Comparison of performance of the 2016 ACR-EULAR classification criteria for primary Sjogren's syndrome with other sets of criteria in Japanese patients

OPEN ACCESS

EXTENDED REPORT

Comparison of performance of the 2016 ACR-EULAR classification criteria for primary Sjögren's syndrome with other sets of criteria in Japanese patients

Hiroto Tsuboi,[1,2] Shinya Hagiwara,[1] Hiromitsu Asashima,[1] Hiroyuki Takahashi,[1] Tomoya Hirota,[1] Hisashi Noma,[3] Hisanori Umehara,[2,4,5] Atsushi Kawakami,[2,6] Hideki Nakamura,[2,6] Hajime Sano,[2,7] Kazuo Tsubota,[2,8] Yoko Ogawa,[2,8] Etsuko Takamura,[2,9] Ichiro Saito,[2,10] Hiroko Inoue,[2,11] Seiji Nakamura,[2,12] Masafumi Moriyama,[2,12] Tsutomu Takeuchi,[2,13] Yoshiya Tanaka,[2,14] Shintaro Hirata,[2,14] Tsuneyo Mimori,[2,5] Isao Matsumoto,[1] Takayuki Sumida[1,2]

Linked

► http://dx.doi.org/10.1136/annrheumdis-2017-211378

CrossMark

ABSTRACT

Objectives To compare the performance of the new 2016 American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) classification criteria for primary Sjögren's syndrome (SS) with 1999 revised Japanese Ministry of Health criteria for diagnosis of SS (JPN), 2002 American-European Consensus Group classification criteria for SS (AECG) and 2012 ACR classification criteria for SS (ACR) in Japanese patients.

Methods The study subjects were 499 patients with primary SS (pSS) or suspected pSS who were followed up in June 2012 at 10 hospitals in Japan. All patients had been assessed for all four criteria of JPN (pathology, oral, ocular, anti-SS-A/SS-B antibodies). The clinical diagnosis by the physician in charge was set as the 'gold standard'.

Results pSS was diagnosed in 302 patients and ruled out in 197 patients by the physician in charge. The sensitivity of the ACR-EULAR criteria in the diagnosis of pSS (95.4%) was higher than those of the JPN, AECG and ACR (82.1%, 89.4% and 79.1%, respectively), while the specificity of the ACR-EULAR (72.1%) was lower than those of the three sets (90.9%, 84.3% and 84.8%, respectively). The differences of sensitivities and specificities between the ACR-EULAR and other three sets of criteria were statistically significant (p<0.001). Eight out of 302 patients with pSS and 11 cases out of 197 non-pSS cases satisfied only the ACR-EULAR criteria, compared with none of the other three sets.

Conclusions The ACR-EULAR criteria had significantly higher sensitivity and lower specificity in diagnosis of pSS, compared with the currently available three sets of criteria.

INTRODUCTION

Sjögren's syndrome (SS) is an autoimmune disease that affects mainly exocrine glands including the salivary and lacrimal glands, and is often associated with extraglandular manifestations, such as interstitial lung and kidney diseases, and neurological, haematological and musculoskeletal involvements.[1] It is characterised by lymphocytic infiltration into the exocrine glands and other organs, leading to dry mouth, dry eyes and various extraglandular symptoms. SS is subcategorised into primary SS (pSS) which is not associated with other well defined connective tissue diseases (CTDs), and secondary SS which is associated with other well defined CTDs.[2]

In Japan, the revised criteria for the diagnosis of SS proposed by the Japanese Ministry of Health (JPN) (1999),[3] as well as the American-European Consensus Group classification criteria for SS (AECG) (2002)[2] have been used commonly in both daily clinical practice and clinical studies in this decade. In 2012, the American College of Rheumatology (ACR) published the 2012 ACR classification criteria for SS, which were proposed by the Sjögren's International Collaborative Clinical Alliance (SICCA).[4] These three sets of criteria have also been applied for the diagnosis or classification of SS in Japan in the last 3 years. We previously analysed 694 Japanese patients with SS or suspected SS, and showed that the sensitivities of JPN, AECG and ACR in the diagnosis of SS were 79.6%, 78.6% and 77.5%, respectively, with respective specificities of 90.4%, 90.4% and 83.5%, when considering the clinical diagnosis as the 'gold standard'.[5] We concluded in that study the superiority of the JPN criteria in the diagnosis of SS in Japanese patients compared with the ACR and AECG criteria.[5]

