



Available online at
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com/en



Original article

What is the best salivary gland ultrasonography scoring methods for the diagnosis of primary or secondary Sjögren's syndromes?

Amélie Martel^{a,b}, Guillaume Coiffier^{c,d,*}, Aurore Bleuzen^{b,e}, Jean Goasguen^f, Michel de Bandt^g, Christophe Deligny^h, Julie Magnant^{a,b}, Nicole Ferreira^{a,b}, Elisabeth Diot^{a,b}, Aleth Perdriger^{c,d}, François Maillot^{a,b}

^a Internal Medicine Department, 37000 Tours, France

^b François-Rabelais University of Tours, 37000 Tours, France

^c Rheumatology Department, 35203 Rennes, France

^d Inserm, UMR 1241, University of Rennes 1, 35203 Rennes, France

^e Medical Imaging Department, 37000 Tours, France

^f Medical Imaging Department, 97200 Fort de France, Martinique, France

^g Rheumatology Department, 97200 Fort-de-France, Martinique, France

^h Internal medicine department, 97200 Fort-de-France, Martinique



suppl.
 Informations

ARTICLE INFO

Article history:

Accepted 25 June 2018

Available online 24 July 2018

Keywords:

Sjögren's syndrome
 Salivary gland
 Ultrasound

ABSTRACT

Objective: To evaluate the performance of salivary gland ultrasonography for the diagnosis of primary and secondary Sjögren's syndromes (pSS and sSS).

Method: Multicenter cross-sectional study on 97 patients with clinical sicca symptoms. The pSS ($n = 22$) met the American-European Consensus Group (AECG) classification criteria. The control patients ($n = 36$) with sicca symptoms did not fulfill the AECG criteria. Four scores were used to evaluate the 4 major salivary gland echostructure: the Salaffi score (0–16), Jousse-Joulin score (0–4), Hocevar score (0–48) and Milic score (0–12).

Results: The medians of ultrasonographic (US) scores were higher in the pSS and sSS groups than in the control group ($P < 0.001$). The receiver-operating characteristic (ROC) curves and the positive likelihood ratio (LR+) of the four scores showed a good diagnostic performance for the US diagnosis of pSS and sSS. Respectively, for pSS and sSS, the AUC were 0.891 (95%CI 0.812–0.970) and 0.824 (95%CI 0.695–0.954) for Hocevar score, 0.885 (95%CI 0.804–0.965) and 0.808 (95%CI 0.673–0.943) for Milic score, 0.915 (95%CI 0.848–0.982) and 0.844 (95%CI 0.724–0.965) for Salaffi score, 0.897 (95%CI 0.821–0.973) and 0.851 (95%CI 0.735–0.968) for Jousse-Joulin score. This study showed an interesting inter-observer reproducibility ($\kappa = 0.714 \pm 0.131$) of the US evaluation with 85.7% agreement between reader to determine the pathological character of the salivary glands.

Conclusion: Salivary gland US is a simple, non-invasive and performant imaging procedure for the diagnosis of pSS and sSS, with Salaffi, Milic and Jousse-Joulin scores.

© 2018 Published by Elsevier Masson SAS on behalf of Société française de rhumatologie.

1. Introduction

The primary Sjögren's Syndrome (pSS) is a chronic autoimmune disease. It is characterized by a lymphocytic infiltration and destruction of the salivary and lacrimal glands, which causes dryness of the eyes (xerophthalmia) and mouth (xerostomia) by loss of glandular functions. Asthenia, inflammatory arthralgia often com-

plete this clinical background. Almost a third of patients have systemic involvement with pulmonary, renal, neurologic, rheumatologic, vascular or lymphatic disorders. The pSS is the systemic disease for which the incidence of non-Hodgkin's Lymphoma (NHL) is the highest [1,2]. Biologically, anti-SSA/SSB antibodies are the most specific markers, but are present in only 50% to 75% of pSS diagnostics. The anti-nuclear antibodies present in 80% of pSS and the rheumatoid factors in 40% to 75% of pSS, are non-specific and can be found in other systemic autoimmune diseases [3–5]. The Sjögren's Syndrome is considered secondary (sSS) when associated with other connective tissue disease.

Like other systemic diseases, there is no single and specific diagnostic test. The diagnosis of the Sjögren's syndrome is based

* Corresponding author: Rheumatology Department, Hôpital Sud, 16 boulevard de Bulgarie, BP90347, 35203 Rennes Cedex 2, France

E-mail address: guillaume.coiffier@chu-rennes.fr (G. Coiffier).

<https://doi.org/10.1016/j.jbspin.2018.06.014>

1297-319X/© 2018 Published by Elsevier Masson SAS on behalf of Société française de rhumatologie.

