Assessment of major salivary gland ultrasonography in Sjögren's syndrome. A comparison between bedside and post-examination evaluations

D.S. Hammenfors[1,2], H. Causevic[1,2], J. Assmus[3], J.G. Brun[2,4], R. Jonsson[1,2], M.V. Jonsson[1,5]

[1]*Broegelmann Research Laboratory, Department of Clinical Science, University of Bergen;* [2]*Department of Rheumatology, Haukeland University Hospital;* [3]*Centre for Clinical Research, Haukeland University Hospital;* [4]*Department of Clinical Science, Section for Rheumatology, University of Bergen;* [5]*Department of Clinical Dentistry, Section for Oral and Maxillofacial Radiology, University of Bergen, Norway.*

Daniel S. Hammenfors, MD, PhD
Haris Causevic, MD
Jörg Assmus, PhD
Johan G. Brun, MD, PhD
Roland Jonsson, DMD, PhD
Malin V. Jonsson, DMD, PhD

Please address correspondence to:
Dr Daniel S. Hammenfors,
Department of Rheumatology,
Haukeland University Hospital,
University of Bergen,
Jonas Lies vei 83,
5021 Bergen, Norway.
E-mail:
sten.daniel.hammenfors@helse-bergen.no

Received on April 24, 2019; accepted in revised form on July 12, 2019.

Clin Exp Rheumatol 2019; 37 (Suppl. 118): S153-S158.

Key words: Sjögren's syndrome, ultrasonography, salivary gland, correlation

Funding: this research was funded by Western Norway Regional Health Authorities (Helse Vest) the University of Bergen) and EU-Horizon 2020 contract: SC1-PM-04-2016 HarmonicSS no. 731944.

Competing interests: none declared.

ABSTRACT

Objective. *Major salivary gland ultrasonography (SGUS) is a suitable diagnostic tool in Sjögren's syndrome (SS). We aimed to determine the more representative gland, projection and format most applicable for reproducible image analysis.*

Methods. *One investigator performed SGUS in patients with SS. Parotid and submandibular glands were examined in longitudinal and transverse planes and evaluated bedside using a simplified scoring system (0-3). Longitudinal and transverse images and videos of all glands were stored and later evaluated/graded by three investigators, at two time-points. Agreement was calculated using intraclass correlation coefficient (ICC).*

Results. *The ICC for static image and video scoring compared to bedside evaluation ranged from 0.131 to 0.882. Average ICC for longitudinal/transverse image was 0.667/0.662, and 0.683/0.510 for longitudinal/transverse video. Interobserver reliability was good to excellent (0.81–0.94). Intraobserver reliability scores ranged from fair to excellent (0.46–0.96). The correlation between image and video evaluations of all modalities and examiners was good to excellent (0.614-0.904). The best mean ICC was found for the longitudinal projection of the left parotid gland (0.861) and the lowest mean ICC was for the transverse projection of the left submandibular gland (0.66).*

Conclusion. *Our study indicates a trend favouring longitudinal video of the parotid gland as preferred projection, gland and storage format.*

Introduction

Primary Sjögren's syndrome (pSS) is a systemic autoimmune disease mainly affecting exocrine glands leading to oral and ocular dryness (xerostomia and keratoconjunctivitis sicca, respectively). Extra-glandular manifestations such as fatigue, arthralgias and xeroderma are also common. Ultrasonography is a cost-effective and non-invasive method for evaluation of the salivary gland component in patients with Sjögren's syndrome (SS). Major salivary gland ultrasonography (SGUS) has proven to have a high diagnostic value (1-6) and enhance proposed classification criteria (7-9). However, lack of international consensus on standardisation makes integration into clinical practice difficult (10-15). A recent meta-analysis describe several different scoring systems used, partly favouring a simplified scoring system (0-4) with lower operator time for use in the clinic (16). So far, SGUS has not been added to the current classification criteria for pSS. Interestingly, a recent study demonstrated that scoring of parenchymal echogenicity and hypoechogenic areas in one parotid and one submandibular salivary gland highly predicted SS according to the ACR-EULAR classification; even scoring only hypoechogenic areas on one side showed good results (17). A study from the same group showed that validity of the ACR-EULAR criteria remained high after incorporation of SGUS, offering a larger array of tests to evaluate fulfilment of the ACR-EULAR criteria (18). In a study with several ultrasonographers, Jousse-Joulin *et al.* demonstrated that homogeneity and echogenicity for static images had the highest interobserver reliability, whereas on acquisition images, echogenicity had higher interobserver reliability than homogeneity (14). Another study with two blinded investigators showed high degree of

interobserver agreement regarding the final SGUS score, as well as in the assessment of glandular homogeneity, echogenicity and hypoechoic areas, when comparing bedside evaluations only (19).