Recently the 2016 new ACR-European League Against Rheumatism (EULAR) classification criteria for pSS (ACR-EULAR) were published.[6][7] Investigators from the SICCA team and the EULAR Sjögren's Task Force formed the International SS Criteria Working Group to develop this single set of classification criteria that combined features of the ACR and AECG criteria, based on methodology consistent with the current ACR and EULAR guidelines.[6][7] The working group adopted the methodology based on both data and expert clinical judgement, and finally defined the new classification criteria comprising five objective tests or items, and a total score of ≥4 as the cut-off for the diagnosis of pSS. The total score is derived from the sum of the weights assigned to each positive

test or item as follows: focal lymphocytic sialadenitis in labial salivary gland with Focus Score (FS) of ≥ 1 (based on number of foci/4 mm^2) and positive anti-SS-A/Ro serology with the highest weights (3 for each positive test), and Ocular Staining Score (OSS) of ≥ 5 (or Van Bijsterveld Score of ≥ 4) on at least one eye, Schirmer's test of ≤ 5 mm/5 min on at least one eye, and unstimulated whole saliva (UWS) flow rate of ≤ 0.1 mL/min with a weight of 1 for each positive test.[6 7]

Comparison of the above four sets of criteria (table 1) shows certain differences in the adopted items. In addition to the adopted items, the purpose of these criteria sets also differs. Importantly, the JPN criteria were formulated for the diagnosis of SS as the diagnostic criteria, while other three sets of criteria (the ACR-EULAR, AECG and ACR criteria) were formulated for research purposes as the classification criteria. The purpose of the present study was to compare the performance of the new ACR-EULAR criteria with the former sets of criteria, such as the JPN, AECG and ACR criteria in Japanese patients.

PATIENTS AND METHODS
Study population
The study subjects were 499 patients (38 men and 461 women) with the diagnosis of pSS or suspected pSS, who had been checked for all four criteria of the JPN (pathology, oral, ocular, anti SS-A/Ro and SS-B/La antibody), and were followed up in June 2012 at 10 hospitals across Japan (Kanazawa Medical University Hospital, Nagasaki University Hospital, Hyogo Medical University Hospital, Keio University Hospital, Tokyo Women's Medical University Hospital, Tsurumi University

Hospital, Kyushu University Hospital, University of Occupational and Environmental Health Hospital, Kyoto University Hospital, and University of Tsukuba Hospital), which form parts of the Research Team for Autoimmune Diseases, The Research Program for Intractable Disease of the Japan Ministry of Health, Labor and Welfare (MHLW).

Data collection and analysis
We collected clinical data through a questionnaire from the above 10 hospitals. We retrospectively examined the clinical diagnosis by the physician in charge, satisfaction of ACR-EULAR, JPN, AECG and ACR criteria. Because the OSS adopted in the ACR-EULAR and ACR criteria is not commonly performed in Japan, we regarded patients with Van Bijsterveld Score ≥ 4 in the Rose Bengal test, lissamine green test or fluorescein staining test to have satisfied OSS in the ACR-EULAR criteria, and patients who had positive Rose Bengal or lissamine green test (Van Bijsterveld Score ≥ 3) or fluorescein staining test to have satisfied OSS in the ACR criteria. Similarly, because numerous cases (116/499 cases) lacked results of the UWS, which was not adopted in the JPN criteria,[3] we regarded patients who had UWS ≤ 0.1 mL/min, gum test ≤ 10 mL/10 min or Saxon test ≤ 2 g/2 min to have satisfied low salivary volume in the ACR-EULAR and AECG criteria.

Moreover, we performed the subanalysis using 383 cases who were examined for UWS, excluding 116 cases who lacked results of the UWS mentioned above. We examined satisfaction for each criteria set more strictly in this subanalysis than in whole analysis of 499 cases. For salivary volume, we regarded patients who had gum test ≤ 10 mL/10 min or Saxon test ≤ 2 g/2 min to have satisfied decreased salivary volume in JPN criteria, while UWS ≤ 0.1 mL/min in the ACR-EULAR and AECG criteria. For ocular staining, we regarded patients with Van Bijsterveld Score ≥ 3 in the Rose Bengal test, lissamine green test or fluorescein staining test, and/or positive fluorescein staining test to have satisfied positive ocular staining in JPN and ACR criteria, while Van Bijsterveld Score ≥ 4 in AECG and ACR-EULAR criteria.