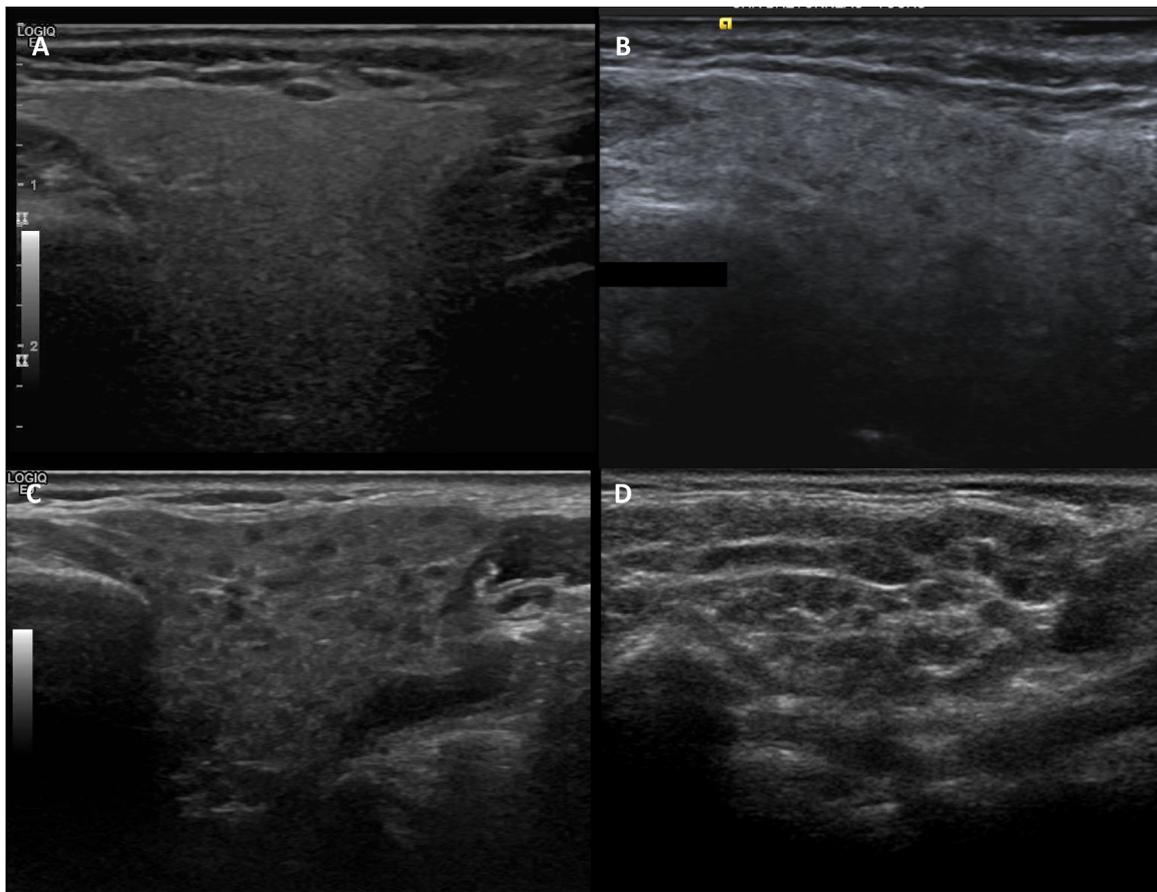


Fig. 1. Examples of Parotid Gland ultrasound in short axis. A. Normal parenchyma (echogenicity comparable to the thyroid) with no inhomogeneity no hypoechoic areas, no hyperechoic band reflexions and regular defined borders. B. Echogenicity was comparable to the thyroid with mild inhomogeneity, a few hypoechoic areas (< 2 mm), a few hyperechoic band reflexions and partly defined borders. (This gland was scored Hocevar = 4, Salaffi = 2, Milic = 1, and Jousse-Joulin = 2). C. Echogenicity was comparable to the thyroid with evident inhomogeneity, several hypoechoic areas (2–6 mm), a several hyperechoic band reflexions and partly defined borders. (This gland was scored Hocevar = 7, Salaffi = 3, Milic = 2, and Jousse-Joulin = 3). D. Echogenicity was decreased to the thyroid parenchyma with grossly inhomogeneity, numerous hypoechoic areas (2–6 mm), numerous hyperechoic band reflexions and ill-defined borders. (This gland was scored Hocevar = 12, Salaffi = 4, Milic = 3, and Jousse-Joulin = 3).

on a range of clinical, biological, morphological and histological symptoms. The most widely used classification criteria were proposed by the American-European Consensus Group (AECG) in 2002 [6]. They are used to define and standardize the diagnosis of primary and secondary SS. Among these criteria, parotid sialography and salivary gland scintigraphy are morphological examinations that are now rarely used in clinical practice, because of the radiating character of sialography and the low specificity of scintigraphy [7]. The ultrasonography (US) appears to be an attractive imaging approach by its non-invasive, inexpensive, non-irradiating and accessible technique. It allows also morphologic evaluation. Major salivary gland US (2 parotid and 2 submandibular glands) has been developed over the past 20 years. Several studies have shown that it provides similar information to sialography while being considerably less invasive [7,8]. MRI showed a high diagnostic performance [9–11] with good correlation between parotid MRI and salivary gland US (kappa agreement = 0.87) [9]. However, its lack of availability and its high cost does not allow its use for diagnostic purpose, but rather for the exploration of a suspicious glandular lesion.

With US B mode, the most specific pathogenic characteristic in SS is the inhomogeneity of the glandular parenchyma. Since the first proposal of US score made in 1992 by De Vita et al. [12], several US scores have been proposed with different items: gland size, contour regularity, clearness of the posterior glandular border, presence of hypoechoic zones, echogenic bands and size of the cysts. The lack of consensual score for SS diagnosis is therefore one of the first limits for the use of salivary gland US in clinical practice. In

another hand, few studies have evaluated this imaging method for sSS diagnosis. Finally, the main disadvantage of US is its operator-dependent character, with currently little data on its inter- and intra-observer reproducibility.

Our objective was to compare the performance of four scoring methods for the diagnosis of pSS and sSS, and to evaluate inter and intra-observer reproducibility of salivary gland US.

2. Methods

This is a multicentre, cross-sectional observational study carried out in three French university hospitals (Rennes, Tours and Fort-de-France).