The aim of this study was to investigate the more representative major salivary gland and projection, and determine the more appropriate storage format for post-examination evaluation of SGUS.

Material and methods

Patient cohort

Consecutive patients with SS (n=32) attending the outpatient clinic of the Department of Rheumatology, Haukeland University Hospital, Bergen, Norway were included in the cohort (May-November 2017). Following oral and written information of the study, all patients signed a consent form. Patients were evaluated using the American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) classification criteria (20) for pSS, including a clinical rheumatological examination, registration of sicca symptoms, unstimulated and stimulated whole salivary flow (UWS and SWS) measurements, minor salivary gland biopsy if necessary, serology and current/previous medical treatment.

Major salivary gland ultrasonography

Upon inclusion, one clinical investigator (B) performed SGUS on all patients using a GE Logiq E9 (GE, USA) with a linear high-frequency transducer (6–15 MHz), 66 dB dynamic range, using 2.5–3.5 cm depth and 3.3–5.5 gain. Both right and left parotid and submandibular glands were examined. All four glands were examined in the longitudinal and transverse planes. All projections were evaluated bedside by the clinical investigator (B), and each gland was given a score (0-3) based on overall impression using a simplified scoring system. Briefly, glandular homogeneity and presence of hypoechogenic areas were evaluated and graded as previously described according to Hocevar *et al*. 2005 (13). The number and size of hypoechogenic and/or anechoic areas as well as the degree of inhomogeneity presented in both the submandibular and the parotid glands formed the basis for an overall score determined by the clinical investigator. In cases where hypo-/anechoic areas were not detected, the score was set to 0. If SGUS showed a few minor focal hypo-/anechoic areas that were considered within normal, the score was set to 1. When the gland was more severely affected with multiple focal hypo-/anechoic areas but some homogenous and normal-appearing salivary gland tissue remained, the score was set to 2. In cases where the whole perimeter of the gland was affected, with minimal remains of normal-appearing glandular tissue, the score was set to 3 (10, 13). Grades 0–1 were considered to correspond to normal/non-specific changes and grades 2–3 were regarded as pathological changes (Fig. 1).