We considered the clinical diagnosis by the physician in charge as the 'gold standard' for the diagnosis of pSS in this study. In all cases, the diagnosis established by the physician in charge was based on clinical findings, laboratory and serological tests of blood and saliva samples, sialography, scintigraphy, and histopathological examination of biopsy material. We regarded the clinical diagnosis by the physician in charge to be appropriate for the 'gold standard', because the clinical diagnosis was decided by senior and experienced clinicians belonging to 10 hospitals which form parts of the Research Team for Autoimmune Diseases, The Research Program for Intractable Disease of the Japan MHLW described above. We compared the sensitivity and specificity between the ACR-EULAR, JPN, AECG and ACR criteria in the diagnosis of pSS.

Statistical analysis
The differences of sensitivities and specificities between all possible pairs of the four sets of criteria were evaluated using the McNemar's test and the Newcombe's square-and-add method. A p value <0.05 denoted the presence of a statistically significant difference.

RESULTS
Diagnosis of pSS and denial of pSS
None of the 499 patients had other well defined CTDs. pSS was diagnosed in 302 patients, whereas pSS was excluded in 197

Table 1 Comparison of items adopted in the 2016 ACR-EULAR, JPN, AECG and ACR criteria

Items	ACR-EULAR	JPN	AECG	ACR
Ocular symptoms	Not adopted	Not adopted	Adopted	Not adopted
Oral symptoms	Not adopted	Not adopted	Adopted	Not adopted
Ocular signs				
Schirmer's test	Adopted (1 point)	Adopted	Adopted	Not adopted
Ocular staining	Adopted (1 point)	Adopted	Adopted	Adopted
Labial salivary gland biopsy	Adopted (3 points)	Adopted	Adopted	Adopted
Salivary gland involvements				
Salivary secretion	Adopted (1 point)	Adopted	Adopted	Not adopted
Sialography	Not adopted	Adopted	Adopted	Not adopted
Scintigraphy	Not adopted	Adopted	Adopted	Not adopted
Autoantibodies				
SS-A/Ro	Adopted (3 points)	Adopted	Adopted	Adopted
SS-B/La	Not adopted	Adopted	Adopted	Adopted
ANA	Not adopted	Not adopted	Not adopted	Adopted
RF	Not adopted	Not adopted	Not adopted	Adopted

ACR, American College of Rheumatology ACR criteria for SS; ACR-EULAR, American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) classification criteria for pSS; AECG, The American-European Consensus Group classification criteria for SS; ANA, antinuclear antibody; JPN, The revised Japanese Ministry of Health criteria for the diagnosis of SS; pSS, primary SS; RF, rheumatoid factor; SS, Sjögren's syndrome; SS-A/Ro, anti-SS-A/Ro antibody; SS-B/La, anti-SS-B/La antibody.

patients by the physician in charge, and these judgements were considered as the 'gold standard' in the present study. For subanalysis using 383 cases that were examined for UWS, pSS was diagnosed in 203 patients, while pSS was excluded in 180 patients based on clinical judgements.

Sensitivity and specificity in diagnosis of pSS by different sets of criteria

For all the 499 patients, the sensitivities of the ACR-EULAR, JPN, AECG and ACR criteria in the diagnosis of pSS were 95.4% (95% CI 93.0% to 97.1%), 82.1% (79.6% to 84.1%), 89.4% (86.8% to 91.6%) and 79.1% (76.2% to 81.6%), respectively, considering the diagnosis by the physician in charge as the 'gold standard' (table 2). The respective specificities were 72.1% (68.4% to 74.7%), 90.9% (87.0% to 93.8%), 84.3% (80.2% to 87.6%) and 84.8% (80.3% to 88.5%) (table 2).

For subanalysis using 383 cases that were examined for UWS, the sensitivities of the ACR-EULAR, JPN, AECG and ACR criteria in the diagnosis of pSS were 94.1% (90.8% to 96.4%), 74.9% (71.3% to 77.6%), 85.7% (82.1% to 88.7%) and 79.8% (75.8% to 83.2%), respectively. The respective specificities were 76.7% (73.0% to 79.3%), 90.6% (86.5% to 93.7%), 86.1% (82.0% to 89.4%) and 81.1% (76.6% to 85.0%) (table 2).