2.1. Patients

All included patients complained of ocular and/or oral dry syndrome. Some already had a diagnosis of pSS or sSS established according to the AECG classification criteria. Others were explored for a suspicion of SS. Patients with sicca symptoms who did not fulfill the AECG criteria were included in the control group. The inclusions have been taking place between January 2014 and November 2016, in the departments of rheumatology in Rennes, in internal medicine in Tours and Fort-de-France. All patients were informed of the objectives and procedure of the study and gave their consent. The study was approved by the ethics committee of Rennes university hospital.

Exclusion criteria were: surgery for the removal of major salivary glands (≥ 2 glands / 4), graft versus host disease, history of sarcoidosis, cervical-facial radiotherapy, hepatitis C, HTLV-1 or HIV infection.

The use of drugs causing sicca symptoms was not considered as an exclusion factor because it does not interfere with the results of salivary gland US.

2.2. Clinical, histopathological evaluations and laboratory investigations

A standardized questionnaire was used to evaluate the severity of the subjective sicca syndrome, rated from 1 to 6 [Appendix A, Supplementary data S1; See the supplementary material associated with this article online]. Age, sex, duration of development of the sicca symptoms (<5 , between 5 and 10, >10 years), medications providing a sicca syndrome and active smoking were notified. Schirmer's test (abnormal if ≤ 5 mm/5mn), salivary flow (abnormal if ≤ 0.5 g/5 min) and minor labial salivary gland biopsy (abnormal if grade 3 or 4 according to Chisholm and Mason or focus score ≥ 1 [13]) were performed unless the patient has already had these examinations. In this case, the results were taken from the medical records.

Anti-SSA/SSB antibodies, anti-nuclear antibodies (ANA) and rheumatoid factor (RF) levels, blood protein electrophoresis assays were investigated in the medical records or performed during sicca syndrome exploration.

2.3. Ultrasonography of the major salivary glands

US examination of the parotid and submandibular salivary glands was performed on every included patient to evaluate the four scoring US methods, with a different US scanner in each center (Philips, General Electrics and Esaote). Images of the four salivary glands were obtained in transverse and longitudinal planes with a 5–18 MHz linear probe. Each of the three examiners was blinded of clinical symptoms and diagnosis: a rheumatologist in Rennes (GC), a radiologist in Tours (AB) and Fort-de-France (JG) with respectively five, fifteen and twenty years of practice of salivary gland US. Four US scores were chosen in this study for their diagnostic performance described in previous studies: Salaffi (0–16) [14], Jousse-Joulin (0–4) [15] [48] [16] and Milic (0–12) [17] (Appendix A, Supplementary data S-12) [17] (Fig. 1 and Appendix 1, Supplementary Data S1). The four scores propose different descriptions to evaluate the glandular parenchyma with several items (hypoechoic areas, echogenic bands, cysts, clearness of glandular borders). The Milic score is a simplified method with only one of these items: the parenchymal inhomogeneity. The Salaffi score evaluates all these items but is less accurate than Hocevar score. Jousse-Joulin score assesses the size of cysts and echogenic bands. All the included patients benefited from a centralized reading of the US images to establish the 4 scores (GC).

The Rennes (GC) and Tours (AB) examiners performed a second reading of their US images, few months after the first reading, to evaluate the intra-observer reproducibility of the Salaffi score. The Rennes examiner (GC) read again all the images from Tours to determine the inter-observer reproducibility of the Milic and Salaffi scores. In the event of a discrepancy between Tours and Rennes results, the US score evaluated by the examiner in Rennes (GC) was used.

2.4. Statistical A

The statistical analysis software used was SPSS-statistics 19.0. The sensitivity (Se), the specificity (Sp) and the positive likelihood ratio (LR+) of the salivary gland US for the SS diagnosis were cal-

Table 1
Characteristics of the study population.