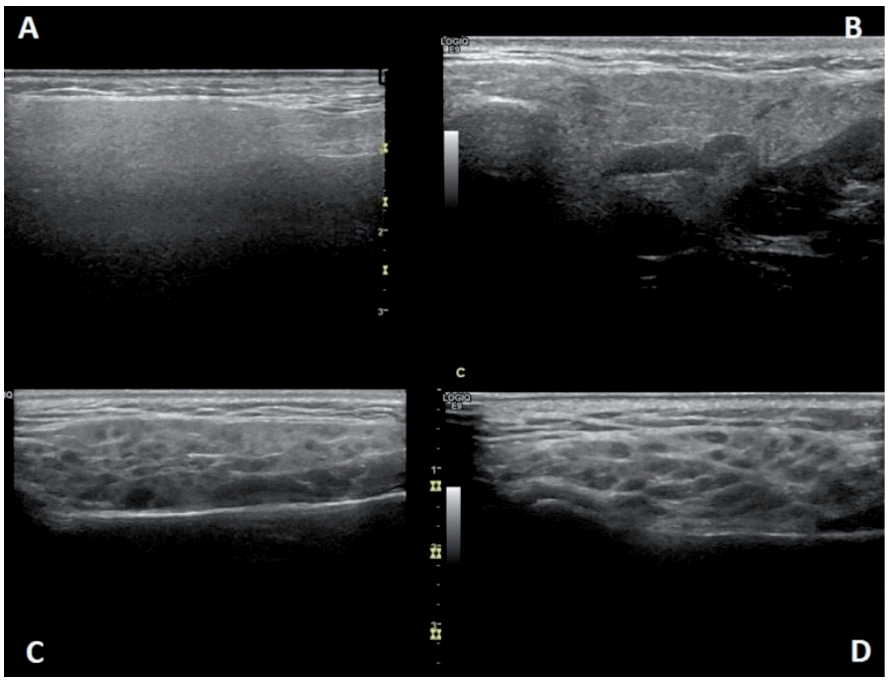

Fig. 1. Ultrasonography of the parotid glands illustrating score 0-3. Score 0 (**A**) and score 1 (**B**) are considered normal-appearing morphology. Score 2 (**C**), with obvious but limited affection of glandular tissue, and score 3 (**D**), with little or no remaining normal-appearing glandular tissue, are considered to correspond with SS-like pathological changes.

Representative videos and static images for both the longitudinal and transverse planes of all four glands were concurrently stored as DICOM files for documentation and later re-evaluation, in total 16 images (eight videos and eight static images) per patient. Using the hospital media storage data-base, videos and images were extracted in de-identified format and randomised using a random number generator (random.org, random integer generator) for evaluation in equal consecutive order by three investigators (A, B, C). Images and videos were converted to .jpg and .mp4 before evaluation.

All videos and images were evaluated and scored by all three investigators at two separate timepoints (round 1 and round 2); November 2017 and January 2018. In a minority of cases two videos and/or images were available from the same gland and projection, the highest/more severe score was then applied. The scores registered in round 1 were used for interclass correlation assessment (bedside compared to image and video analyses). The scores registered in round 2 were used for calculating intraobserver reliability.

Statistical analysis

Descriptive methods were used to characterise the cohort. The inter- and intraobserver reliability was assessed by intraclass correlation coefficient (ICC), two-way mixed, absolute agreement. Gold standard was defined as the score obtained bed-side at inclusion, including both right and left parotid and submandibular glands, in the longitudinal

and transverse planes; each gland given a composite score (0-3) based on overall impression. Blinded to the bedside results, three investigators (A, B, C) scored anonymised images and videos at two different time-points. For inter-observer reliability we compared image and video scoring from round 1 to bedside evaluation, using the mean ICC from all investigator comparisons to determine the best gland, projection and storage modality. The individual ICC were also calculated to determine whether there were differences between the investigators. Comparing results from round 1 and round 2, the intra-observer reliability was calculated for each investigator.

ICC less than 0.40 were considered poor, 0.40–0.59 as fair, 0.60–0.74 as good, and greater than or equal to 0.75 as excellent (21).

All statistical analyses were performed using SPSS 24 software (IBM Corp, Armonk, NY). The graphics were created using GraphPad Prism v. 7 software (GraphPad Software, La Jolla, CA).

Ethical considerations

This study was performed in accord with the Regional Medical and Health Research Ethics regulations, and necessary applications were approved by the regional Committees in Norway: 145/96-44.96, 242.06 2009/686. Informed consent was obtained from all participants. None of the patients withdrew from the study.

Results

The study included 32 patients with SS. The cohort characteristics are summarised in Table I. Eight patients did not fulfill the 2016 ACR-EULAR classification criteria for pSS; four were indeed positive for anti-Ro/SSA autoantibodies, whereas two lacked the minor salivary gland biopsy, and two had a minor salivary gland biopsy but the focus score was less than 1. One patient did have a focus score of greater than or equal to 1 but was anti-Ro/SSA negative. Three patients had not performed a lip biopsy. Among these, one patient was anti-Ro/SSA negative; upon inclusion this patient had reduced saliva secretion and reduced tear flow.

Table I. Patient characteristics.

Characteristics	Total no.	Value
Age (years, mean [min-max])	32	56 [21-83]
Gender (Female)	32	28 (88 %)
UWS ≤1.5 ml/15 min	32	19 (64 %)
SWS ≤3.5 ml/5 min	32	13 (41 %)
Schirmer's I-test ≤5mm/5 min	31	16 (51 %)
Anti-Ro/SSA and/or anti-La/SSB	32	24 (75 %)
Focus score ≥1	20	15 (75%)
ACR[1]-EULAR[2] classification criteria for pSS	32	24 (75 %)
Pathological SGUS (bed-side)	32	14 (44 %)

Age and gender distribution is in line with other reports as well as registered positive autoantibodies. However, we did not register as many abnormal ocular or oral tests, and not as many positive minor salivary gland biopsies (23). This might be explained by the fact that not all our patients fulfill the ACR-EULAR classification criteria for pSS, a common inclusion criterion in studies.
[1]ACR: American College of Rheumatology.
[2]EULAR: European League Against Rheumatism.