The sensitivity of the ACR-EULAR criteria was statistically significantly higher than those of other three sets of criteria in both whole analysis and subanalysis (p<0.001) (table 3). The specificity of the ACR-EULAR criteria was statistically significantly lower than those of other three sets of criteria in whole analysis (p<0.001) (table 3). Although the specificity of the ACR-EULAR criteria was statistically significantly lower than those of the JPN and AECG criteria in subanalysis (p<0.001), the difference between the ACR-EULAR and ACR criteria was not statistically significant (p=0.117) (table 3).

These findings indicate that the ACR-EULAR criteria have higher sensitivity and lower specificity in the diagnosis of pSS, compared with the JPN, AECG and ACR in both whole analysis and subanalysis.

Agreement of ACR-EULAR criteria with the other three sets of criteria

Table 4 shows the satisfaction for these four sets of criteria in each case. The data showed that the ACR-EULAR criteria were satisfied by much more cases than the other three sets of criteria for both pSS (288 cases) and non-pSS (55 cases) groups. Many

pSS (220/302 cases, 72.8%) satisfied all four sets of criteria, while many non-pSS (135/197 cases, 68.5%) satisfied none of four sets of criteria (table 4). Although 8 non-pSS cases satisfied all four sets of criteria, 11 patients with pSS did not satisfy any set of criteria (table 4). There was no case that satisfied all the other three sets of criteria except for the ACR-EULAR criteria among both pSS and non-pSS groups (table 4).

Importantly, 8 out of the 302 patients with pSS diagnosed by the physician in charge and 11 cases out of the 197 clinically non-pSS cases satisfied only the ACR-EULAR criteria, compared with none of the other three sets of criteria. These 19 cases explained the low agreement between the ACR-EULAR criteria and the other three sets of criteria. Further analysis of positivity for each item adopted in the ACR-EULAR criteria among the 8 patients with pSS who satisfied only the ACR-EULAR criteria indicated that they had positive FS (62.5%, 5/8 cases) or positive SS-A/Ro (37.5%, 3/8 cases), together with decreased salivary (87.5%, 7/8 cases) or lacrimal (12.5%, 1/8 cases) secretion, resulting in a total score of 4 in these 8 patients (figure 1A). The 11 non-pSS cases who satisfied only the ACR-EULAR criteria had positive FS (54.5%, 6/11 cases) or positive SS-A/Ro (45.5%, 5/11 cases), together with decreased salivary secretion (100%, 11/11 cases), resulting in a total score of 4 in these 11 cases (figure 1B).

Considered together, the above analyses suggest that judgement by the ACR-EULAR criteria for both diagnosis and exclusion of pSS was different from those by the JPN, AECG and ACR criteria (table 4).

DISCUSSION

In the present study, we compared the sensitivity and specificity of the 2016 ACR-EULAR criteria for pSS with those of the JPN, AECG and ACR criteria, for the diagnosis of pSS using clinical data of 499 Japanese patients with pSS or suspected pSS. The results showed clearly that the 2016 ACR-EULAR criteria had higher sensitivity and lower specificity in the diagnosis of pSS, compared with the other three sets of criteria in both whole analysis (n=499) and subanalysis (n=383) using cases that were examined for UWS. Moreover, the degree of agreement of the ACR-EULAR criteria with the three sets of criteria for both diagnosis and exclusion of pSS was low. These results are different from those reported in a recent study by Shiboski et al,[6 7] which showed high sensitivity (96%, 95% CI 92% to 98%) and specificity (95%, 95% CI 92% to 97%) for the ACR-EULAR criteria and high agreement rate with both AECG