	pSS n (%)	sSS n (%)	Controls n (%)
<i>n</i>	39	22	36
Age (\pm SD years)	59.1 (13.4)	55.3 (15.0)	55.8 (11.9)
Female	36 (92.3)	22 (100)	34 (94.4)
Current tabagism	10 (27.7)	2 (9.1)	5 (12.8)
Use of drugs causing sicca symptoms	4 (10.3)	1 (4.5)	3 (8.3)
Clinical characteristics			
Raynaud's phenomenon	20 (51.3)	9 (41.0)	17 (47.2)
Inflammatory arthralgia	33 (84.6)	22 (100)	29 (80.6)
Clinical parotitis/submaxillitis	23 (59.0)	5 (22.7)	11 (30.5)
Interstitial lung disease	4 (10.2) ^a	4 (18.2) ^b	1 (2.8) ^{ab}
Pleurisy	0	1 (4.5)	0
Pericarditis	1 (2.6)	1 (4.5)	0
Peripheral neuropathy	5 (12.8)	2 (9.1)	1 (2.8)
Central neuropathy	1 (2.6)	1 (4.5)	0
Tubulo-interstitial nephropathy	0	1 (4.5)	0
Glomerulonephritis	3 (7.7)	4 (18.2)	1 (2.8)
(lupus/cryoglobulinemia)			
Thrombopenia	2 (5.1)	1 (4.5)	0
Leucopenia	4 (10.2)	0	0
Non-Hodgkin's lymphoma	3 (7.7)	0	0
Vascular purpura	3 (7.7)	1 (4.5)	1 (2.8)
Cutaneous lupus	6 (15.4)	7 (31.8)	3 (8.3)
Cryoglobulinemic vasculitis	3 (7.7)	2 (9.0)	1 (2.8)
Secondary antiphospholipid syndrome	0	2 (9.0)	0
Immunosuppressive therapy	6 (15.4)	13 (59.1)	3 (8.3)
Severe visceral disease	7 (17.9) ^{ab}	8 (36.4) ^a	3 (8.3) ^b
Connective tissue disease associated			
None (<i>n</i> = 62)	39 (100)	–	23 (63.9)
Rheumatoid arthritis (<i>n</i> = 15)	–	9 (41.0)	6 (16.7)
Systemic lupus erythematosus (<i>n</i> = 13)	–	9 (41.0)	4 (11.1)
Systemic sclerosis (<i>n</i> = 3)	–	2 (9.0)	1 (2.8)
Sharp's disease (<i>n</i> = 3)	–	2 (9.0)	1 (2.8)
Overlap syndrome (<i>n</i> = 1)	–	0	1 (2.8)
Clinical sicca syndrome			
Duration of the sicca symptoms			
< 5 years	12 (30.8)	6 (27.3)	17 (47.2)
5–10 years	14 (35.9)	4 (18.2)	11 (30.6)
≥ 10 years	13 (33.3) ^{ab}	12 (54.5) ^a	8 (22.2) ^a
Xérophtalmia	38 (97.4)	21 (95.5)	35 (97.2)
Schirmer's test (\pm SD)	5.36 (7.1)	1.27 (1.83)	9.32 (9.31)
Schirmer's test < 5 mm	28 (71.8) ^a	22 (100) ^b	14 (38.9) ^{ab}
Xerostomia	38 (97.4) ^a	21 (95.5) ^b	24 (66.7) ^{ab}
Salivary flow (\pm SD)	1.3 (1.26)	1.2 (0.96)	2.1 (1.50)
Salivary flow < 0.5 mL/5 min	9/25 (36.0) ^a	8/19 (42.0) ^b	0 ^{ab}
Immunologic disorders			
ANA $\geq 1/320$	32 (82.1) ^a	19 (86.4) ^b	11 (30.5) ^{ab}
Anti-SSA or anti-SSB	30 (76.9) ^a	11 (50.0) ^{ab}	9 (23.1) ^{ab}
Rheumatoid factor	18 (46.2) ^a	11 (50.0) ^b	8 (22.2) ^{ab}

sSS: secondary Sjögren's syndrome; pSS: primary Sjögren's syndrome.

^a Groups are significantly different ($P < 0.05$).

^b Groups are significantly different ($P < 0.05$).

culated with the AECG criteria as gold standard. ROC curves were constructed by computing the Se and Sp of different US scores in the pSS and sSS groups. For each score, the plot determined the area under the curve (AUC). The AUC was used to evaluate the diagnostic performance. We computed also the optimal cut-off US point for each score. It corresponds to the score for which the sum of sensitivity and specificity is the highest and with the most significant LR+ for SS diagnosis. The agreement of the intra- and inter-observer US results was evaluated by Cohen's non-parametric Kappa test. Kruskal-Wallis non parameter test was used to compare continue variables. Chi² adjusted by Bonferroni method was used to compare qualitative variables. Differences were considered significant if P

Table 2
Median of salivary gland US scores for each group.

Salivary gland US score	pSS		sSS		Controls	
	n = 39		n = 22		n = 36	
	Median	[IQR]	Median	[IQR]	Median	[IQR]
Hocevar	23.0	[13.0–33.0] ^a	22.5	[5.75–25.25] ^b	4.0	[2.0–6.75] ^{a,b}
Milic	7.0	[2.0–9.0] ^a	5.5	[0.0–7.0] ^b	0.0	[0.0–0.0] ^{a,b}
Salaffi	11.0	[4.0–13.0] ^a	8.0	[1.5–11.0] ^b	0.0	[0.0–0.0] ^{a,b}
Jousse-Joulin	2.0	[1.0–3.0] ^a	1.5	[1.0–3.0] ^b	0.0	[0.0–0.0] ^{a,b}

primary Sjögren's syndrome; sSS: secondary Sjögren's syndrome. IQR: interquartile range
Kruskal-Wallis non-parametric test.

^a p

^b p

A.

US scores	AUC (CI95%)	Optimal cut off point	Se (%) (CI95%)	Sp (%) (CI95%)	LR+
Hocevar	0.891 (0.812-0.970)	≥ 17	61.5 (45.9-75.1)	91.7 (78.2-97.1)	7.4
Milic	0.885 (0.804-0.965)	≥ 3	65.8 (49.9-78.8)	94.4 (81.9-98.5)	11.8
Salaffi	0.915 (0.848-0.982)	≥ 7	63.2 (48.1-76.3)	94.4 (81.9-98.5)	11.3
Jousse-Joulin	0.897 (0.821-0.973)	≥ 2	65.8 (49.9-78.8)	94.4 (81.9-98.5)	11.8

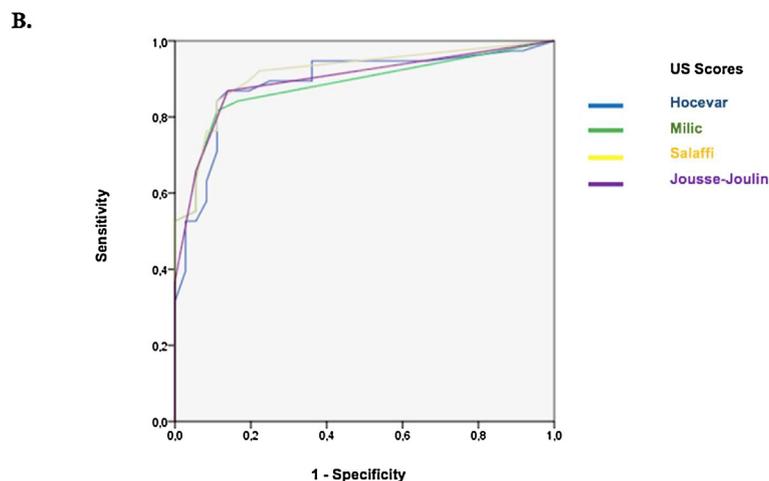


Fig. 2. Performance for each of the four US scores for the pSS diagnosis in comparison to the control group. A. Table with AUC, optimal cut-off point, Sensitivity (Se), Specificity (Sp) and Positive Likelihood Ratio (LR+) for each of the four scores. B: ROC curves of the four US scores for pSS diagnosis.

