University of Groningen

A comprehensive overview of major salivary gland ultrasonography in (suspected) primary Sjögren's syndrome

Mossel, Esther

DOI:
[10.33612/diss.172042654](https://doi.org/10.33612/diss.172042654)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Mossel, E. (2021). *A comprehensive overview of major salivary gland ultrasonography in (suspected) primary Sjögren's syndrome*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.172042654>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

**A comprehensive overview
of major salivary gland
ultrasonography
in (suspected) primary
Sjögren's syndrome**

Esther Mossel

The research presented in this thesis was performed at the Sjögren's expterise centre, University Medical Center Groningen.

The printing and distribution of this thesis was supported by:

Rijksuniversiteit Groningen

The Graduate School of Medical Sciences, University Medical Center Groningen

Nationale Vereniging Sjögrenpatiënten

© E. Mossel, 2021

No parts of this thesis may be reproduced, stored in a retrieval system or transmitted in any form by any means, without permission of the author, or, when appropriate, of the publisher of the publication or illustration material.

auteur Esther Mossel

ontwerp Rik Ontwerpt

print Ipskamp Printing

ISBN 978-94-6421-354-6

A comprehensive overview of major salivary gland ultrasonography in (suspected) primary Sjögren's syndrome

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. C. Wijmenga
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 2 juni 2021 om 9.00 uur

door

Esther-Jellina Mossel

geboren op 15 oktober 1990
te Hurdegaryp

Promotores

Prof. dr. H. Bootsma

Prof. dr. F.G.M. Kroese

Copromotor

Dr. S. Boezeroij-Arends

Beoordelingscommissie

Prof. dr. A.W.J.M. Glaudemans

Prof. dr. V. Devauchelle-Pensec

Prof. dr. J.M. Van Laar

Paranimfen

Mw. G.S. Van Zuiden

Mw. M.S. Van Ginkel

Content

Chapter 1	General introduction	9
Chapter 2a	Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome.	25
Chapter 2b	The parotid gland connection: ultrasound and biopsies in primary Sjögren's syndrome.	53
Chapter 3	Scoring hypoechogenic areas in one parotid and one submandibular gland increases feasibility of ultrasound in primary Sjögren's syndrome.	61
Chapter 4	Histopathology, salivary flow and ultrasonography of the parotid gland: three complementary measurements in primary Sjögren's syndrome.	89
Chapter 5	Incorporation of salivary gland ultrasonography into the ACR-EULAR criteria for primary Sjögren's syndrome.	117
Chapter 6	Clinical phenotyping of primary Sjögren's patients using salivary gland ultrasonography – data from the RESULT cohort.	143
Chapter 7	Can ultrasound of the major salivary glands assess histopathological changes induced by treatment with rituximab in primary Sjögren's syndrome?	179

Chapter 8	Abatacept treatment for patients with early active primary Sjögren's syndrome: a single-centre, randomised, double-blind, placebo-controlled, phase 3 trial (ASAP-III study).	187
Chapter 9	One year in review: recent insights in the potential role of imaging modalities for diagnosing patients with primary Sjögren's syndrome.	229
Chapter 10	General discussion	243
Appendices	Summary Samenvatting Dankwoord Curriculum vitae List of publications List of abbreviations	265

A decorative border with a pattern of pink and white polka dots surrounds the central text area.

CHAPTER 1

General introduction

Primary Sjögren's Syndrome

Primary Sjögren's syndrome (pSS) is a common, systemic, auto-immune disease, with an estimated prevalence of 0.04% in the general population.^{1,2} Females are affected nine times more often compared with men, and the peak incidence lies between the age of forty and sixty years.²

The disease is characterized by lymphocytic infiltration of the exocrine glands, which can involve all tear and salivary glands, and leads to a sensation of dry eyes (keratoconjunctivitis sicca) and dry mouth (xerostomia).³ Besides sicca complaints, patients often experience extreme fatigue and arthralgia.^{1,3} Moreover, in about 30-40% of pSS patients, systemic manifestations like arthritis, vasculitis and peripheral nerve involvement can occur.^{1,3}

The lymphocytic infiltrates in the salivary glands are in particular situated around striated ducts and form periductal foci.^{4,5} Foci consist of clusters of ≥ 50 mononuclear cells, and the number of foci per 4 mm² of glandular tissue is calculated into a focus score.⁶⁻⁸ Another histopathological feature that can be found in the salivary glands of pSS patients is a relative decrease in the number of immunoglobulin A (IgA) expressing plasma cells, due to an influx of IgG and IgM expressing plasma cells.⁹⁻¹¹ Furthermore, presence of germinal centers (GCs), and lymphoepithelial lesions (LELs), which are composed of hyperplastic ductal epithelial cells, with intra-epithelial B and T lymphocytes are frequently seen.^{5,12-15}

Overall, T cell-dependent B cell hyperactivity seems to be characteristic in pSS. This hyperactivity is expressed in the form of auto-antibodies, like anti-nuclear antibodies (ANAs), including anti-SSA/Ro and anti-SSB/La antigens, but can also manifest as hypergammaglobulinemia or presence of cryoglobulins.³

Unfortunately, the exact pathophysiologic mechanism of pSS still remains unclear.³ Thus far it appears to be multifactorial, with both genetic and environmental factors contributing to the disease initiation.^{3,16}

Classification and diagnosis

In recent decades, different classification criteria sets for pSS have been developed.¹⁷⁻²⁰ In 2002, the American European Consensus Group (AECG) criteria were introduced.¹⁷ The AECG criteria consisted of 6 items, of which 2 were subjective and 4 were objective. Cornerstones were presence of anti-SSA/Ro and/or anti-SSB/La antibodies, and a focus score of ≥ 1 (i.e. ≥ 50 lymphocytes per 4mm² of glandular tissue) in a salivary gland biopsy. In order to be classified as a pSS patient, at least 1 of these items had to be positive. The other items were sicca complaints of eyes and mouth (subjective), and functional impairment of the tear- and salivary glands

(objective). A patient was classified as pSS patient if at least 4 out of 6 items were present, including at least one of the cornerstones, or 3 out of 4 objective items were present.¹⁷ For many years, the AECG criteria have been used in daily clinical practice, but questions were raised about the importance of subjective sicca symptoms in the classification of pSS.²¹ Therefore, in 2012, the American College of Rheumatology (ACR) criteria were developed, in order to better define the Sjögren's syndrome phenotype and properly select individuals for clinical trials.¹⁸

The ACR criteria solely consisted of objective items, of which at least 2 items were needed to be present to be classified as pSS patient. Presence of anti-SSA/Ro and/or anti-SSB/La antibodies, or positive rheumatoid factor in combination with an anti-nuclear antibody (ANA) titer of $\geq 1:320$, and a focus score of ≥ 1 in a salivary gland biopsy were key items, complemented by the ocular staining score to measure tear gland involvement.¹⁸ Neither the AECG nor the ACR criteria were endorsed by the ACR and European League Against Rheumatism (EULAR), which again rose the need for international consensus regarding the classification criteria for pSS. This was especially the case since sialography and scintigraphy, both included in the AECG criteria for the measurement of objective salivary gland involvement, had become a bit old-fashioned.²² This gap was filled with the development of the 2016 ACR-EULAR criteria (table 1) by the International Sjögren's Syndrome Criteria Working Group, endorsed by both ACR and EULAR.^{19,20}

The purpose of classification criteria is to help define homogeneous study groups for research purposes.²³ In daily clinical practice, these criteria are often helpful for diagnostic purposes, but expert opinion should be leading for the final diagnosis.

The ACR-EULAR criteria consist of 5 objective items. Again, presence of anti-SSA/Ro antibodies and a focus score ≥ 1 in a salivary gland biopsy are cornerstones, and presence of each of these items accounts for 3 points. For the remaining items, i.e., Schirmer's test, OSS and measurement of unstimulated whole saliva flow (UWS), 1 point is obtained for every positive test.^{19,20} A patient is classified with pSS if ≥ 4 points are obtained. With regard to the eyes, both the tear gland function (Schirmer's test), as well as the involvement (i.e., disease activity and damage) of the eyes are measured (OSS). With regard to the salivary glands, however, only the salivary gland function is measured. This implies that there is a need to also measure salivary gland involvement (i.e., disease activity and damage), which may be possible with ultrasound.

Table 1.

2016 American College of Rheumatology/ European League Against Rheumatism classification criteria^{19,20}

The classification of primary Sjögren's syndrome (SS) applies to any individual who meets inclusion criteria,* does not have any of the conditions listed as exclusion criteria,† and has a score of ≥ 4 when the weights from the five criteria items below are summed

Item	Weight/score
Labial salivary gland with focal lymphocytic sialadenitis and focus score of ≥ 1 foci/4mm ²	3
Anti-SSA/Ro-positive	3
Ocular Staining Score ≥ 5 (or van Bijsterveld score ≥ 4) in at least one eye	1
Schirmer's test ≤ 5 mm/5 min in at least one eye‡	1
Unstimulated whole saliva flow rate ≤ 0.1 mL/min‡	1

* These inclusion criteria are applicable to any patient with at least one symptom of ocular or oral dryness, defined as a positive response to at least one of the following questions: (1) Have you had daily, persistent, troublesome dry eyes for more than 3 months? (2) Do you have a recurrent sensation of sand or gravel in the eyes? (3) Do you use tear substitutes more than three times a day? (4) Have you had a daily feeling of dry mouth for more than 3 months? (5) Do you frequently drink liquids to aid in swallowing dry food? or in whom there is suspicion of Sjögren's syndrome (SS) from the European League Against Rheumatism SS Disease Activity Index questionnaire (at least one domain with a positive item).