Table 2 Sensitivity and specificity for the diagnosis of pSS by the four sets of criteria

| Analysis | Criteria sets | Diagnosis by the physician in charge as the 'gold standard' | | | |
		Sensitivity (%)	95% CI	Specificity (%)	95% CI
Whole n=499	ACR-EULAR	95.4 (288/302)	93.0 to 97.1	72.1 (142/197)	68.4 to 74.7
	JPN	82.1 (248/302)	79.6 to 84.1	90.9 (179/197)	87.0 to 93.8
	AECG	89.4 (270/302)	86.8 to 91.6	84.3 (166/197)	80.2 to 87.6
	ACR	79.1 (239/302)	76.2 to 81.6	84.8 (167/197)	80.3 to 88.5
Subanalysis n=383	ACR-EULAR	94.1 (191/203)	90.8 to 96.4	76.7 (138/180)	73.0 to 79.3
	JPN	74.9 (152/203)	71.3 to 77.6	90.6 (163/180)	86.5 to 93.7
	AECG	85.7 (174/203)	82.1 to 88.7	86.1 (155/180)	82.0 to 89.4
	ACR	79.8 (162/203)	75.8 to 83.2	81.1 (146/180)	76.6 to 85.0

Of the 499 enrolled patients, pSS was diagnosed in 302 patients, whereas pSS was excluded in 197 patients by the physician in charge. For subanalysis using 383 cases that were examined for unstimulated whole saliva (UWS), pSS was diagnosed in 203 patients, while pSS was excluded in 180 patients based on clinical judgements.
ACR-EULAR, American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) classification criteria for pSS; AECG, American-European Consensus Group classification criteria for SS; JPN, The revised Japanese Ministry of Health criteria for the diagnosis of SS; pSS, primary SS; SS, Sjögren's syndrome.

Table 3 Differences of the sensitivities and specificities (with 95% CI) for the diagnosis of pSS among the four sets of criteria

	Competitor	JPN	AECG	ACR
Whole n=499				
Sensitivities	ACR-EULAR	13.2 (9.2 to 17.7) p<0.001	6.0 (3.0 to 9.4) p<0.001	16.2 (12.2 to 20.7) p<0.001
	JPN	–	−7.3 (−11.8 to −2.9) p=0.001	3.0 (−0.3 to 6.4) p=0.072
	AECG	–	–	10.3 (5.6 to 15.1) p<0.001
Specificities	ACR-EULAR	−18.8 (−24.7 to −13.1) p<0.001	−12.2 (−17.2 to −7.3) p<0.001	−12.7 (−18.6 to −6.8) p<0.001
	JPN	–	6.6 (0.8 to 12.5) p=0.024	6.1 (2.4 to 10.3) p=0.001
	AECG	–	–	0.0 (−6.9 to 5.9) p=0.873
Subanalysis n=383				
Sensitivities	ACR-EULAR	19.2 (13.7 to 25.2) p<0.001	8.4 (4.3 to 13.1) p<0.001	14.3 (9.6 to 19.6) p<0.001
	JPN	–	−10.8 (−17.3 to −4.4) p=0.001	−4.9 (−10.2 to 0.3) p=0.059
	AECG	–	–	5.9 (−0.3 to 12.1) p=0.058
Specificities	ACR-EULAR	−13.9 (−19.6 to −8.6) p<0.001	−9.4 (−14.4 to −4.8) p<0.001	−4.4 (−10.2 to 1.2) p=0.117
	JPN	–	4.4 (−1.5 to 10.5) p=0.131	9.4 (4.9 to 14.5) p<0.001
	AECG	–	–	5.0 (−1.6 to 11.6) p=0.128

Of the 499 enrolled patients, pSS was diagnosed in 302 patients, whereas pSS was excluded in 197 patients by the physician in charge. For subanalysis using 383 cases that were examined for unstimulated whole saliva (UWS), pSS was diagnosed in 203 patients, while pSS was excluded in 180 patients based on clinical judgements.
The p values and confidence limits were computed by the McNemar's test and the Newcombe's square-and-add method.
–, not examined; ACR-EULAR, American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) classification criteria for pSS; AECG, American-European Consensus Group classification criteria for SS; JPN, The revised Japanese Ministry of Health criteria for the diagnosis of SS; pSS, primary SS; SS, Sjögren's syndrome.