The two other patients who were anti-Ro/SSA negative and lacked the minor salivary gland lip biopsy did display reduced tear flow, and one patient also had pathologically reduced saliva secretion.

Major salivary gland ultrasonography examination showed pathological findings (score 2–3) in at least two glands in 14 (44 %) patients (Fig. 2), for the remaining 18 patients imaging findings were considered within normal. Among the patients with pathological changes characteristic for SS, 13/14 patients fulfilled the ACR-EULAR criteria. One of the eight patients not fulfilling ACR-EULAR classification criteria also presented with SGUS changes characteristic for SS; the patient was anti-Ro/SSA positive but lacked the minor salivary gland biopsy. Among the patients with pathological SGUS findings, seven patients had pathological changes in all four glands, whereas one patient had pathological changes in both submandibular glands but only one parotid gland. One patient had changes in both parotid glands but only one submandibular gland. In the remaining patients, two displayed pathological changes only in the parotid glands, whereas three had pathological changes only in the submandibular glands. The composite bedside score served as gold-standard for further comparison to results from round 1 and round 2.

Not all images and videos were considered representative by all investigators and were discarded, in total eleven images and eight videos. Evaluation (round

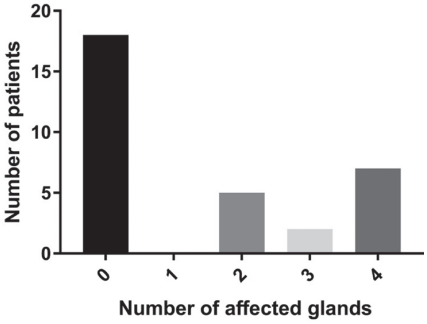

Fig. 2. Number of affected major salivary glands in patients examined bed-side. Major salivary gland ultrasonography examination showed pathological findings (score 2-3) in at least two glands in 14/32 (44 %) patients. Seven patients had pathological changes in all four glands, whereas one patient had pathological changes in both submandibular glands but only one parotid gland. One patient had changes in both parotid and only one submandibular gland. In the remaining patients, two displayed pathological changes only in the parotid glands, whereas three had pathological changes only in the submandibular glands. Pathological changes in only one of the major salivary glands was not observed in this patient material.

1) of SGUS static images and videos revealed pathological changes recognised by at least one of the evaluators in 276/723 (38%) of SGUS static images and in 308/713 (43%) of the SGUS videos. In 412/674 (61%) cases, static image score coincided with video score for the same gland, investigator and projection. Normal-appearing (score 0–1) and SS-like pathological changes (score 2–3) coincided in 583/674 (86%) cases. Re-evaluation (round 2) of the SGUS static images revealed pathological changes in 316/736 (43%), and in 334/710 (47%) of the SGUS videos. In 434/684 (63%) cases static image

score coincided with video score for the same gland, investigator and projection. Non-pathologic (score 0–1) and pathologic (score 2–3) scores coincided in 581/684 (85%) cases.

Taken together, the correlation between image and video evaluation of all projections and examiners was good to excellent, with ICC values varying from 0.614 to 0.904 (mean 0.796). The highest mean was calculated for the longitudinal projection of the left parotid gland (0.861), whereas the lowest mean was calculated for the transverse projection of the left submandibular gland (0.660). For longitudinal projections of the parotid and the submandibular glands the mean ICC was 0.853 (parotid gland ICC 0.854 and submandibular gland ICC 0.852). For the transverse plane, the mean ICC was 0.739 (parotid gland ICC 0.725 and submandibular gland 0.753).

Comparing interobserver agreement the ICC between investigators was overall excellent with scores ranging from 0.81 to 0.93 in round 1 and 0.84 to 0.94 in round 2. The intraobserver reliability scores had a wider range of 0.46 to 0.96, depending on the gland (parotid/submandibular), projection (longitudinal/transverse) and format (image/video). In general, the scores were good with a slightly higher intraobserver reliability score mean (investigator A, B and C) for images (0.81) compared to video (0.74) and longitudinal (0.80) compared to transverse (0.76) projection. The difference between the parotid and submandibular glands was negligible (0.78 and 0.78).