† Exclusion criteria include prior diagnosis of any of the following conditions, which would exclude diagnosis of SS and participation in SS studies or therapeutic trials because of overlapping clinical features or interference with criteria tests: (1) history of head and neck radiation treatment, (2) active hepatitis C infection (with confirmation by PCR), (3) AIDS, (4) sarcoidosis, (5) amyloidosis, (6) graft-versus-host disease, (7) IgG4-related disease.

‡ Patients who are normally taking anticholinergic drugs should be evaluated for objective signs of salivary hypofunction and ocular dryness after a sufficient interval without these medications in order for these components to be a valid measure of oral and ocular dryness.

Figure 1. Range of sound frequencies.

B-mode ultrasonography

Ultrasound refers to sound waves above the range of frequencies that can be heard by humans (20Hz – 20KHz; figure 1), and enables visualization of deep body structures.²⁴ The tissue structures return a sudden amount of waves, which produce echo's and are reproduced as an image. For diagnostic imaging, frequencies between 2MHz and 22MHz are used. The most commonly used ultrasonography mode is B-mode, which stands for brightness, and is used for real-time evaluation.^{24,25}

B-mode salivary gland ultrasonography (SGUS) appears to be an promising imaging technique to assess salivary gland involvement in pSS.²⁶ SGUS is patient-friendly, as it is non-invasive and does not involve radiation. In addition, ultrasonographic devices are already widely available in the outpatient rheumatologic clinic, making SGUS an inexpensive and easily accessible tool.^{25,27} Furthermore, SGUS might be of value in the detection of mucosa-associated lymphoid tissue (MALT)-lymphoma, which occurs in 5-10% of pSS patients.²³ However, training regarding SGUS and scoring of the images is required, since it is an operator-dependent technique.

At the University Medical Center Groningen (UMCG), B-mode SGUS is performed using the MyLabSeven scanner (Esaote, Genova, Italy), equipped with a high-resolution linear probe (4-13 MHz). Both parotid and submandibular glands are examined and the Hocevar scoring system is applied investigating: (i) parenchymal echogenicity compared with the thyroid gland (graded 0-1); (ii) homogeneity (graded 0-3); (iii) presence of hypoechoogenic areas (graded 0-3); (iv) hyperechoogenic reflections (graded 0-3 in parotid glands and 0-1 in submandibular glands); and (v) clearness of salivary gland border (graded 0-3).²⁸ Total ultrasound score is calculated as the sum of these five domains and can range from 0 to 48.²⁸

Measurements of disease activity and damage

Several tools to measure different aspects of disease activity and disease damage have been developed in pSS.

Symptom burden is assessed by the EULAR Sjögren's Syndrome Patient-Reported Index (ESSPRI).²⁹ ESSPRI is a patient-administered questionnaire and consists of three questions about experiencing symptoms of dryness, fatigue and pain in the past 2 weeks.²⁹ All items are collected on a visual analogue scale ranging from 0 to 10 and the ESSPRI total score is calculated as the sum of the three questions divided by three.²⁹

Systemic disease activity is assessed by the EULAR Sjögren Syndrome Disease Activity Index (ESSDAI).³⁰ The ESSDAI consists of the following 12 domains: cutaneous, respiratory, renal, articular, muscular, peripheral nervous system (PNS), central nervous system (CNS), hematological, glandular, constitutional, lymphadenopathy and biological domain and within each domain three or four levels of activity (no, low, moderate, high) are possible.³⁰ The sum of the different domains results in the total score (range 0-123), with a total score of 0-4 defined as low activity, 5-13 as moderate activity and ≥ 14 as high activity.³⁰

Both ESSDAI and ESSPRI are being applied as outcome measurements in clinical trials, as well as in observational cohort studies. Due to the poor correlation between both outcome parameters it is justified to indeed use both, as they appear complementary to each other.³¹

Since disease flares can also cause irreversible damage in pSS patients, the Sjögren's Syndrome Disease Damage Index (SSDDI) has been developed. The SSDDI consists of the following 6 domains: oral/salivary damage, ocular damage, neurologic damage, pleuropulmonary damage, renal impairment, and lymphoproliferative disease, and within each domain different items could be scored. The predetermined weight of the items varies between 1, 2 or 5 points, and the total score ranges from 0-16.³²

Treatment options in pSS

There is an unmet need when it comes to curative treatment options in pSS. Regarding dryness symptoms, some relieve can be obtained with tear and saliva substitutes or non-hormonal moisturisers.³³ Besides symptomatic therapy, a broad spectrum of conventional and biological disease modifying anti-rheumatic drugs (DMARDs) have previously been considered as possible treatment options, including abatacept and rituximab, though none has been approved (yet) by the Food and Drug Administration (FDA) and/or European Medicines Agency (EMA).^{33,34}

Initial open-label studies showed promising effects of rituximab, a humanized anti-CD20 monoclonal antibody, and abatacept, a fully human soluble fusion protein of cytotoxic T lymphocyte antigen coupled to the Fc tail of IgG.³⁵⁻⁴² Unfortunately, four large randomized, double-blind, placebo-controlled trials investigating the effectiveness of abatacept (n=2), and rituximab (n=2) failed to meet their primary endpoint.⁴³⁻⁴⁶ Rituximab is, however, still used off-label in a subgroup of patients with specific systemic manifestations, as the debate considering the potential efficacy of rituximab is still ongoing.^{35,47} Regrettably, it is unknown how many patients are treated with off-label systemic therapy, and for whom it is of benefit.⁴⁷

Of note, however, other potential drugs for systemic treatment are currently under development.⁴⁸⁻⁵⁰

RESULT cohort

Since pSS is a very heterogeneous systemic disease, this implies that not every symptom is present in every single patient. Prospective cohort studies are becoming more important, as information is gathered that could help to predict the long-term course of pSS.⁵¹ Within the UMCG, there is a unique multidisciplinary collaboration in order to provide optimal care for patients diagnosed with or suspected of having pSS. The REgistry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) cohort is an ongoing prospective observational cohort study with standardized follow-up visits. The main objective of the RESULT cohort is to identify biomarkers and clinical parameters that determine and predict the longitudinal course of pSS. All different aspects of the disease are assessed in a protocollary manner, including systemic disease activity, patient-reported outcome measurements, serological parameters and ultrasonographic evaluations. This in order to provide information on the long-term outcome of pSS.⁵²

References

1. Mariette X, Criswell A. Primary Sjögren's syndrome. *N Engl J Med* 2018;378:931-9.
2. Qin B, Wang J, Yang Z, *et al.* Epidemiology of primary Sjögren's syndrome: a systematic review and meta-analysis. *Ann Rheum Dis* 2015;74:1983-9.
3. Fox RI Sjögren's syndrome. *Lancet* 2005;366:321-31.
4. Christodoulou MI, Kapsogeorgou EK, Moutsopoulos HM Characteristics of the minor salivary gland infiltrates in Sjögren's syndrome. *J Autoimmun* 2010;34:400-7.
5. Van Ginkel MS, Haacke EA, Bootsma H, *et al.* Presence of intra-epithelial B-lymphocytes is associated with the formation of lymphoepithelial lesions in salivary glands of primary Sjögren's syndrome patients. *Clin Exp Rheumatol* 2019;37Suppl118:42-8.
6. Fisher BA, Jonsson R, Daniels T, *et al.* Standardisation of labial salivary gland histopathology in clinical trials in primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1161-8.
7. Chisholm DM, Mason DK. Labial salivary gland biopsy in Sjögren's disease. *J Clin Pathol* 1968;21:656-60.
8. Greenspan JS, Daniels TE, Talal N, *et al.* The histopathology of Sjögren's syndrome in labial salivary gland biopsies. *Oral Surgery, Oral Med Oral Pathol* 1974;37:217-29.
9. Zandbelt M, Wentink J, De Wilde P, *et al.* The synergistic value of focus score and IgA% score of sublabial salivary gland biopsy for the accuracy of the diagnosis of Sjögren's syndrome: a 10-year comparison. *Rheumatology (Oxford)* 2002;41:819-23.
10. De Wilde P, Vooy's G, Baak J, *et al.* Quantitative immunohistologic and histomorphometric diagnostic criteria for Sjögren's syndrome. *Pathol Res Pr* 1989;185:778-80.
11. Bodeutsch C, De Wilde P, Kater, *et al.* Quantitative immunohistologic criteria are superior to the lymphocytic focus score criterion for the diagnosis of Sjögren's syndrome. *Arthritis Rheum* 1992;35:1075-97.
12. Kroese FGM, Haacke EA, Bombardieri M. The role of salivary gland histopathology in primary Sjögren's syndrome: promises and pitfalls. *Clin Exp Rheumatol* 2018;36:222-33.