2.5. Role of the funding source

Not applicable.

3. Results

3.1. Characteristics of the studied population

Five patients were excluded (3 patients were not at US appointment, 1 patient did not have the questionnaire, 1 patient died). Ultimately, 97 patients were included for the analysis. Thirty-nine of them met the AECG criteria for the diagnosis of pSS and 22 patients met the AECG criteria for the diagnosis of sSS (9 patients with systemic lupus, 9 with rheumatoid arthritis, 2 with systemic sclerosis and 2 with Sharp syndrome). There were 36 control patients with sicca symptoms who did not fulfill the AECG criteria (13 patients with associated connective tissue disease and 23 without associated systemic disease).

There was no significant difference between groups regarding the mean age, sex, tabagism, use of medications causing a sicca syndrome (Table 1).

3.2. Comparison of four US scores (Table 2)

The medians for each of the four US scores (Salaffi, Hocevar, Milic, Jousse-Joulin) was significantly higher in the pSS and sSS groups than in the control group ($P < 0.001$). The medians scores were higher in the pSS group compared to the sSS group but with no statistically significant difference.

3.3. Comparison of performance of the four scores for diagnosis of pSS (Fig. 2)

Diagnostic performance was higher for the Salaffi score with AUC of 0.915 (95%CI 0.848-0.982), followed by the Jousse-Joulin score with AUC of 0.897 (95%CI 0.821-0.973), Hocevar score of 0.891 (95%CI 0.812-0.970) and Milic score of 0.885 (95%CI 0.804-0.965). The LR+ was significant (> 11.0) for the Milic, Salaffi and Jousse-Joulin scores.

3.4. Comparison of performance of the four US scores for diagnosis of sSS (Fig. 3)

The Jousse-Joulin score had the best performance with an AUC of 0.851 (95% CI: 0.735-0.968), followed by Salaffi with AUC of

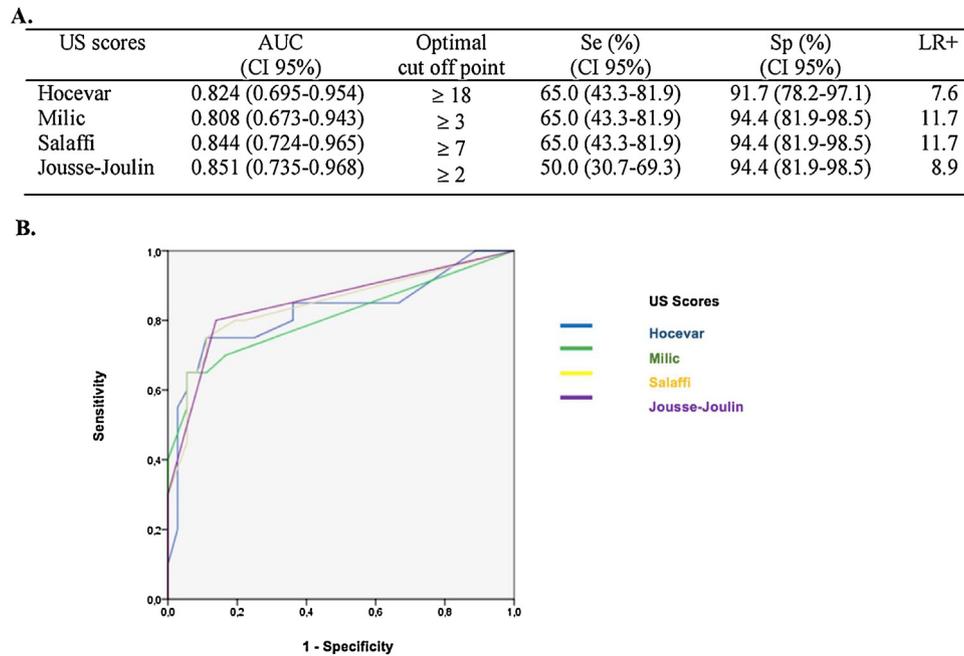


Fig. 3. Performance of each of the four US scores for the sSS diagnosis in comparison to the control group. A. Table with AUC, optimal cut-off point, Sensitivity (Se), Specificity (Sp) and Positive Likelihood Ratio (LR+) for each of the four scores. B: ROC Curves of the four US scores for sSS diagnosis.

0.844 (95%CI 0.724–0.965), Hocevar with AUC of 0.824 (95% CI: 0.695–0.954) and Milic with AUC of 0.808 (95%CI 0.673–0.943). The sensitivity was 65% for the Hocevar, Milic and Salaffi scores with respective optimal cut-off points ≥ 18 , 3 and 7. For the Jousse-Joulin score, the optimal cut-off point was ≥ 2 and was associated with 50% sensitivity. The specificity of the four scores was excellent, between 91.7% and 94.4%. The LR+ was significant (> 11.0) only for the Milic and Salaffi scores.