With regard to our gold standard, results from round 1 compared to bedside evaluation, the ICC ranged from 0.131 to 0.882 (Fig. 3), with an average ICC score 0.577–0.674 for the three investigators (A: 0.674, B: 0.577, and C: 0.641). The highest average score between bedside assessment and image/video re-evaluation was observed for longitudinal video of the right parotid gland (0.772). The lowest correlation for all investigators between bedside scoring and image/video re-evaluation was seen for transverse video of the left submandibular gland (0.209), compared to the correlation for the transverse image of the left submandibular gland (0.564) and transverse video of the left parotid gland (0.637) and transverse image of the left parotid gland (0.718). The ICC average for longitudinal video was 0.683, longitudinal image 0.667, transverse image 0.662, and transverse video 0.510.

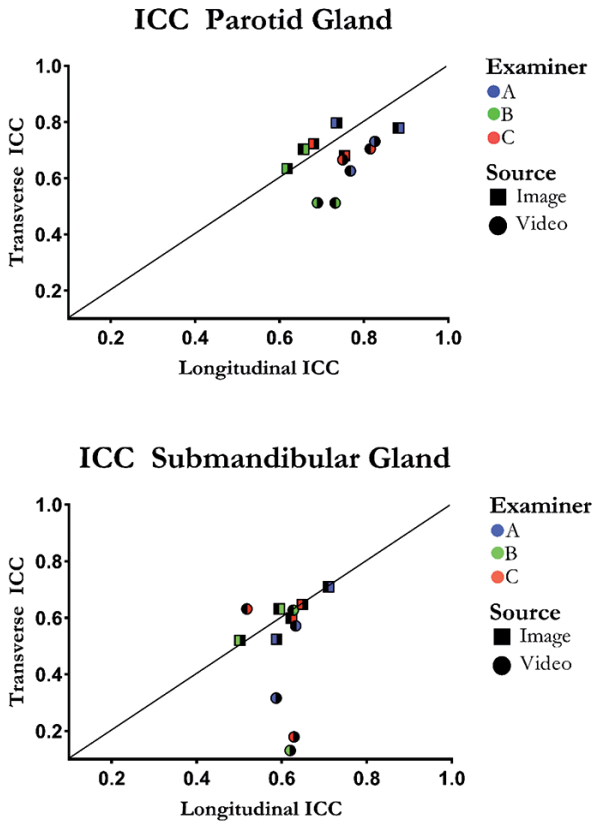

Fig. 3. ICC for the parotid and submandibular gland for the three investigators (A, B and C) with video (O) and image (□) of both glands. Both right side (right side of symbol coloured) and left side (left side of symbol coloured) glands are illustrated. Our results indicate a tendency favouring longitudinal videos of the parotid glands, but this finding was not significant.
Transverse video of the left submandibular gland had the lowest ICC.

Discussion

Ultrasonography is a cost-effective and non-invasive method for evaluation of the salivary gland component in patients with SS. Major salivary gland ultrasonography has proven to have a high diagnostic value and enhance proposed classification criteria (7). To our knowledge, there are no studies comparing bedside SGUS evaluation to post-examination scoring of longitudinal and transverse digital stored images and videos of both parotid and submandibular glands.

The aim of this study was to determine the more representative major salivary gland and projection for clinical use, and to find the most appropriate storage format for post-examination evaluation of SGUS. We evaluated the inter- and intraobserver reliability of three investigators scoring static images and videos of the parotid and submandibular salivary glands. Bedside evaluation performed by one of the investigators served as gold standard.

Major salivary gland ultrasonography examination showed pathological findings in at least two glands in 14 (44%) patients. Similar numbers (38–43%) were obtained when combining the results from the three investigators. Interestingly, static image findings coincided with video in 412/674 (61%) cases, for the same gland, investigator and projection. Normal-appearing morphology (score 0–1) and SS-like pathological changes (score 2–3) coincided in 583/674 (86%) cases. In the re-evaluation (round 2) pathological changes were observed in 43-47%, and in 434/684 (63%) cases static image score coincided with video score for the same gland, investigator and projection. Normal-appearing morphology (score 0–1) and SS-like pathological changes (score 2–3) coincided in 581/684 (85%) cases.