13. Risselade AP, Looije MF, Kruize AA, *et al.* The role of ectopic germinal centers in the immunopathology of primary Sjögren's syndrome: a systematic review. *Semin Arthritis Rheum* 2013;42:368-76.
14. Ihrler S, Adam P, Guntinas-Lichius O, *et al.* Pattern recognition in the differential diagnosis of salivary lymphoepithelial lesions. *Pathologe* 2009;30:431-41.
15. Ihrler S, Zietz C, Sendelhofert A, *et al.* Lymphoepithelial duct lesions in Sjögren-type sialadenitis. *Virchows Arch* 1999;434:315-23.
16. Lessard CJ, Li H, Adrianto I, *et al.* Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjögren's syndrome. *Nat Genet* 2013;45:1-26.
17. 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.
18. Shiboski SC, Shiboski CH, Criswell LA, *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.
19. 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.
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. *Arthritis Rheumatol* 2017;69:35-45.
21. Van Nimwegen JF, Van Ginkel MS, Arends S, *et al.* Validation of the ACR-EULAR criteria for primary Sjögren's syndrome in a Dutch prospective diagnostic cohort. *Rheumatology (Oxford)* 2018;57:818-25.
22. Baldini C, Zabotti A, Filipovic N, *et al.* Imaging in primary Sjögren's syndrome: the 'obsolete and the new'. *Clin Exp Rheumatol* 2018;36Suppl112:215-21.
23. Bootsma H, Spijkervet FKL, Kroese FGM, *et al.* Toward new classification criteria for Sjögren's syndrome? *Arthritis Rheum* 2013;65:21-3.
24. Terkawi AS, Karakitsos D, Elbarbary M, *et al.* Ultrasound for the anesthesiologists: present and future *Scientific World Journal* 2013; published online: 13 November 2013.
25. Delli K, Dijkstra PU, Stel AJ, *et al.* Diagnostic properties of ultrasound of major salivary glands in Sjögren's syndrome: a meta-analysis. *Oral Dis* 2015;21:792-800.

26. Mossel E, Delli K, Van Nimwegen JF, *et al.* Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1883-9.
27. Jousse-Joulin S, Milic V, Jonsson MV, *et al.* Is salivary gland ultrasonography a useful tool in Sjögren's syndrome? A systematic review. *Rheumatology (Oxford)* 2016;55:789-800.
28. Hocevar A, Ambrozic A, Rozman B, *et al.* Ultrasonographic changes of major salivary glands in primary Sjögren's syndrome. Diagnostic value of a novel scoring system. *Rheumatology (Oxford)* 2005;44:768-72.
29. Seror R, Ravaud P, Mariette X, *et al.* EULAR Sjögren's Syndrome Patient Reported Index (ESSPRI): development of a consensus patient index for primary Sjögren's syndrome. *Ann Rheum Dis* 2011;70:968-72.
30. Seror R, Bowman SJ, Brito-Zeron P, *et al.* EULAR Sjögren's syndrome disease activity index (ESSDAI): a used guide. *RMD open* 2015;1:e000022.
31. Seror R, Theander E, Brun JG, *et al.* Validation of EULAR primary Sjögren's syndrome disease activity (ESSDAI) and Patient indexes (ESSPRI). *Ann Rheum Dis* 2015;74:859-66.
32. Vitali C, Palombi G, Baldini C, *et al.* Sjögren's Syndrome Disease Damage Index and Disease Activity Index Scoring Systems for the Assessment of Disease Damage and Disease Activity in Sjögren's Syndrome, derived from an analysis of a cohort of Italian patients. *Arthritis Rheum* 2007;56:2223-31.
33. Price EJ, Baer AN. How to treat Sjögren's syndrome. *Rheumatology (Oxford)* 2019; published online: 15 February 2019.
34. Ramos-Casals M, Brito-Zerón P, Bombardieri S, *et al.* EULAR recommendations for the management of Sjögren's syndrome with topical and systemic therapies. *Ann Rheum Dis* 2020;79:3-18.
35. Verstaappen GM, Van Nimwegen JF, Vissink A, *et al.* The value of rituximab treatment in primary Sjögren's syndrome. *Clin Immunol* 2017;182:62-71.
36. Meiners PM, Vissink A, Kroese FGM, *et al.* Abatacept treatment reduces disease activity in early primary Sjögren's syndrome (open-label proof of concept ASAP study). *Ann Rheum Dis* 2014;73:1393-6.
37. Adler S, Körner M, Förger F, *et al.* Evaluation of histologic, serologic, and clinical changes in response to abatacept treatment of primary sjögren's syndrome: a pilot study. *Arthritis Care Res* 2013;65:1862-8.
38. Machado AC, Dos Santos LC, Fidelix T, *et al.* Effectiveness and safety of abatacept for the treatment of patients with primary Sjögren's syndrome. *Clin Rheumatol* 2020;39:243-8.

39. Meijer JM, Meiners PM, Vissink A, *et al.* Effectiveness of rituximab treatment in primary Sjögren's syndrome: a randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2010;62:960-8.
40. Meiners PM, Arends S, Brouwer E, *et al.* Responsiveness of disease activity indices ESSPRI and ESSDAI in patients with primary Sjögren's syndrome treated with rituximab. *Ann Rheum Dis* 2012;71:1297-302.
41. Dass S, Bowman SJ, Vital EM, *et al.* Reduction of fatigue in Sjögren's syndrome with rituximab: Results of a randomised, double-blind, placebo-controlled pilot study. *Ann Rheum Dis* 2008;67:1541-4.
42. Carubbi F, Cipriani P, Marrelli A, *et al.* Efficacy and safety of rituximab treatment in early primary Sjögren's syndrome: A prospective, multi-center, follow-up study. *Arthritis Res Ther* 2013;15:R172.
43. Bowman SJ, Everett CC, O'Dwyer JL, *et al.* Randomized controlled trial of rituximab and cost-effectiveness analysis in treating fatigue and oral dryness in primary Sjögren's syndrome. *Arthritis Rheumatol* 2017;69:1440-50.
44. Devauchelle-Pensec V, Mariette X, Jousse-Joulin S, *et al.* Treatment of primary Sjögren's syndrome with rituximab – a randomized trial. *Ann Intern Med* 2014;160:233-42.
45. Van Nimwegen JF, Mossel E, Van Zuiden GS, *et al.* Abatacept treatment for patients with early active primary Sjögren's syndrome: a single-centre, randomised, double-blind, placebo-controlled, phase 3 trial (ASAP-III) study. *Lancet Rheumatol* 2020;9913:1-11.
46. Baer AN, Gottenberg J-E, St. Clair EW, *et al.* Efficacy and safety of abatacept in active primary Sjögren's syndrome: results of a phase III, randomised, placebo-controlled trial. *Ann Rheum Dis* 2020; published online first: 9 November 2020.
47. Van den Hoogen LL, Van Laar JM. Targeted therapies in systemic sclerosis, myositis, antiphospholipid syndrome, and Sjögren's syndrome. *Best Pract Res Clin Rheumatol* 2020;34:101485.
48. Fisher BA, Szanto A, Ng WF, *et al.* Assessment of anti-CD40 antibody iscalimab in patients with primary Sjögren's syndrome: a multicenter, randomised, double-blind, placebo-controlled, proof of concept study. *Lancet Rheumatol* 2020;2:e142-52.
49. Dörner T, Posch MG, Li Y, *et al.* Treatment of primary Sjögren's syndrome with ianalumab (VAY736) targeting B cells by BAFF receptor blockade coupled with enhanced, antibody-dependent cellular cytotoxicity. *Ann Rheum Dis* 2019;78:641-7.

50. Radstake TRDJ, Van der Heijden EHM, Moret FM, *et al.* Clinical efficacy of leflunomide/hydroxychloroquine combination therapy in patients with primary Sjögren's syndrome: results of a placebo-controlled double-blind randomized clinical trial [abstract]. *Arthritis Rheumatol* 2018;70:suppl110.
51. Rischmueller M, Tieu J, Lester S. Primary Sjögren's syndrome. *Best Pract Res Clin Rheumatol* 2016;30:189-220.
52. Mossel E, Van Nimwegen JF, Stel AJ, *et al.* Clinical phenotyping of primary Sjögren's patients using salivary gland ultrasonography – data from the RESULT cohort. *J Rheumatol* 2020; published online first: 1 October 2020.

Aim and outline of this thesis

AIM

The overall aim of the research described in this PhD thesis was to assess the potential role of salivary gland ultrasonography (SGUS) in patients diagnosed with or suspected for primary Sjögren's syndrome (pSS).

OUTLINE

In **Chapter 2A**, the validity of salivary gland ultrasonography was assessed by comparing SGUS with parotid gland and labial gland biopsies, sialometry, anti-SSA/Ro antibody status and classification criteria in patients clinically suspected with pSS.

In **Chapter 2B**, we comment on the letter to the editor that was received on our article (Chapter 2A), with as main message that the parotid gland should not be ignored and that it is an important organ in the evaluation of Sjögren's syndrome.

In **Chapter 3**, it was assessed whether ultrasonographic scoring of (i) both parotid and submandibular glands and (ii) all five components of the Hoyer scoring system,²⁸ are needed for classifying patients as pSS.

In **Chapter 4**, it was assessed how ultrasonography, histopathology and salivary secretion, all from the same parotid gland, are associated with each other in pSS patients, who are classified according to the ACR-EULAR criteria.