3.5. Reproducibility of US diagnosis

Inter-observer reproducibility was calculated for the Salaffi and Milic scores for 28 patients. The results showed low reproducibility for the accurate count of the Salaffi and Milic scores, respectively 0.274 ± 0.083 and 0.173 ± 0.069 . On the other hand, the inter-observer reproducibility was better in determining the pathological character of the US as a function of the optimal cut-off point. For the Salaffi score, the inter-observer agreement showed a kappa-value of 0.714 ± 0.131 with the optimal cut-off point ≥ 7 . For the Milic score, the kappa-value was 0.721 with the optimal cut-off point ≥ 3 . The agreement between reader in determining the pathological character of US was 24/28 (85.7%).

The intra-observer reproducibility of the Salaffi score was evaluated on 79 patients. The results showed low intra-observer reproducibility for the accurate measurement of the Salaffi score (0.516 ± 0.098 (AB) and 0.412 ± 0.064 (GC)). However, there was a better intra-observer reproducibility to determine the pathological character of US according to the optimal cut-off score. The kappa agreement was 0.68 ± 0.127 (GC) and 0.714 ± 0.132 (AB).

4. Discussion

Our study confirms the high diagnostic performance of salivary gland US to discriminate pSS, sSS and symptomatic control patients with 3 US scores (Milic, Salaffi, and Jousse-Joulin). The medians of the scores were significantly higher in the pSS and sSS groups than in the control group. The Milic, Salaffi and Jousse-Joulin scores were significant and clinically relevant (LR+ > 11.0) for the diagnosis of

pSS. The LR+ was also greater than 11 with the Milic and Salaffi scores for the diagnosis of sSS.

For the diagnosis of pSS, the optimal cut-off scores and the AUC were in agreement with previous studies for three scoring systems (Jousse-Joulin, Hocevar and Salaffi). Jousse-Joulin et al. [15] also had an optimal US cut-off score ≥ 2 , with a lower AUC (0.817) than in our AUC study (0.897). In the study of Hocevar et al. [16], the optimal cut-off score was the same and the AUC was 0.860 ± 0.031 , just below the AUC of our study at 0.891. Zhang et al. [18] used the Salaffi score and found the same optimal cut-off score ≥ 7 and the AUC was 0.939, slightly higher than in our study (AUC was 0.915). On the other hand, we noted a difference of results for the simplified score system of Milic et al. [17]. The AUC was 0.96 ± 0.01 with an optimal cut-off score ≥ 6 in their study, while we found an optimal cut-off score ≥ 3 with an AUC of 0.885.

These four studies had also used the AECG criteria for the diagnosis of pSS, which allows a good comparison of patient groups with a pSS. Lin et al. [19] also assessed 3 US scoring systems (Hocevar, Milic and Salaffi). Hocevar score had the best LR+ (LR+ 10, AUC 0.888), whereas Milic score was recommended for its excellent balance between the diagnostic value (LR+ 6.3, AUC 0.866) and inter/intra-observer agreement value. The AUC for Salaffi score was 0.851 and LR+ to 3.8 was lower. However, this study did not use the same inclusion criteria from AECG (scintigraphy and Rose Bengal) than in our study. Furthermore, results are not comparable, because Lin et al. included sSS patients in their control group.

For the diagnosis of sSS, the AUC was evaluated between 0.808 and 0.851 according to the scoring method. Very few studies have evaluated the salivary glands US during sSS. In 1992, De Vita et al. [12] obtained a sensitivity of 56.8% and a specificity of 92.2%, i.e. an LR+ of 6.45 compared to the control group, but with different diagnostic criteria for sSS (Copenhagen 1975). Wernicke et al. [20] used an US score derived from De Vita et al. with AECG criteria as gold standard. They showed a sensitivity of 63.6%, a specificity of 98% and a LR+ of 31 for the diagnosis of sSS, well beyond our results.

According to our study, the major salivary glands US appears to be efficient to discriminate a sSS from a patient with connective tissue disease or a sicca syndrome. In clinical practice, the diagnosis of sSS does not change the treatment of the associated

connective tissue disease (except for local treatment). A Danish study suggested anyway that the sSS had no impact on primary disease activity or on treatment [21]. However, some studies have shown that patients with an sSS had an increased risk of developing NHL (twofold) and higher mortality compared to the group of patients with rheumatoid arthritis without sSS [22,23]. The establishment of the sSS diagnosis should therefore encourage a more regular follow-up of these patients to detect a possible NHL.

The high diagnostic performance of salivary gland US to discriminate pSS and sSS was demonstrated in our study even if in our group control, 13 (36%) patients had connective tissue disease without AECG criteria. A cross-sectional study [24] consecutively included 109 patients (55 with pSS and 54 with undifferentiated connective tissue diseases) UCTD showed a higher Milic score in pSS (mean 2.2 ± 1.8 vs. 0.2 ± 0.5). The Milic score cut-off ≥ 2 showed a LR+ of 16.25 for pSS diagnosis in this study.

Few studies have evaluated the reproducibility of salivary gland US during the pSS. Salaffi et al. [14] have measured the inter-observer reproducibility of their score with two blind radiologists with a high kappa agreement of 0.88 for the submandibular and 0.93 for the parotids. Hocevar et al. [25] have obtained a good inter-observer agreement (with two blind observers) with a global kappa of 0.90. A high level of agreement between the two observers was seen in evaluation of echogenicity, inhomogeneity and the presence of hypoechoic areas ($k = 0.88, 0.90$, and 0.88 , respectively).