Table 4 Satisfaction of the ACR-EULAR, JPN, AECG and ACR criteria in clinically diagnosed pSS and non-pSS cases

Cases						
pSS	non-pSS		ACR-EULAR	JPN	AECG	ACR
220	8		O	O	O	O
14	0		O	O	O	X
11	8		O	O	X	O
7	3		O	X	O	O
0	0		X	O	O	O
0	1		O	O	X	X
27	19		O	X	O	X
1	5		O	X	X	O
2	0		X	O	O	X
0	1		X	O	X	O
0	0		X	X	O	O
8	11		O	X	X	X
1	0		X	O	X	X
0	1		X	X	O	X
0	5		X	X	X	O
11	135		X	X	X	X
Total 302	197	pSS	288	248	270	239
		non-pSS	55	18	31	30
		Total numbers of cases that satisfied each criteria set				

Of the 499 enrolled patients, pSS was diagnosed in 302 patients, whereas pSS was excluded in 197 patients by the physician in charge.
ACR-EULAR, American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) classification criteria for pSS; AECG, American-European Consensus Group classification criteria for SS; JPN, The revised Japanese Ministry of Health criteria for the diagnosis of SS; pSS, primary SS; SS, Sjögren's syndrome; O, satisfaction; X, non-satisfaction.

cases had positive FS (6/11 cases) or positive anti-SS-A/Ro (5/11 cases), and decreased salivary volume (11/11 cases). These 11 non-pSS cases might cause the low specificity of the ACR-EULAR criteria. However, one has to pay enough attention to the adopted methods for assessment of salivary volume in this study. As mentioned in the Patients and methods section, we regarded patients with UWS ≤0.1 mL/min, gum test ≤10 mL/10 min or Saxon test ≤2 g/2 min to have satisfied decreased salivary volume in the ACR-EULAR and AECG criteria, because numerous cases (116/499 cases) lacked the results of UWS, which was not adopted in the JPN criteria.[3] Importantly, 6 out of 11 non-pSS cases who satisfied only the ACR-EULAR criteria, were considered to have decreased salivary volume based on the gum and/or Saxon test instead of UWS (data not shown). These six cases might explain the low specificity of the ACR-EULAR criteria in the present study. Actually, in the subanalysis of 383 cases that were examined for UWS, the specificity (76.7%) of the ACR-EULAR criteria was higher than in the whole analysis (72.1%). On the other hand, eight patients with pSS who satisfied only the ACR-EULAR criteria but none of other three sets of criteria might cause the high sensitivity of the ACR-EULAR criteria. These eight patients with pSS had positive FS (5/8 cases) or positive anti-SS-A/Ro (3/8 cases), accompanied with decreased salivary (7/8 cases) or lacrimal (1/8 cases) volume. Collectively, these 19 cases (8 patients with pSS and 11 non-pSS cases) that satisfied only the ACR-EULAR criteria but none of other three sets of criteria seem to lead to the disagreement between the ACR-EULAR criteria and other three sets of criteria. All of these 19 cases had positive FS (11/19 cases) or positive anti-SS-A/Ro (8/19 cases) accompanied by decreased salivary (18/19 cases) or lacrimal (1/19 cases) volume. Thus these 19 patients, even including clinically judged 11 non-pSS cases, seem to have a high probability of suffering from 'true pSS'. This means that the ACR-EULAR criteria might allow to correctly select patients with pSS who are misclassified by other criteria sets. Importantly, the highest sensitivity of the ACR-EULAR criteria might offer some advantages considering that more patients with true pSS could be selected for clinical and therapeutic trials. However, on the other hand, the lowest

(κ coefficient: 0.91) and ACR criteria (κ coefficient: 0.82). These disagreements seem to be somewhat related to our 8 patients with pSS and 11 non-pSS cases who did not satisfy any of the JPN, AECG and ACR criteria, while they satisfied only the ACR-EULAR criteria. Among these 19 cases, 11 non-pSS

Figure 1 Satisfaction for each item adopted in the American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) classification criteria for pSS (ACR-EULAR) among 19 cases that satisfied only the ACR-EULAR criteria. (A) Eight patients with primary Sjögren's syndrome (pSS) who only satisfied the ACR-EULAR criteria had positive Focus Score (FS) (62.5%, 5/8 cases) or pSS-A/Ro (37.5%, 3/8 cases), together with decreased salivary (87.5%, 7/8 cases) or lacrimal (12.5%, 1/8 cases) secretion, resulting in Total Score 4. (B) Eleven non-pSS cases who only satisfied the ACR-EULAR criteria had positive FS (54.5%, 6/11 cases) or positive SS-A/Ro (45.5%, 5/11 cases), together with decreased salivary secretion (100%, 11/11 cases), resulting in Total Score 4. White boxes: non-satisfaction, shadowed boxes: satisfaction (dark shadow: 3 points, light shadow: 1 point). FS-LSG, FS $\geq$1 foci/4 mm^2 in labial salivary gland; SS-A/Ro, positive anti-SS-A/Ro antibody; Ocular staining, Van Bijsterveld Score $\geq$4 in Rose Bengal test, lissamine green test or fluorescein staining test; Schirmer, Schirmer's test $\leq$5 mm/5 min; Saliva, unstimulated whole saliva (UWS) $\leq$0.1 mL/min, gum test $\leq$10 mL/10 min or Saxon test $\leq$2 g/2 min.