In **Chapter 5**, the incorporation of SGUS into the ACR-EULAR criteria was investigated. Both, addition and replacement of current items, were evaluated in a large cohort of patients clinically suspected of pSS.

In **Chapter 6**, our REgistry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) cohort was introduced. In a population of classified pSS patients, salivary gland ultrasound (SGUS) abnormalities were investigated in relation to clinical phenotype and patients characteristics, disease activity and disease damage in patients with pSS.

In **Chapter 7**, we comment on an article about the effect of Rituximab on salivary gland ultrasonography in pSS and stress the importance of the evaluation of baseline histopathological characteristics of the parotid glands in determining a personalized medicine approach.

In **Chapter 8**, the efficacy and safety of subcutaneous injections with Abatacept was investigated in a randomized placebo-controlled phase III trial of patients with pSS. International accepted outcome measurements were used as outcome measurements. In addition, SGUS evaluations are currently performed to analyze whether Abatacept treatment leads to a reduction in ultrasonographic abnormalities.

In **Chapter 9**, a review article, an overview is given about recent insights regarding to the potential role of imaging modalities for diagnosing patients with pSS.

In **Chapter 10**, the conclusions of this thesis are discussed in more detail and future perspectives about the use of SGUS in (suspected) pSS are addressed.

CHAPTER 2^A

Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome

Esther Mossel^{1*}, Konstantina Delli^{2*}, Jolien F van Nimwegen¹, Alja J Stel¹, Frans GM Kroese¹, Fred KL Spijkervet², Arjan Vissink², Suzanne Arends¹, Hendrika Bootsma¹, on behalf of the EULAR US-pSS study group

** both authors contributed equally*

¹ Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

² Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Edited version of Ann Rheum Dis 2017;76:1883 – 1889.

Abstract

Objective

To assess the validity of ultrasound of major salivary glands (SGUS) compared with parotid and labial gland biopsies, sialometry, anti-SSA/Ro antibody status and classification criteria in patients clinically suspected with primary Sjögren's syndrome (pSS).

Methods

103 consecutive outpatients with clinically suspected pSS underwent SGUS. Parenchymal echogenicity, homogeneity, hypoechogenic areas, hyperechogenic reflections and clearness of salivary gland border were scored according to the Hocevar scoring system. Total ultrasound score was calculated as the sum of these domains (range 0-48).

Results

Absolute agreement between SGUS and parotid (83%) and labial (79%) gland biopsy outcome was good. Negative SGUS predicts negative parotid gland biopsy, and positive SGUS predicts positive labial gland biopsy.

Compared with the American European Consensus Group (AECG) classification, SGUS showed an absolute agreement of 82%, sensitivity of 71% and specificity of 92%. Compared with the American College of Rheumatology (ACR) classification, absolute agreement was 86%, sensitivity was 77% and specificity was 92%. Compared with the ACR-European League Against Rheumatism (EULAR) classification, absolute agreement was 80%, sensitivity was 67% and specificity was 94%. Positive SGUS predicts classification, but negative SGUS does not exclude classification. The combination of positive SGUS with presence of anti-SSA/Ro antibodies or negative SGUS with absence of anti-SSA/Ro antibodies showed a high predictive value for classification as pSS or non-pSS.

Conclusion

In our prospective inception cohort study derived from daily clinical practice, absolute agreement between SGUS and salivary gland biopsies was slightly higher for parotid compared with labial gland biopsies. The combination of positive SGUS and presence of anti-SSA/Ro antibodies highly predicts classification according to the AECG, ACR and ACR-EULAR classification criteria.

Introduction

Primary Sjögren's syndrome (pSS) is a chronic, systemic autoimmune disease characterised by inflammation of the exocrine glands, with an estimated prevalence of 0.05% in the general population.¹ Most patients with pSS suffer from xerostomia, keratoconjunctivitis sicca and extreme fatigue.² In addition, different extraglandular manifestations, most frequently arthralgia, arthritis and myalgia, may be present.²

Currently, multiple criteria sets are available for the classification of pSS. In 2002, the American European Consensus Group (AECG) criteria were developed, and although not endorsed by the American College of Rheumatology (ACR) or European League Against Rheumatism (EULAR) these are yet most commonly used in daily clinical practice.^{3,4} In 2012, the ACR criteria were developed and provisionally approved by the ACR,⁵ but these criteria were not well received by many Sjögren's syndrome (SS) experts.⁶ In order to develop international consensus on classification criteria, the ACR-EULAR criteria were recently introduced, endorsed by both EULAR and ACR.^{4,7} In all three sets, salivary gland biopsies and presence of anti-SSA/Ro antibodies play a significant role in classifying patients as pSS.³⁻⁵ In the AECG and ACR-EULAR criteria, salivary gland involvement is also assessed by unstimulated whole saliva flow (UWS).^{3,4}

Ultrasound of major salivary glands (SGUS) is an upcoming diagnostic method to assess involvement of major salivary glands in pSS.^{8,9} SGUS is well tolerated, non-invasive, inexpensive, non-irradiating and widely available in the rheumatological outpatient clinics, but its reliability depends greatly on its operator. A recent meta-analysis assessing the diagnostic properties of SGUS in pSS reported a pooled sensitivity of 69% and specificity of 92%. This meta-analysis also revealed a large clinical and methodological heterogeneity between studies, which hampered interpretation of pooled outcomes and influenced the results reported in the various studies.⁸ Thus, the possible role of SGUS in the diagnosis of pSS remains unclear.⁸⁻¹⁰

This study assesses the validity of SGUS compared with parotid and labial gland biopsies, sialometry, anti-SSA/Ro antibody status and classification criteria in patients clinically suspected with pSS.

Patients and Methods

Patients

The present cross-sectional study is based on prospective data from the multidisciplinary Sjögren's expertise centre in the University Medical Center Groningen, a tertiary referral centre. One hundred and ten consecutive patients clinically suspected with pSS, who underwent SGUS as part of the diagnostic work-up and were over 18 years of age, were included. All patients underwent the diagnostic work-up. All domains of the AECG, ACR and ACR-EULAR criteria were assessed, including parotid and/or labial gland biopsy.

Primary assessment: ultrasonography

All patients were examined with the same ultrasonographic scanner (Esaote MyLab Seven, Genova, Italy), equipped with a high-resolution linear scanner (4-13MHz). Patients were examined in supine position with their neck slightly extended and turned away from the examined side.^{11,12} The Hocevar et al.¹² scoring system was used investigating: (1) parenchymal echogenicity compared with the thyroid gland, graded 0-1; (2) homogeneity, graded 0-3; (3) presence of hypoechogenic areas, graded 0-3; (4) hyperechogenic reflections, graded 0-3 in parotid glands and 0-1 in submandibular glands; and (5) clearness of the salivary gland border, graded 0-3, in both parotid and submandibular salivary glands. Total ultrasound score was the sum of these five domains and can range from 0 to 48.¹²

Other assessments

Parotid and/or labial gland biopsies were considered positive if the focus score (defined as the number of mononuclear infiltrates containing ≥ 50 lymphocytes/4mm² of glandular tissue) was ≥ 1 .¹³⁻¹⁵ UWS was evaluated by measuring the saliva production in 15 minutes.³ UWS ≤ 1.5 mL/15min was considered abnormal.¹⁶ Serum levels of anti-SSA/Ro and anti-SSB/La antibodies were assessed with ELISA tests.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics V.23. Descriptive parameters were expressed as number of patients (%) or mean (SD).

Receiver operating characteristic analysis was performed to determine the accuracy of SGUS to predict parotid or labial gland biopsies, UWS, anti-SSA/Ro antibody status and classification as pSS. Area under the curve (AUC) was interpreted as no discrimination (0–0.5), poor accuracy (0.5–0.7), fair (0.7–0.8), good (0.8–0.9) or excellent (0.9–1.0).¹⁷ The optimal cut-off point for SGUS positivity was determined according to the highest combination of sensitivity and specificity.

The percentage of absolute agreement between SGUS outcome and parotid or labial gland biopsies, UWS, anti-SSA/Ro antibody status and classification according to the classification criteria was determined. The association between ultrasound and UWS was analysed using Spearman correlation coefficient (ρ), and interpreted as poor agreement (0.0–0.2), fair (0.2–0.4), moderate (0.4–0.6), good (0.6–0.8) or excellent (0.8–1.0).¹⁸ Furthermore, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated.

Mann-Whitney U tests were used to evaluate differences in total ultrasound score between patients with: (1) positive versus negative parotid or labial gland biopsies, (2) UWS ≤ 1.5 ml/15 min versus UWS > 1.5 ml/15 min, (3) presence versus absence of anti-SSA/Ro and/or anti-SSB/La antibodies and (4) pSS versus non-pSS according to the classification criteria. p Values < 0.05 were considered statistically significant.

Results

Figure 1 Flow chart of number of patients with available data per analysis. * 43 patients underwent parotid gland as well as labial gland biopsy. SS, Sjögren's syndrome; UWS, unstimulated whole saliva.

