 It remains to be seen how other prospect-
ive cohorts evaluate these two scoring systems vis à vis, in order
to determine what the optimal threshold should be. It is note-
worthy that one of the main goals of the development of new
classification criteria by the sicca consortium was to come up
with a system that has high specificity to avoid exposing
unaffected individuals to the potentially serious adverse effects
of novel investigational therapies.21

The two tests that performed the best across all comparison
groups were the minor salivary gland biopsy and anti-Ro/La ser-
ology, which performed similarly to reports in previous
studies.21,28 The Schirmer’s and WUSF tests while less useful in
distinguishing true SS patients from participants with
non-Sjögren’s sicca syndrome, are easy to perform and are non-
invasive. It has recently been suggested that more emphasis
should be given to tests that in addition to identifying true cases
and excluding unaffected individuals, can be done at early
stages, multiple times, and with minimal distress to the partici-

The Schirmer’s and the WUSF tests can be performed in a stand-
ard medical office without the need for sophisticated equipment
or medical specialists. Thus, patients can be assessed for subject-
ive dry eyes and dry mouth, the presence of autoantibodies
along with Schirmer’s and WUSF testing by a rheumatologist. If
AECG criteria are not met with such an assessment, then biopsy
and eye examination can be pursued. In some clinical care set-
goings or research situations that do not include exposing the
selected participants to the risk of significant adverse events
(such as some therapeutic trials), a stepwise approach such as
this may be useful and cost effective.

The comparison of the AEG criteria with the proposed
ACR classification demonstrates that neither system is clearly
superior to the other when classifying a patient with SS;
a finding already reported in the initial publication of the ACR
criteria.21 The lack of highly sensitive, specific and reproducible
criteria may, in part, be due to our current limited understand-
ing of SS physiopathology; such knowledge would provide the
most rational basis for disease classification. In the current
setting, the ACR criteria may be best suited for stricter studies
focused on high specificity to reduce the risk of drug-related
toxicity, while the AECG criteria may be applicable to broader
use, particularly in less risky medical research, or in non-
treatment clinical or translational research settings. Moreover
our findings of similar gene expression profiles across all pos-
sible patients affected by SS, which is different from what is
observed in healthy controls, supports our notion that modify-
ing classification using only clinical criteria is not likely to lead
to consequential improvements in our ability to identify patients
with SS. We believe that such improvements in diagnostic
acumen will require a more fundamental understanding of the
pathogenic mechanisms than is at present available.

Author affiliations
1Arthritis and Clinical Immunology Research Program, Oklahoma Medical Research
Foundation, Oklahoma City, Oklahoma, USA
2Department of Pathology, University of Oklahoma Health Sciences Center,
Oklahoma City, Oklahoma, USA
3Department of Oral Diagnosis and Radiology, University of Oklahoma College of
Dentistry, Oklahoma City, Oklahoma, USA
4Department of Ophthalmology, University of Oklahoma Health Sciences Center,
Oklahoma City, Oklahoma, USA
5Helfer Eye Care and Optical Center, Oklahoma City, Oklahoma, USA
6Center for Autoimmune Diseases Research (CREA), Universidad del Rosario,
Bogotá, Colombia
7Hard Tissue Research Laboratory, University of Minnesota School of Dentistry,
Minneapolis, Minnesota, USA
8Division of Oral Pathology, Department of Developmental and Surgical Science,
University of Minnesota School of Dentistry, Minneapolis, Minnesota, USA
9Department of Ophthalmology, Harvard Medical School, Boston, Massachusetts,
USA
10Division of Rheumatology, Cincinnati Children’s Hospital Medical Center,
Cincinnati, Ohio, USA
11US Department of Veterans Affairs Medical Center, Cincinnati, Ohio, USA
12Division of Oral and Maxillofacial Surgery, Department of Developmental and
Surgical Science, University of Minnesota School of Dentistry, Minneapolis, Minnesota, USA
13Department of Oral Surgery, University of Minnesota School of Dentistry, Minneapolis, Minnesota, USA
14Hennepin County Medical Center, Minneapolis, Minnesota, USA
15Broegelmann Research Laboratory, Department of Clinical Science, University of
Bergen, Bergen, Norway
16Department of Rheumatology, Haukeland University Hospital, Bergen, Norway
17Department of Medicine, University of Oklahoma Health Sciences Center,
Oklahoma City, Oklahoma, USA
18Department of Veterans Affairs Medical Center, Oklahoma City, Oklahoma, USA
19Department of Veterans Affairs Medical Center, Oklahoma City, Oklahoma, USA
20Department of Veterans Affairs Medical Center, Oklahoma City, Oklahoma, USA
21Department of Veterans Affairs Medical Center, Oklahoma City, Oklahoma, USA
22Department of Veterans Affairs Medical Center, Oklahoma City, Oklahoma, USA

Correction notice This article has been corrected since it was published Online
First. Occurrences of ‘STICCA’ have been changed to lower case (‘sicca’).
Clinical and epidemiological research

Acknowledgements We are grateful to all the individuals with SS and those serving as healthy controls who participated in this study. We would like to thank the following individuals for their help in the collection and ascertainment of the samples used in this study: Erin Rothrock, Judy Harris, Shuron Johnson, Sarah Ciof, Nicole Weber, Dominique Williams, Wes Daniels, Cheryl Pritchett-Frazee, Kylia Crouch, Laura Battiest, Justin Rodgers, James Robertson, Thuan Nguyen, Amanda Crosbie, Ellen James, Carolyn Meyer, Amber McElroy, Eshrat Emamian, Julie Emmer, Kristine Rohlf, Joanlis Leon, Anita Petersen, Daniele Hartle, Jill Novikze, Ward Ortman, Carl Espy, Beth Cobb, Gudlaug Kristjansdottir and Marnanne Eidsheim. We would also like to thank Stuart Glenn and Jared Ning for their ongoing assistance in developing and maintaining the computational infrastructure used to perform this study.

Contributors All authors of the manuscript contributed to: conception and design, or analysis and interpretation of data, drafting the article or revising it critically for important intellectual content and final approval of the version to be published.

Funding This publication was made possible by grants SR01 AR50782 (KLS), P50 AR0608040 (KLS, CJL, RHS, and ADP), SU19 AI02714 (KLS, CJL), SR01 DE018209 (KLS, JBH), SR37AD024717-22S1 (JBH, AR). The contents are the sole responsibility of the authors and do not necessarily represent the official views of the NIH. Additional funding was obtained from the Philipona Foundation (KLS) and the Oklahoma Medical Research Foundation (CJL and KLS). DUS received funding from an unrestricted grant from Research to Prevent Blindness to the University of Oklahoma Department of Ophthalmology.

Competing interests None.

Ethics approval Oklahoma Medical Research Foundation Internal Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The authors would consider requests for the data generated in this project through collaborative arrangements.

REFERENCES