El Miedany et al. [9] evaluated intra-observer reproducibility in only 13 patients (7 pSS and 6 controls) and using another US score (0–3) about the parotid glands. Patients were examined twice. Kappa value was 0.8. In the study of Salaffi et al. [7], intra-observer reproducibility was evaluated for only 12 patients with sicca symptoms without pSS or with pSS, for the parenchymal homogeneity ($k = 0.852$), echogenicity ($k = 0.821$), size of the glands and posterior glandular border clearness ($k = 0.804$). The second reading was evaluated from the recorded images.

Our inter-observer reproducibility was lower on the overall score, suggesting the importance of a common learning in measuring the scores. The need of developing a consensual system is the main objective of the international US-pSS study group who reported on acquisition images, a substantial parotid glands inter-observer reliability ($\kappa = 0.62$) for echogenicity and moderate ($\kappa = 0.52$) for homogeneity. Submandibular glands inter-observer reliability was moderate for homogeneity ($\kappa = 0.46$) and fair for echogenicity ($\kappa = 0.38$) [26].

We did not find an association in our study between the duration of evolution of dry syndrome in Sjogren's syndrome and a high US score of the salivary glands. Salivary glands US abnormalities assessed (Jousse-Joulin score) did not change significantly during a follow-up (nearly 2 years after an initial evaluation) for suspected pSS [27].

Mossel E et al. [28] have even proposed to score hypoechoic areas in a parotid and submandibular gland increases the feasibility of ultrasound in primary Sjögren's syndrome. The performances for the positive diagnosis are not diminishing. However, this practice in daily clinic ask question because differential diagnoses isn't taking into account such as salivary gland lymphomas, other tumors, lithiasis or other inflammatory diseases of the salivary glands as sarcoidosis which, in our opinion, require evaluation of the 4 main salivary glands.

The score of Jousse-Joulin seems to have a good balance between simplicity and diagnostic performance based on our results. However, the cysts were the only criterion of heterogeneity of the glandular parenchyma and not the fat involution, quite frequent in the submandibular glands. This fat involution is characterized by a difficulty in visualizing the salivary gland in its environment with disappearance of the gland borders and the presence of numerous thick hyperechoic bands. These characteristics are not taken

into account in the Jousse-Joulin score resulting in a scoring of 2, whereas the scoring in the Salaffi score of the glands would be 4 (Appendix A).

Finally, the score proposed by Salaffi et al. [14] has the advantage of having several items to better define the parenchymal echotexture (inhomogeneity, hypoechogenic areas, echogenic bands, cysts, visibility of the posterior glandular border), contrary to the Milic score which proposes a less detailed definition in 3 grades with a higher risk of subjectivity, and the Jousse-Joulin score which preferentially evaluates the cystic degeneration of the glandular parenchyma. The Hocevar score is very detailed but longer to carry out, with no supplementary performance compared to the scores of Salaffi and Jousse-Joulin, according to our results.

The US of the major salivary glands showed a good diagnostic performance to differentiate a primary or a secondary SS from a control patient with sicca symptoms. According to some authors, it could replace sialography and salivary scintigraphy within AECG criteria without modifying them [7,8]. Cornec et al. [15] showed that the addition of US to the AECG criteria increased the sensitivity of the AECG criteria from 77.9% to 87% without altering the specificity (96.1% vs. 98.7%) for the pSS diagnosis. They also found that the addition of salivary gland US to the 2012 American College of Rheumatology (ACR) classification criteria increased their sensitivity from 64.4% to 84.4% without modifying the specificity (89.3% vs. 91%) [29]. Takagi et al. [30] suggested that the US could replace all ACR criteria, but without increasing the diagnostic performance of the 2012 ACR criteria, otherwise US can improve the accuracy of the integrated ACR/EULAR system and US score (93%) by compared to ACR/EULAR criteria only (86%) [31]. In a study [32] with 290 patients prospectively included [125 (43%) met ACR/EULAR criteria and 114 (39%) also met AECG criteria], US salivary gland was abnormal in 12% of the 165 patients fulfilling no criteria set. Including salivary gland US score among the ACR/EULAR criteria increased sensitivity from 87.4% to 91.1% when physician diagnosis was the reference standard.

In conclusion, the salivary gland US is a simple, non-invasive and efficient method for the diagnosis of pSS and sSS. However, setting up a validated consensual score, and training practitioners to evaluate this US score will still be necessary to improve its reproducibility, before using it in clinical practice within diagnostic criteria.

Contributorship

All listed authors have had substantial contributions in the acquisition, analysis, and interpretation of data for the work; they were involved in revising it critically and give final approval for the version to be published. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Disclosure of interest

The authors declare that they have no competing interest.

Acknowledgments

The authors would like to thank the patients who participated in the study and the Doctor Christophe Boulnois for his proofreading and translation work in the English language.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at <https://doi.org/10.1016/j.jbspin.2018.06.014>.