A flow chart of inclusion and exclusion of patients and information about the number of patients included in the analyses on salivary gland biopsies, sialometry and anti-SSA/Ro antibody status is presented in figure 1. Of the 103 included patients, the mean age was 50 years (15), 90% were female and the mean total ultrasound score was 15 (10) (supplementary table 1). For research purposes, 43 patients underwent parotid and labial gland biopsies. Of the remaining patients, 35 had a parotid gland biopsy only and 13 had a labial gland biopsy only.

Ultrasound versus salivary gland biopsy

The accuracy of SGUS to predict a parotid gland biopsy outcome was good, with an AUC of 0.849 (95% CI 0.746 to 0.952) and optimal cut-off point of 15 (supplementary table 2). The absolute agreement between SGUS outcome and parotid gland biopsy was 83% (65/78), with a sensitivity of 75% (21/28), specificity of 88% (44/50), PPV of 78% (21/27) and NPV of 86% (44/51) (table 1).

The accuracy of SGUS to predict a labial gland biopsy outcome was good, with an AUC of 0.824 (95% CI 0.714 to 0.934) and optimal cut-off point of 14 (supplementary table 2). The absolute agreement between SGUS outcome and labial gland biopsy was 79% (44/56), with a sensitivity of 72% (21/29), specificity of 85% (23/27), PPV of 84% (21/25) and NPV of 74% (23/31) (table 1).

Total ultrasound score was significantly higher in patients with positive parotid or labial gland biopsies compared with patients with negative parotid or labial gland biopsies ($p < 0.001$; figure 2).

Table 1.

Ultrasound of major salivary glands versus salivary gland biopsy, sialometry and anti-SSA/Ro antibodies status

	Parotid gland biopsy (n=78)	Labial gland biopsy (n=56)	UWS (n=98)	Anti-SSA/Ro antibodies (n=103)
Optimal cut-off point	15	14	15	15
% absolute agreement	83.3% (65/78)	78.6% (44/56)	66.3% (65/98)	81.6% (84/103)
Sensitivity	75.0% (21/28)	72.4% (21/29)	54.5% (30/55)	69.8% (37/53)
Specificity	88.0% (44/50)	85.2% (23/27)	81.4% (35/43)	94.0% (47/50)
PPV	77.8% (21/27)	84.0% (21/25)	78.9% (30/38)	92.5% (37/40)
NPV	86.3% (44/51)	74.2% (23/31)	58.3% (35/60)	74.6% (47/63)

NPV, negative predictive value; PPV, positive predictive value; UWS, unstimulated whole saliva.

Figure 2 Ultrasound total score compared with salivary gland biopsy. (A) Positive versus negative parotid gland biopsy. (B) Positive versus negative labial gland biopsy. Histopathology: positive parotid or labial gland biopsy was defined as focus score ≥ 1 .

Ultrasound versus sialometry

The accuracy of SGUS to predict UWS outcome was poor, with AUC of 0.696 (95% CI 0.593 to 0.799). The absolute agreement between SGUS outcome and UWS was 66% (65/98) (table 1).

Total ultrasound score was significantly higher in patients with UWS $\leq 1.5\text{mL}/15\text{min}$ compared to patients with UWS $>1.5\text{mL}/15\text{min}$ ($p<0.001$; supplementary figure 1A). There was a fair reversed association between total ultrasound score and UWS total flow ($\rho=-0.366$) (supplementary figure 1B).

Ultrasound versus auto-antibodies

The accuracy of SGUS to predict anti-SSA/Ro antibody status was good, with an AUC of 0.803 (95% CI 0.711 to 0.894). The absolute agreement between SGUS outcome and anti-SSA/Ro antibody status was 82% (84/103) (table 1).

Total ultrasound score was significantly higher in patients with anti-SSA/Ro antibodies compared to patients without anti-SSA/Ro antibodies ($p<0.001$; supplementary figure 2).

Ultrasound versus classification criteria

For the following analyses, SGUS is compared with classification criteria with the outcome of parotid gland biopsy considered as an item of these criteria. For the comparison of SGUS with classification criteria with the outcome of labial gland biopsy considered as an item, see table 2, supplementary table 3 and figure 3.

The accuracy of SGUS to predict AECG classification was good, with an AUC of 0.826 (95% CI 0.735 to 0.918) and optimal cut-off point of 15 (supplementary table 3). The absolute agreement between SGUS outcome and AECG classification was 82% (80/97), with a sensitivity of 71% (32/45), specificity of 92% (48/52), PPV of 89% (32/36) and NPV of 79% (48/61) (table 2).

The accuracy of SGUS to predict ACR classification was good, with an AUC of 0.862 (95% CI 0.777 to 0.947) and optimal cut-off point of 15 (supplementary table 3). The absolute agreement between SGUS outcome and ACR classification was 86% (83/97), with a sensitivity of 77% (34/44), specificity of 92% (49/53), PPV of 89% (34/38) and NPV of 83% (49/59) (table 2).

The accuracy of SGUS to predict ACR-EULAR classification was good, with an AUC of 0.802 (95% CI 0.710 to 0.894) and optimal cut-off point of 15 (supplementary table 3). The absolute agreement between SGUS outcome and ACR-EULAR classification was 80% (79/99), with a sensitivity of 67% (35/52), specificity of 94% (44/47), PPV of 92% (35/38) and NPV of 72% (44/61) (table 2).

Total ultrasound score was significantly higher in pSS versus non-pSS according to the classification criteria ($p<0.001$; figure 3).

Table 2.

Ultrasound of major salivary glands versus classification criteria

	AECG	
	(parotid gland biopsy)	(labial gland biopsy)
Optimal cut-off point	15/16	15/16
% absolute agreement	82.4% (80/97)	79.2% (76/96)
Sensitivity	71.1% (32/45)	67.3% (33/49)
Specificity	92.3% (48/52)	91.4% (43/47)
PPV	88.9% (32/36)	89.2% (33/37)
NPV	78.7% (48/61)	72.9% (43/59)

	ACR	
	(parotid gland biopsy)	(labial gland biopsy)
Optimal cut-off point	15	15
% absolute agreement	85.6% (83/97)	81.7% (76/93)
Sensitivity	77.3% (34/44)	69.8% (30/43)
Specificity	92.4% (49/53)	92.0% (46/50)
PPV	89.4% (34/38)	88.2% (30/34)
NPV	83.1% (49/59)	78.0% (46/59)

	ACR-EULAR	
	(parotid gland biopsy)	(labial gland biopsy)
Optimal cut-off point	15	15
% absolute agreement	79.8% (79/99)	75.3% (73/97)
Sensitivity	67.3% (35/52)	61.8% (34/55)
Specificity	93.6% (44/47)	92.9% (39/42)
PPV	92.1% (35/38)	91.9% (34/37)
NPV	72.1% (44/61)	65.0% (39/60)

In the left column parotid gland biopsies are considered an item of the AECG, ACR and ACR-EULAR classification criteria. In the right column labial gland biopsies are considered an item of the AECG, ACR and ACR-EULAR criteria.

ACR, American College of Rheumatology; AECG, American European Consensus Group; EULAR, European League Against Rheumatism; NPV, negative predictive value; PPV, positive predictive value

Predictive value of the combination of SGUS and antibody status

In patients with positive SGUS combined with anti-SSA/Ro antibodies, 78% (14/18) had a positive parotid gland biopsy and 94% (17/18) had a positive labial gland biopsy.

In patients with negative SGUS combined with absence of anti-SSA/Ro antibodies, 93% (37/40) had a negative parotid gland biopsy and 77% (17/22) had a negative labial gland biopsy.

In patients with positive SGUS combined with anti-SSA/Ro antibodies, 94% (32/34) fulfilled the AECG, 97% (34/35) fulfilled the ACR and 97% (35/36) fulfilled the ACR-EULAR criteria, considering the outcome of parotid gland biopsies as an item for classification. Percentages are equal when the outcome of labial gland biopsy is considered as an item for classification (figure 4).

Figure 3 Ultrasound total score compared with classification criteria for pSS. In A-C, when ‘including parotid gland biopsy’ is added, parotid gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. In D-F, when ‘including labial gland biopsy’ is added, labial gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. (A) Positive versus negative of fulfilling AECG criteria, (B) positive versus negative of fulfilling ACR criteria, (C) positive versus negative of fulfilling ACR-EULAR criteria, (D) positive versus negative of fulfilling AECG criteria, (E) positive versus negative of fulfilling ACR criteria and (F) positive versus negative of fulfilling ACR-EULAR criteria. In most patients who underwent only one salivary gland biopsy, classification as pSS or non-pSS could be determined, as the biopsy would not have been decisive for classification. Therefore, ≥90% (93/103) of the patients could be included in the analysis of the agreement between SGUS and the different criteria sets including parotid or labial gland biopsy. ACR, American College of Rheumatology; AECG, American European Consensus Group; EULAR, European League Against Rheumatism; pSS, primary Sjögren’s syndrome; SGUS, ultrasound of major salivary glands.

In patients with negative SGUS combined with absence of anti-SSA/Ro antibodies, 98% (45/46) did not fulfil the AECG, 100% (45/45) did not fulfil the ACR and 98% (44/45) did not fulfil the ACR-EULAR criteria, considering the outcome of parotid gland biopsies as an item for classification (figure 4). Percentages are lower when the outcome of labial gland biopsy is considered as an item for classification: 89% (40/45) did not fulfil the AECG, 93% (43/46) did not fulfil the ACR and 89% (39/44) did not fulfil the ACR-EULAR criteria (figure 4).