In line with our findings, earlier studies have shown good inter- and intraobserver reliability for SGUS, reviewed in (4). Interobserver reliability was overall excellent with scores ranging

from 0.81 to 0.93 in round 1 and 0.84 to 0.94 in round 2, indicating consistent image/video evaluation and scoring. The intraobserver reliability was lower and ranged from 0.46 to 0.96, but overall in accordance with an acquisition study by Jousse-Joulin *et al*. (14). Damjanov *et al*. (15) showed that interobserver reliability was higher among experienced investigators compared to those less experienced. Similarly, lower ICC between static images, videos and bedside evaluation were calculated for the least experienced investigator in our study. The same investigator had indeed performed the bedside SGUS examination, interpretation, and image acquisition, raising an interesting question of whether experience in interpreting SGUS images post-examination is of greater importance for the ICC, than actually performing both the bed-side acquisition and scoring, and subsequent post-examination analysis. This could in part explain the lower intraobserver reliability as compared to interobserver reliability.

In cases where there is a discrepancy in experience, and thus reliability, larger variation in the evaluations may be expected. In a study such as the current, intra- and interobserver variation is calculated to illustrate this variation. However, in daily clinical practice, differences may not be so obvious. Hence, regular calibration exercises among ultrasonographers should be recommended. Comparing three raters, the scores for interobserver reliability were less influenced by the one investigator with lower reliability and the ICC scores were indeed better. This raises the question of how much variation in SGUS image evaluation is clinically significant, when findings of normal-appearing morphology and SS-like pathological findings coincided in 85-86% of the cases.

In a recent study by Jousse-Joulin *et al*. (22) consensual "OMERACT" definitions were developed. In short, it was agreed that the parotid and submandibular glands should be evaluated and scored based upon greyscale lesions and anechoic/hypoechoic foci in a semi-quantitative matter. The simplified scoring system applied to still-images and videos in our study much resembles the OMERACT definitions applied to video clips of the major salivary glands, with high inter- and intrareader reliability (22).

To conclude, when performing SGUS in a clinical setting, storing images or videos is of high value to enable later image evaluation with regard to diagnostics, second opinion and disease progression. It is also possible that future treatment effects will be evaluated by progression/reversal of SGUS changes. The findings in our study indicate a trend favouring longitudinal video of the parotid gland as preferred projection, gland and storage format. It also supports the notion that experience and training is an important factor when evaluating post-examination videos and images.

Possible limitations include a limited number of patients. The bedside evaluation was performed by one investigator, enabling standardisation of image acquisition and selection of videos and images, but possibly limiting objective evaluation. Future studies could benefit from a higher number of patients, and acquisition and bed-side evaluation of each patient performed by several investigators, using the full Hocevar *et al*. scoring system (19) and by regression analysis determining the essential items when comparing bed-side scoring to post examination evaluations.