References

- [1] Theander E, Henriksson G, Ljungberg O, et al. TMandIRManthorpeLTJacobssonLymphoma and other malignancies in primary Sjögren's syndrome: a cohort study on cancer incidence and lymphoma predictors. *Ann Rheum Dis* 2006;65:796–803.
- [2] Zintzaras E, Voulgarelis M, HM. The risk of lymphoma development in autoimmune diseases: a meta-analysis. *Arch Intern Med* 2005;165:2337–44.
- [3] Cornec D, Saraux A, Pers JO, et al. Diagnostic accuracy of blood B-cell subset profiling and autoimmunity markers in Sjögren's syndrome. *Arthritis Res Ther* 2014;16:R15.
- [4] Fauchais AL, Martel C, Gondran G, et al. MLambertDLaunayMOJauberteaulmunological profile in primary Sjögren syndrome: clinical significance, prognosis and long-term evolution to other auto-immune disease. *Autoimmun Rev* 2010;9:595–9.
- [5] Bourmia VK, Vlachoyiannopoulos PG. Subgroups of Sjögren syndrome patients according to serological profiles. *J Autoimmun* 2012;39:15–26.
- [6] 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.
- [7] Salaffi F, Carotti M, Iagnocco A, et al. Ultrasonography of salivary glands in primary Sjogren's syndrome: a comparison with contrast sialography and scintigraphy. *Rheumatology (Oxford)* 2008;47:1244–9.
- [8] 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 Sjogren's syndrome. *Rheumatology (Oxford)* 2012;51:1081–5.
- [9] El Miedany YM, Ahmed I, Mourad HG, et al. Quantitative ultrasonography and magnetic resonance imaging of the parotid gland: can they replace the histopathologic studies in patients with Sjogren's syndrome? *Joint Bone Spine* 2004;71:29–38.
- [10] Makula E, Pokorny G, Kiss M, et al. The place of magnetic resonance and ultrasonographic examinations of the parotid gland in the diagnosis and follow-up of primary Sjogren's syndrome. *Rheumatology (Oxford)* 2000;39:97–104.
- [11] Niemelä RK, Takalo R, Pääkkö E, et al. Ultrasonography of salivary glands in primary Sjogren's syndrome. A comparison with magnetic resonance imaging and magnetic resonance sialography of parotid glands. *Rheumatology (Oxford)* 2004;43:875–9.
- [12] De Vita S, Lorenzon G, Rossi G, et al. Salivary gland echography in primary and secondary Sjögren's syndrome. *Clin Exp Rheumatol* 1992;10:351–6.
- [13] Chisholm DM, Mason DK. Labial salivary gland biopsy in Sjögren's disease. *J Clin Pathol* 1968;21:656–60.
- [14] Salaffi F, Argalia G, Carotti M, et al. Salivary gland ultrasonography in the evaluation of primary Sjögren's syndrome. Comparison with minor salivary gland biopsy. *J Rheumatol* 2000;27:1229–36.
- [15] Cornec D, Jousse-Joulin S, Pers JO, et al. Contribution of salivary gland ultrasonography to the diagnosis of Sjogren's syndrome: toward new diagnostic criteria? *Arthritis Rheum* 2013;65:216–25.
- [16] Hocevar A, Ambrozic A, Rozman B, et al. Ultrasonographic changes of major salivary glands in primary Sjogren's syndrome. Diagnostic value of a novel scoring system. *Rheumatology (Oxford)* 2005;44:768–72.
- [17] Milic VD, Petrovic RR, Boricic IV, et al. Major salivary gland sonography in Sjogren's syndrome: diagnostic value of a novel ultrasonography score (0–12) for parenchymal inhomogeneity. *Scand J Rheumatol* 2010;39:160–6.
- [18] Zhang X, Zhang S, He J, et al. Ultrasonographic evaluation of major salivary glands in primary Sjögren's syndrome: comparison of two scoring systems. *Rheumatology (Oxford)* 2015;54:1680–7.
- [19] Lin D, Yang W, Guo X, et al. Cross-sectional comparison of ultrasonography scoring systems for primary Sjogren's syndrome. *Int J Clin Exp Med* 2015;8:19065–71.
- [20] Wernicke D, Hess H, Gromnica-Ihle E, et al. Ultrasonography of salivary glands – a highly specific imaging procedure for diagnosis of Sjogren's syndrome. *J Rheumatol* 2008;35:285–93.
- [21] Haga HJ, Naderi Y, Moreno AM, et al. A study of the prevalence of sicca symptoms and secondary Sjogren's syndrome in patients with rheumatoid arthritis, and its association to disease activity and treatment profile. *Int J Rheum Dis* 2012;15:284–8.
- [22] Kauppi M, Pukkala E, Isomäki H. Elevated incidence of hematologic malignancies in patients with Sjogren's syndrome compared with patients with rheumatoid arthritis (Finland). *Cancer Causes Control* 1997;8:201–4.
- [23] Martens PB, Pillemer SR, Jacobsson LT, et al. Survivorship in a population based cohort of patients with Sjogren's syndrome, 1976–1992. *J Rheumatol* 1999;26:1296–300.
- [24] 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 (Oxford)* 2015;54:2198–204.
- [25] Hocevar A, Rainer S, Rozman B, et al. Ultrasonographic changes of major salivary glands in primary Sjogren's syndrome. Evaluation of a novel scoring system. *Eur J Radiol* 2007;63:379–83.
- [26] 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.
- [27] Gazeau P, Cornec D, Jousse-Joulin S, et al. Time-course of ultrasound abnormalities of major salivary glands in suspected Sjögren's syndrome. *Joint Bone Spine* 2018;85:227–32.
- [28] 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.
- [29] Cornec D, Jousse-Joulin S, Marhadour T, et al. Salivary gland ultrasonography improves the diagnostic performance of the 2012 American College of Rheumatology classification criteria for Sjögren's syndrome. *Rheumatology (Oxford)* 2014;53:1604–7.
- [30] Takagi Y, Sumi M, Nakamura H, et al. Ultrasonography as an additional item in the American College of Rheumatology classification of Sjögren's syndrome. *Rheumatology (Oxford)* 2014;53:1977–83.
- [31] 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.
- [32] 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.