Figure 4 Percentage of patients either or not fulfilling the AECG, ACR or ACR-EULAR criteria including only parotid gland biopsy or only labial gland biopsy. When 'parotid gland' is mentioned, parotid gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. When 'labial gland' is mentioned, labial gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. (A) Fulfilling the AECG, ACR or ACR-EULAR classification criteria, with on the left positive SGUS in combination with presence of anti-SSA/Ro antibodies and on the right negative SGUS in combination with absence of anti-SSA/Ro antibodies. (B) NOT fulfilling the AECG, ACR or ACR-EULAR classification criteria, with on the left positive SGUS in combination with presence of anti-SSA/Ro antibodies and on the right negative SGUS in combination with absence of anti-SSA/Ro antibodies. ACR, American College of Rheumatology; AECG, American European Consensus Group; EULAR, European League Against Rheumatism; SGUS, ultrasound of major salivary glands.

Discussion

This study assessed the validity of SGUS in a representative population of patients with clinically suspected pSS. It is the first study that directly compared the validity of SGUS with parotid gland biopsy outcome, and to the best of our knowledge the first that compared the validity of SGUS with the ACR-EULAR criteria.

Since the Hocevar scoring system was the most comprehensive and valid, it was incorporated in our daily clinical practice in order to address all relevant aspects of SGUS. Previously, three studies evaluated the optimal cut-off of the Hocevar scoring system.¹² These studies used different gold standards, viz AECG criteria,¹² labial gland biopsy¹⁹ or AECG and ACR criteria.²⁰ The cut-off points ranged from 15 to 19. In the present study, the optimal cut-off point for the Hocevar score was found to be 15 in almost all analyses. Applying a higher cut-off value, for example, 17 as used by Hocevar et al.,¹² would lead to increased specificity and PPV for parotid and labial gland biopsies, as well as for the AECG, ACR and ACR-EULAR classification criteria, at the expense of sensitivity and NPV (supplementary table 1-2). Furthermore, when comparing the PPV and NPV reported in different studies, it is important to keep in mind that both strongly depend on the prevalence of the disease in the investigated population, while sensitivity and specificity are less influenced.

To date, no consensus has been reached about the optimal scoring system and cut-off point for the ultrasonographic evaluation of the major salivary glands in patients suspected with pSS. Consensus is certainly needed to further elucidate the role of SGUS in the diagnosis of pSS and is prerequisite for being officially included in the classification criteria for pSS. Additionally, it will make the comparison between similar studies easier. SGUS showed good absolute agreement and specificity using salivary gland biopsies as gold standard. Both were slightly higher when parotid gland instead of labial gland biopsies were used, possibly because the parotid gland is included in the SGUS evaluation, whereas the labial gland is not.

Interestingly, the parotid gland biopsies were negative in most patients with a negative SGUS, but the labial gland biopsies were positive in 26% of patients with a negative SGUS. On the other hand, positive SGUS predicts positive labial gland biopsies, while 22% of patients with a positive SGUS had a negative parotid gland biopsy. When comparing the results of parotid and labial gland biopsies, it is important to keep in mind not all patients had both biopsies performed and the populations in which the agreement of SGUS with parotid and labial gland biopsies were evaluated were not the same. However, this is the first study to compare SGUS to parotid as well as labial gland biopsies.

There was a fair reversed association between total ultrasound score and UWS total flow. In other words, patients with more pronounced abnormalities on ultrasound tend to have a reduced UWS production. However, there was also a significant amount of patients with few abnormalities on ultrasound who did have a reduced UWS production. This group may consist of patients with early pSS, where UWS is lowered but no SGUS abnormalities are yet seen or another condition is causing the decrease in UWS. A next step would be to assess whether the performance of the different classification criteria will improve, when UWS is replaced by salivary gland ultrasonography.

The specificity of SGUS was high when anti-SSA/Ro antibody status was used as gold standard, confirming a previous study.²¹

Our findings regarding the sensitivity and specificity of SGUS outcome compared with the classification criteria for pSS are similar with the results described in a recent meta-analysis.⁸ Sensitivity of SGUS was lowest when compared with the recently published ACR-EULAR criteria, when either parotid or labial gland biopsies were considered as an item for classification. It is currently unknown if SGUS is sensitive enough to detect changes in the major salivary glands early in the disease course. In our inception cohort, more patients with low SGUS scores fulfil the ACR-EULAR criteria compared with the AECG and ACR criteria. This may either suggest that patients are classified as pSS at an earlier stage of disease according to the ACR-EULAR criteria or that the ACR-EULAR criteria are too liberal.⁷ Finally, the PPV of SGUS compared with the classification criteria was higher than the NPV. Thus, positive SGUS predicts fulfilment of the AECG, ACR and ACR-EULAR criteria, but negative SGUS does not exclude classification. When applying classification criteria, we should keep in mind that they are developed for research purposes to define homogeneous study groups, although in clinical practice they are often used for diagnostic purposes.²²

In accordance with our findings, Astorri et al.²³ reported that positive SGUS was highly predictive of positive labial gland biopsies. Therefore, one could consider not performing a labial gland biopsy in patients with a positive SGUS. This previous study also showed that negative SGUS was highly predictive of negative labial gland biopsies in patients with sicca symptoms. However, we were unable to confirm this observation with our data, as we found a moderate NPV of SGUS for labial gland biopsies. Therefore, in patients with a negative SGUS the result of labial gland biopsies could not fully be predicted. There are some possible explanations for this discrepancy. None of the 'non-Sjögren sicca patients' in the study of Astorri et al.²³ had a positive labial gland biopsy, while some of our 'non-Sjögren sicca

patients' did have a positive biopsy. These biopsies may be false-positive,²⁴⁻²⁵ or these patients have SS not meeting the classification criteria. Moreover, Astorri et al.²³ did not mention the time interval between SGUS and labial gland biopsy, which may be longer than in our study and a different SGUS scoring system was used.

Astorri et al.²³ stated that labial gland biopsies should not be performed in extractable nuclear antigen (ENA)-negative patients with negative SGUS, unless there are other strong clinical indications for SS. Based on our data, we cannot support this conclusion. It is well established that negative serology occurs in 10%-50% of patients with pSS and correlates with milder disease.²⁶⁻²⁸ Interestingly, when labial gland biopsies are considered an item of the classification criteria, 7%-11% of our patients with the combination of negative SGUS and absence of anti-SSA/Ro antibodies were classified as pSS according to the different criteria sets. In these anti-SSA/Ro negative patients, positive biopsies are decisive for classification.²⁹ Thus, ultrasound cannot fully replace salivary gland biopsies as there is a risk of underdiagnosing serological negative patients. Consequently, we recommend that physicians should still consider performing biopsies in patients with absence of anti-SSA/Ro antibodies and negative SGUS, especially when gland function (eg, abnormal Schirmer or UWS) is impaired or when there are other signs and symptoms pointing to pSS.

Strikingly, almost all of our patients with both a positive SGUS and presence of anti-SSA/Ro antibodies fulfilled the classification criteria for pSS. In this patient group, physicians could consider to skip the salivary gland biopsy as the combination of positive SGUS and presence of anti-SSA/Ro antibodies is already highly suggestive of pSS. These rather interesting results are to be confirmed in other cohorts.

The main strength of our study is that consecutive patients clinically suspected with pSS were included. Thus, the study population clearly represents the clinical circumstances in daily clinical practice. Moreover, our Sjögren's expertise centre is one of the few centres in which both parotid and labial gland biopsies can be performed.¹⁴ Having access to parotid gland biopsies has several advantages, in particular that repeated biopsies of the same parotid gland can be performed (eg, for monitoring treatment efficacy) and mucosa-associated lymphoid tissue (MALT)-lymphoma might be identified at an earlier stage.¹⁴

In conclusion, in our prospective inception cohort study derived from daily clinical practice, absolute agreement between SGUS and salivary gland biopsies was good, although slightly higher for parotid compared with labial gland biopsies.

1. Positive SGUS predicts classification according to the AECG, ACR and ACR-EULAR classification criteria, but negative SGUS does not exclude classification.
2. Positive SGUS in combination with presence of anti-SSA/Ro antibodies highly predicts classification according to the AECG, ACR and ACR-EULAR criteria. The combination of negative SGUS and absence of anti-SSA/Ro antibodies highly excludes classification when parotid gland biopsy outcome is considered as an item for classification, but when the outcome of labial gland biopsy is considered as an criteria item, combining negative SGUS with absence of anti-SSA/Ro antibodies does not exclude classification.

Acknowledgements

The authors would like to acknowledge the participating patients.

Part of this study has been presented at the ACR 2016 in Washington and the corresponding abstract has been published as a conference abstract.