specificity of the ACR-EULAR criteria for Japanese patients with pSS was confirmed in both whole analysis and subanalysis using 383 cases that were examined for UWS, whereas the specificity was higher (76.7%) in subanalysis than in whole analysis (72.1%). A high specificity is the most critical aspect because this prevents subjects without pSS from entering clinical studies or trials. Therefore, if we apply the ACR-EULAR criteria to Japanese clinical studies targeted on pSS, we have to pay enough attention to this low specificity.

The 2016 ACR-EULAR criteria focused on pSS only but not on secondary SS, which is associated with other well defined CTDs.[6 7] However, the International SS Criteria Working Group indicated that these criteria could also be applicable for secondary SS, and recommended further studies for secondary SS to confirm this.[6 7] Thus, we also need to compare the performance of these four sets of criteria, including the ACR-EULAR, in the diagnosis of secondary SS because we targeted only patients with pSS or suspected pSS who were free of other CTDs.

The present study has certain limitations. First, we applied the diagnosis by the physician in charge as the 'gold standard'. However, it is better to decide disease case or non-case status by

expert clinical judgement based on clinical vignettes for the 'gold standard' diagnosis, which has been employed by newly developed classification criteria, such as the 2016 ACR-EULAR criteria for pSS[6 7] and the proposed new classification criteria for systemic lupus erythematosus (SLE) by the Systemic Lupus International Collaborating Clinics.[8] Moreover, since this study was entirely carried out in Japan, it might be predicted that the physicians who commonly used JPN criteria could be strongly influenced in their mind by this conceptual habit when they clinically defined cases as having pSS or not. Second, the methods adopted for ocular staining and salivary tests varied among the participating institutions, forcing us to modify certain items in some criteria as described in the Patients and methods section. For these reasons, we need a more sophisticated validation cohort study, using expert clinical judgement based on clinical vignettes as the 'gold standard' and integrated methodology for ocular staining and salivary measurement that fits completely with the items adopted in the criteria.

In addition to the classification criteria for clinical studies, we also need diagnostic criteria for daily clinical practice. The JPN criteria were established for the diagnosis of SS including both pSS and secondary SS, and are always used in daily clinical

practice in Japan. Moreover, we reported previously that the JPN criteria had the highest sensitivity (79.6%) and specificity (90.4%) for all SS, including both primary and secondary SS, among the JPN, AECG and ACR criteria in Japan.[5] However, the JPN criteria adopt sialography, scintigraphy, and gum and Saxon tests for assessment of salivary volume, which are neither adopted in the 2016 ACR-EULAR nor ACR criteria.[3 4 6 7] Furthermore, the JPN criteria use a different cut-off value for ocular staining (Van Bijsterveld Score ≥3) from those used by the ACR-EULAR (Van Bijsterveld Score ≥4 or OSS ≥5), AECG (Van Bijsterveld Score ≥4) and ACR criteria (OSS ≥3).[2–4 6 7] Thus, we need to examine the performance of the JPN criteria using alternative items that are adopted in the ACR-EULAR criteria, such as UWS for gum and Saxon tests, and Van Bijsterveld Score ≥4 or OSS ≥5 for Van Bijsterveld Score ≥3 by a prospective validation cohort study in the near future.

In conclusion, although this study has certain limitations, the results showed that the ACR-EULAR criteria have higher sensitivity and lower specificity in the diagnosis of pSS, compared with the JPN, AECG and ACR criteria. Furthermore, the degree of agreement of the ACR-EULAR criteria with the other three sets of criteria for the diagnosis and denial of pSS was low.