References

1. WERNICKE D, HESS H, GROMNICA-IHLE E, KRAUSE A, SCHMIDT WA: Ultrasonography of salivary glands -- a highly specific imaging procedure for diagnosis of Sjögren's syndrome. *J Rheumatol* 2008; 35: 285-93.
2. MILIC V, PETROVIC R, BORICIC I *et al*.: Ultrasonography of major salivary glands could be an alternative tool to sialoscintigraphy in the American-European classification criteria for primary Sjögren's syndrome. *Rheumatology* 2012; 51: 1081-5.
3. LUCIANO N, BALDINI C, TARANTINI G *et al*.: Ultrasonography of major salivary glands: a highly specific tool for distinguishing primary Sjögren's syndrome from undifferentiated connective tissue diseases. *Rheumatology* 2015; 54: 2198-204.
4. JOUSSE-JOULIN S, MILIC V, JONSSON MV *et al*.: Is salivary gland ultrasonography a useful tool in Sjögren's syndrome? A systematic review. *Rheumatology* 2016; 55: 789-800.
5. LUCIANO N, FERRO F, BOMBARDIERI S, BALDINI C: Advances in salivary gland ultrasonography in primary Sjögren's syndrome. *Clin Exp Rheumatol* 2018; 36 (Suppl. 114): S159-64.
6. BALDINI C, ZABOTTI A, FILIPOVIC N *et al*.: Imaging in primary Sjögren's syndrome: the 'obsolete and the new'. *Clin Exp Rheumatol* 2018; 36 (Suppl. 112): S215-21.
7. TAKAGI Y, NAKAMURA H, SUMI M *et al*.: Combined classification system based on ACR/EULAR and ultrasonographic scores for improving the diagnosis of Sjögren's syndrome. *PloS One* 2018; 13: e0195113.
8. LE GOFF M, CORNEC D, JOUSSE-JOULIN S *et al*.: Comparison of 2002 AECG and 2016 ACR/EULAR classification criteria and added value of salivary gland ultrasonography in a patient cohort with suspected primary Sjögren's syndrome. *Arthritis Res Ther* 2017; 19: 269.
9. JONSSON MV, BALDINI C: Major salivary gland ultrasonography in the diagnosis of Sjögren's syndrome: a place in the diagnostic criteria? *Rheum Dis Clin North Am* 2016; 42: 501-17.
10. HAMMENFORS DS, BRUN JG, JONSSON R, JONSSON MV: Diagnostic utility of major salivary gland ultrasonography in primary Sjögren's syndrome. *Clin Exp Rheumatol* 2015; 33: 56-62.
11. MILIC VD, PETROVIC RR, BORICIC IV *et al*.: Diagnostic value of salivary gland ultrasonographic scoring system in primary Sjögren's syndrome: a comparison with scintigraphy and biopsy. *J Rheumatol* 2009; 36: 1495-500.
12. QI X, SUN C, TIAN Y *et al*.: Comparison of the diagnostic value of four scoring systems in primary sjögren's syndrome patients. *Immunol Lett* 2017; 188: 9-12.
13. HOCEVAR A, AMBROZIC A, ROZMAN B, KVEDER T, TOMSIC M: Ultrasonographic changes of major salivary glands in primary Sjögren's syndrome. Diagnostic value of a novel scoring system. *Rheumatology* (Oxford) 2005; 44: 768-72.
14. JOUSSE-JOULIN S, NOWAK E, CORNEC D *et al*.: Salivary gland ultrasound abnormalities in primary Sjögren's syndrome: consensual US-SG core items definition and reliability. *RMD Open* 2017; 3: e000364.
15. DAMJANOV N, MILIC V, NIETO-GONZALEZ JC *et al*.: Multiobserver reliability of ultrasound assessment of salivary glands in patients with established primary Sjögren syndrome. *J Rheumatol* 2016; 43: 1858-63.
16. ZHOU M, SONG S, WU S *et al*.: Diagnostic accuracy of salivary gland ultrasonography with different scoring systems in Sjögren's syndrome: a systematic review and meta-analysis. *Sci Rep* 2018; 8: 17128.
17. MOSSEL E, ARENDS S, VAN NIMWEGEN JF *et al*.: Scoring hypoechogenic areas in one parotid and one submandibular gland increases feasibility of ultrasound in primary Sjögren's syndrome. *Ann Rheum Dis* 2018; 77: 556-62.
18. VAN NIMWEGEN JF, MOSSEL E, DELLI K *et al*.: Incorporation of salivary gland ultrasonography into the ACR-EULAR criteria for primary Sjögren's syndrome. *Arthritis Care Res* (Hoboken). 2019 Jun 29 [Epub ahead of print].
19. HOCEVAR A, RAINER S, ROZMAN B, ZOR P, TOMSIC M: Ultrasonographic changes of major salivary glands in primary Sjögren's syn-

drome. Evaluation of a novel scoring system. *Eur J Radiol* 2007; 63: 379-83.

20. 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.
21. CICCHETTI DV: Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment* 1994; 6: 284-90.
22. JOUSSE-JOULIN S, D'AGOSTINO MA, NICOLAS C *et al*.: Video clip assessment of a salivary gland ultrasound scoring system in Sjögren's syndrome using consensual definitions: an OMERACT ultrasound working group reliability exercise. *Ann Rheum Dis* 2019; 78: 967-73.
23. BRITO-ZERON P, ACAR-DENIZLI N, ZEHER M *et al*.: Influence of geolocation and ethnicity on the phenotypic expression of primary Sjögren's syndrome at diagnosis in 8310 patients: a cross-sectional study from the Big Data Sjögren Project Consortium. *Ann Rheum Dis* 2017; 76: 1042-50.