References

1. Qin B, Wang J, Yang Z, *et al.* Epidemiology of primary Sjögren's syndrome: a systematic review and meta-analysis. *Ann Rheum Dis* 2015;74:1983-9.
2. Fox RI. Sjögren's syndrome. *Lancet* 2005;366:321-31.
3. 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.
4. Shiboski CH, Shiboski SC, Seror R, *et al.* American College of Rheumatology/ European League Against Rheumatism classification criteria for primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:9-16.
5. Shiboski SC, Shiboski CH, Criswell LA, *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.
6. Vitali C, Bootsma H, Bowman SJ, *et al.* Classification criteria for Sjögren's syndrome: we actually need to definitively resolve the long debate on the issue. *Ann Rheum Dis* 2013;72:476-8.
7. Vissink A, Bootsma H. Refining the classification criteria for primary Sjögren's syndrome. *Nat Rev Rheumatol* 2017;13:10-2.
8. Delli K, Dijkstra PU, Stel AJ, *et al.* Diagnostic properties of ultrasound of major salivary glands in Sjögren's syndrome: a meta-analysis. *Oral Dis* 2015;21:792-800.
9. Jousse-Joulin S, Milic V, Jonsson MV, *et al.* Is salivary gland ultrasonography a useful tool in Sjögren's syndrome? A systematic review. *Rheumatology (Oxford)* 2016;55:789-800.
10. 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;53:1604-07.
11. Theander E, Mandl T. Primary Sjögren's syndrome: diagnostic and prognostic value of salivary gland ultrasonography using a simplified scoring system. *Arthritis Care Res* 2014;66:1102-7.
12. Hocevar A, Ambrozic A, Rozman B, *et al.* Ultrasonographic changes of major salivary glands in primary Sjögren's syndrome. Diagnostic value of a novel scoring system. *Rheumatology (Oxford)* 2005;44:768-72.

13. Chisholm DM, Mason DK. Labial salivary gland biopsy in Sjögren's disease. *J Clin Pathol* 1968;21:656-60.
14. Pijpe J, Kalk WWI, van der Wal JE, *et al.* Parotid gland biopsy compared with labial biopsy in the diagnosis of patients with primary Sjögren's syndrome. *Rheumatology (Oxford)* 2007;46:335-41.
15. Greenspan JS, Daniels TE, Talal N, *et al.* The histopathology of Sjögren's syndrome in labial salivary gland biopsies. *Oral Surgery, Oral Med Oral Pathol* 1974;37:217-29.
16. Jensen SB, Vissink A. Salivary Gland Dysfunction and Xerostomia in Sjögren's Syndrome. *Oral Maxillofac Surg Clin NA* 2014;26:35-53.
17. Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. Third edition 2013.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
19. 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.
20. 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.
21. 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. *Rheumatol (Oxford)* 2015;54:2198-204.
22. Bootsma H, Spijkervet FKL, Kroese FKL, *et al.* Toward new classification criteria for Sjögren's syndrome? *Arthritis Rheum* 2013;65:21-3.
23. Astorri E, Sutcliffe N, Richards PS, *et al.* Ultrasound of the salivary glands is a strong predictor of labial gland biopsy histopathology in patients with sicca symptoms. *J Oral Pathol Med* 2015;44:450-4.
24. Radfar L, Kleiner DE, Fox PC, *et al.* Prevalence and clinical significance of lymphocytic foci in minor salivary glands of healthy volunteers. *Arthritis Rheum* 2002;47:520-4.
25. Segerberg-Kontinen M, Kontinen YT, Bergroth V. Focus score in the diagnosis of Sjögren's syndrome. *Scand J Rheumatol Suppl* 1986;61:47-51.
26. Peri Y, Agmon-Levin N, Theodor E, *et al.* Sjögren's syndrome, the old and the new. *Best Pract Res Clin Rheumatol* 2012;26:105-17.
27. Ramos-Casals M, Solans R, Rosas J, *et al.* Primary Sjögren Syndrome in Spain. *Medicine (Baltimore)* 2008;87:210-9.

28. Oni C, Mitchell S, James K, *et al.* Eligibility for clinical trials in primary Sjögren's syndrome: lessons from the UK Primary Sjögren's Syndrome Registry. *Rheumatology (Oxford)* 2016;55:544-52.
29. Barone F, Campos J, Bowman S, *et al.* The value of histopathological examination of salivary gland biopsies in diagnosis, prognosis and treatment of Sjögren's syndrome. *Swiss Med Wkly* 2015;145:w14168.

Supplementary Table 1.

Characteristics of the study population.

Total group (n=103)				
Age (years)	50 (15)			
Females	93 (90%)			
Ultrasound total score	15 (10)			
Biopsy with focus score ≥ 1				
<i>Parotid gland biopsy</i>	28 (36%)			
<i>Labial gland biopsy</i>	29 (52%)			
Anti-SSA antibodies	53 (52%)			
UWS ≤ 1.5 ml/15min	55 (56%)			
OSS ≥ 5	30 (29%)			
Schirmer ≤ 5 mm/5min	74 (72%)			
	AECG – incl. parotid gland biopsy (n=97)		AECG – incl. labial gland biopsy (n=96)	
	Positive	Negative	Positive	Negative
Age (years)	50 (16)	50 (14)	51 (15)	49 (15)
Females	42 (93%)	46 (88%)	47 (96%)	41 (87%)
Ultrasound total score	22 (11)	10 (5)	22 (11)	10 (5)
Biopsy with focus score ≥ 1	25 (76%)	3 (7%)	26 (84%)	3 (12%)
Anti-SSA antibodies	44 (98%)	5 (10%)	44 (90%)	5 (11%)
UWS ≤ 1.5 ml/15min	33 (77%)	17 (35%)	34 (72%)	15 (34%)
OSS ≥ 5	25 (56%)	5 (10%)	26 (53%)	4 (9%)
Schirmer ≤ 5 mm/5min	40 (89%)	30 (58%)	42 (86%)	27 (57%)

	ACR – incl. parotid gland biopsy (n=97)		ACR – incl. labial gland biopsy (n=93)	
	Positive	Negative	Positive	Negative
Age (years)	48 (16)	51 (14)	50 (14)	51 (14)
Females	40 (91%)	47 (89%)	39 (91%)	44 (88%)
Ultrasound total score	23 (11)	9 (4)	22 (11)	9 (4)
Biopsy with focus score ≥ 1	25 (78%)	3 (7%)	24 (83%)	5 (19%)
Anti-SSA antibodies	44 (100%)	5 (9%)	40 (93%)	4 (8%)
UWS ≤ 1.5 ml/15min	29 (71%)	23 (45%)	27 (68%)	22 (46%)
OSS ≥ 5	25 (57%)	4 (8%)	25 (58%)	3 (6%)
Schirmer ≤ 5 mm/5min	38 (86%)	32 (60%)	36 (84%)	30 (60%)

	ACR-EULAR – incl. parotid gland biopsy (n=99)		ACR-EULAR – incl. labial gland biopsy (n=97)	
	Positive	Negative	Positive	Negative
Age (years)	50 (16)	50 (14)	51 (15)	50 (14)
Females	48 (92%)	42 (89%)	51 (93%)	37 (88%)
Ultrasound total score	21 (12)	9 (4)	20 (12)	9 (4)
Biopsy with focus score ≥ 1	26 (72%)	2 (5%)	26 (79%)	3 (13%)
Anti-SSA antibodies	51 (98%)	1 (2%)	50 (91%)	1 (2%)
UWS ≤ 1.5 ml/15min	37 (76%)	16 (36%)	38 (73%)	14 (35%)
OSS ≥ 5	25 (48%)	4 (9%)	26 (47%)	3 (7%)
Schirmer ≤ 5 mm/5min	45 (87%)	26 (55%)	48 (87%)	23 (55%)

Values are presented as number of patients (%) or mean (SD). When ‘including parotid gland biopsy’ is added, parotid gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. When ‘including labial gland biopsy’ is added, labial gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. UWS = unstimulated whole saliva; OSS = ocular staining score.

Supplementary Table 2.

Different cut-off points for ultrasound total score compared to salivary gland biopsy.

Cut-off point	17	16	15	14	13
Parotid gland biopsy (n=78)					
% absolute agreement	83.3% (65/78)	83.3% (65/78)	83.3% (65/78)	82.1% (64/78)	78.2% (61/78)
Sensitivity	71.4% (20/28)	75.0% (21/28)	75.0% (21/28)	82.1% (23/28)	82.1% (23/28)
Specificity	90.0% (45/50)	88.0% (44/50)	88.0% (44/50)	82.0% (41/50)	76.0% (38/50)
PPV	80.0% (20/25)	77.8% (21/27)	77.8% (21/27)	71.9% (23/32)	65.7% (23/35)
NPV	84.9% (45/53)	86.3% (44/51)	86.3% (44/51)	89.1% (41/46)	88.4% (38/43)
Labial gland biopsy (n=56)					
% absolute agreement	69.6% (39/56)	71.4% (40/56)	71.4% (40/56)	78.6% (44/56)	76.8% (43/56)
Sensitivity	48.3% (14/29)	55.2% (16/29)	55.2% (16/29)	72.4% (21/29)	72.4% (21/29)
Specificity	92.6% (25/27)	88.9% (24/27)	88.9% (24/27)	85.2% (23/27)	81.4% (22/27)
PPV	87.5% (14/16)	84.2% (16/19)	84.2% (16/19)	84.0% (21/25)	80.8% (21/26)
NPV	62.5% (25/40)	64.9% (24/37)	64.9% (24/37)	74.2% (23/31)	73.3% (22/30)

PPV = positive predictive value; NPV = negative predictive value.