Author affiliations
[1]Department of Internal Medicine, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan
[2]The Research Team for Autoimmune Diseases, The Research Program for Intractable Disease of the Ministry of Health, Labor and Welfare (MHLW)
[3]Department of Data Science, The Institute of Statistical Mathematics, Tokyo, Japan
[4]Department of Hematology and Immunology, Kanazawa Medical University, Kanazawa, Japan
[5]Department of Rheumatology and Clinical Immunology, Kyoto University Graduate School of Medicine, Kyoto, Japan
[6]Unit of Translational Medicine, Department of Immunology and Rheumatology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan
[7]Division of Rheumatology, Department of Internal Medicine, Hyogo College of Medicine, Nishimoniya-city, Japan
[8]Department of Ophthalmology, School of Medicine, Keio University, Tokyo, Japan
[9]Department of Ophthalmology, School of Medicine, Tokyo Women's Medical University
[10]Department of Pathology, Tsurumi University School of Dental Medicine, Tsurumi, Japan
[11]Department of Pharmacology, Nihon Pharmaceutical University
[12]Section of Oral and Maxillofacial Oncology, Division of Maxillofacial Diagnostic and Surgical Sciences, Faculty of Dental Science, Kyushu University, Fukuoka, Japan
[13]Division of Rheumatology, Department of Internal Medicine, School of Medicine, Keio University, Tokyo, Japan
[14]The First Department of Internal Medicine, School of Medicine, University of Occupational and Environmental Health, Kitakyushu, Japan

REFERENCES

1 Tsuboi H, Asashima H, Takai C, et al. Primary and secondary surveys on epidemiology of Sjögren's syndrome in Japan. *Mod Rheumatol* 2014;24:464–70.
2 Vitali C, Bombardieri S, Jonsson R, et al. Classification criteria for Sjögren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. *Ann Rheum Dis* 2002;61:554–8.
3 Fujibayashi T, Sugai S, Miyasaka N, et al. Revised Japanese criteria for Sjögren's syndrome (1999): availability and validity. *Mod Rheumatol* 2004;14:425–34.
4 Shiboski SC, Shiboski CH, Criswell L, et al. American College of Rheumatology classification criteria for Sjögren's syndrome: a data-driven, expert consensus approach in the Sjögren's International Collaborative Clinical Alliance cohort. *Arthritis Care Res* 2012;64:475–87.
5 Tsuboi H, Hagiwara S, Asashima H, et al. Validation of different sets of criteria for the diagnosis of Sjögren's syndrome in Japanese patients. *Mod Rheumatol* 2013;23:219–25.
6 Shiboski CH, Shiboski SC, Seror R, et al. 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome: a consensus and data-driven methodology involving three international patient cohorts. *Ann Rheum Dis* 2017;76:9–16.
7 Shiboski CH, Shiboski SC, Seror R, et al. 2016 American College of Rheumatology/European League Against Rheumatism Classification Criteria for Primary Sjögren's syndrome: a consensus and data-driven methodology involving three international patient cohorts. *Arthritis Rheumatol* 2017;69:35–45.
8 Petri M, Orbai AM, 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.

ARD

Comparison of performance of the 2016 ACR-EULAR classification criteria for primary Sjögren's syndrome with other sets of criteria in Japanese patients

Hiroto Tsuboi, Shinya Hagiwara, Hiromitsu Asashima, Hiroyuki
Takahashi, Tomoya Hirota, Hisashi Noma, Hisanori Umehara, Atsushi
Kawakami, Hideki Nakamura, Hajime Sano, Kazuo Tsubota, Yoko
Ogawa, Etsuko Takamura, Ichiro Saito, Hiroko Inoue, Seiji Nakamura,
Masafumi Moriyama, Tsutomu Takeuchi, Yoshiya Tanaka, Shintaro
Hirata, Tsuneyo Mimori, Isao Matsumoto and Takayuki Sumida

Updated information and services can be found at:

http://ard.bmj.com/content/76/12/1980

These include:

References	This article cites 8 articles, 2 of which you can access for free at: **http://ard.bmj.com/content/76/12/1980**#BIBL
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 and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.
Topic Collections	Articles on similar topics can be found in the following collections Editor's choice (164) Open access (652)

Notes

To request permissions go to:

http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:

http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:

http://group.bmj.com/subscribe/