The bold columns show the optimal cut-off point(s).

Cut-off ≥ 15 and cut-off ≥ 16 were found to have identical diagnostic accuracy for parotid gland biopsy.

Supplementary Table 3.

Different cut-off points of ultrasound total score compared to classification criteria.

Cut-off point	17	16	15	14	13
AECG criteria (incl. parotid gland biopsy)					
% absolute agreement	81.4% (79/97)	82.4% (80/97)	82.4% (80/97)	81.4% (79/97)	78.4% (76/97)
Sensitivity	66.7% (30/45)	71.1% (32/45)	71.1% (32/45)	77.8% (35/45)	77.8% (35/45)
Specificity	94.2% (49/52)	92.3% (48/52)	92.3% (48/52)	84.6% (44/52)	78.8% (41/52)
PPV	90.9% (30/33)	88.9% (32/36)	88.9% (32/36)	81.4% (35/43)	76.1% (35/46)
NPV	76.6% (49/64)	78.7% (48/61)	78.7% (48/61)	81.4% (44/54)	80.3% (41/51)
AECG criteria (incl. labial gland biopsy)					
% absolute agreement	78.1% (75/96)	79.2% (76/96)	79.2% (76/96)	79.2% (76/96)	76.0% (73/96)
Sensitivity	63.3% (31/49)	67.3% (33/49)	67.3% (33/49)	73.4% (36/49)	73.4% (36/49)
Specificity	93.6% (44/47)	91.4% (43/47)	91.4% (43/47)	85.1% (40/47)	78.7% (37/47)
PPV	91.2% (31/34)	89.2% (33/37)	89.2% (33/37)	83.7% (36/43)	78.3% (36/46)
NPV	71.0% (44/62)	72.9% (43/59)	72.9% (43/59)	75.4% (40/53)	74.0% (37/50)

Cut-off point	17	16	15	14	13
ACR criteria (incl. parotid gland biopsy)					
% absolute agreement	82.4% (80/97)	84.5% (82/97)	85.6% (83/97)	83.5% (81/97)	80.4% (78/97)
Sensitivity	68.2% (30/44)	75.0% (33/44)	77.3% (34/44)	81.8% (36/44)	81.8% (36/44)
Specificity	94.3% (50/53)	92.4% (49/53)	92.4% (49/53)	84.9% (45/53)	79.2% (42/53)
PPV	90.9% (30/33)	89.2% (33/37)	89.4% (34/38)	81.8% (36/44)	76.6% (36/47)
NPV	78.1% (50/64)	81.7% (49/60)	83.1% (49/59)	84.9% (45/53)	84.0% (42/50)
ACR criteria (incl. labial gland biopsy)					
% absolute agreement	78.4% (73/93)	80.6% (75/93)	81.7% (76/93)	80.6% (75/93)	77.4% (72/93)
Sensitivity	60.4% (26/43)	67.4% (29/43)	69.8% (30/43)	76.7% (33/43)	76.7% (33/43)
Specificity	94.0% (47/50)	92.0% (46/50)	92.0% (46/50)	84.0% (42/50)	78.0% (39/50)
PPV	89.7% (26/29)	87.9% (29/33)	88.2% (30/34)	80.4% (33/41)	75.0% (33/44)
NPV	73.4% (47/64)	76.7% (46/60)	78.0% (46/59)	80.8% (42/52)	79.6% (39/49)
ACR-EULAR (incl. parotid gland biopsy)					
% absolute agreement	77.8% (77/99)	78.8% (78/99)	79.8% (79/99)	78.8% (78/99)	75.8% (75/99)
Sensitivity	61.5% (32/52)	65.4% (34/52)	67.3% (35/52)	73.1% (38/52)	73.1% (38/52)
Specificity	95.7% (45/47)	93.6% (44/47)	93.6% (44/47)	85.1% (40/47)	78.7% (37/47)
PPV	94.1% (32/34)	91.9% (34/37)	92.1% (35/38)	84.4% (38/45)	79.2% (38/48)
NPV	69.2% (45/65)	71.0% (44/62)	72.1% (44/61)	74.1% (40/54)	72.5% (37/51)

Cut-off point	17	16	15	14	13
ACR-EULAR (incl. labial gland biopsy)					
% absolute agreement	73.2% (71/97)	74.2% (72/97)	75.3% (73/97)	75.3% (73/97)	72.2% (70/97)
Sensitivity	56.4% (31/55)	60.0% (33/55)	61.8% (34/55)	67.3% (37/55)	67.3% (37/55)
Specificity	95.2% (40/42)	92.9% (39/42)	92.9% (39/42)	85.7% (36/42)	78.6% (33/42)
PPV	93.9% (31/33)	91.7% (33/36)	91.9% (34/37)	86.0% (37/43)	80.4% (37/46)
NPV	62.5% (40/64)	63.9% (39/61)	65.0% (39/60)	66.7% (36/54)	64.7% (33/51)

PPV = positive predictive value; NPV = negative predictive value.

The bold columns show the optimal cut-off point(s). Cut-off ≥ 15 and cut-off ≥ 16 were found to have identical diagnostic accuracy for AECG criteria including either parotid gland biopsy or labial gland biopsy. When 'including parotid gland biopsy' is added, parotid gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria. When 'including labial gland biopsy' is added, labial gland biopsy outcome is considered an item of the AECG, ACR and ACR-EULAR criteria.

Supplementary Figure 1. Ultrasound total score compared to sialometry. A. UWS flow $\leq 1.5\text{mL}/15\text{min}$ versus $> 1.5\text{mL}/15\text{min}$; B. association with UWS total flow mL/min; UWS, unstimulated whole saliva.

Supplementary Figure 2. Ultrasound total score compared to presence or absence of anti-SSA/Ro and anti-SSB/La antibodies. In patients with absence of anti-SSA/Ro antibodies, anti-SSB/La antibodies were also absent. In 29 patients anti-SSB/La antibodies were present.

CHAPTER 2^B

The parotid gland connection: ultrasound and biopsies in primary Sjögren's syndrome

Esther Mossel¹, Konstantina Delli², Jolien F van Nimwegen¹,
Alja J Stel¹, Erlin A Haacke^{1,3}, Frans G M Kroese¹, Fred KL Spijkervet²,
Arjan Vissink², Suzanne Arends¹, Hendrika Bootsma¹

¹ Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

² Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

³ Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Edited version of Ann Rheum Dis 2018;77:e38.

We would like to thank Alegria et al.¹ for their letter to the editor commenting on our recent publication entitled 'Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome'.² Our study was the first that directly compared the validity of ultrasound of major salivary glands (SGUS) with parotid gland biopsy outcome. We showed that the agreement between SGUS and parotid as well as labial gland biopsies was good but was slightly higher for the former.

As noted by Alegria et al.,¹ we found different results for the positive predictive value (PPV) and negative predictive value (NPV) between SGUS versus parotid and SGUS versus labial gland biopsies.² Although these observations may indeed be caused by discordance between the parotid and labial gland biopsy outcome, there are other factors that may equally well contribute to this discrepancy. First, the parotid glands are assessed during the SGUS examination and included in the scoring, whereas the labial glands are not.² Second, in 6%-15% of the general population, the labial gland biopsy is positive, while in only 5% of the general population the parotid gland biopsy is positive.³⁻⁶ This implies that when biopsies of parotid and labial glands are taken simultaneously, at least up to 10% of the biopsies may be discordant. Third, the patient cohort in which SGUS was compared with parotid gland biopsies was not exactly the same as the patient cohort in which SGUS was compared with labial gland biopsies. However, when only taking into account the 43 patients who underwent both a parotid and labial gland biopsy, there still was a discrepancy between the predictive values of parotid and labial gland biopsies. In this group of 43 patients, compared with parotid gland biopsies, SGUS showed a PPV of 64% (7/11) and NPV of 88% (28/32), and when compared with labial gland biopsies, PPV was 73% (11/15) and NPV was 75% (21/28).

Alegria et al.¹ requested more information about the group of 43 patients who had both a parotid and labial gland biopsy, that is, the agreement between both biopsies and the correlation between SGUS and the focus score in the parotid and labial gland biopsies. We fully agree with Alegria et al.¹ that it is important to make a direct comparison between parotid and labial salivary gland biopsies and not only to compare both to SGUS. Such a detailed comparison of the results obtained with parotid and labial gland biopsies is currently in progress in a larger cohort of patients. In this study, the focus score of the parotid and labial gland biopsies will be compared, as suggested by Alegria et al.,¹ and the other histopathological characteristics of primary Sjögren's syndrome in both types of biopsies.

In the current study, in the 43 patients who underwent a double biopsy, the correlation between SGUS and parotid gland focus score was $\rho=0.376$ (figure 1A),

Figure 1
 (A) Ultrasound total score compared with focus score in parotid gland biopsies.
 (B) Ultrasound total score compared with focus score in labial gland biopsies.

while the correlation between SGUS and labial gland focus score was $p=0.412$ (figure 1B). Previously, Cornec et al. reported a correlation of $p=0.61$ between SGUS and focus score in labial salivary glands of participants in the TEARS trial.⁷ However, all patients in their cohort were already classified as primary Sjögren's syndrome according to the American European Consensus Group criteria, in contrast to our study, in which patients clinically suspected with pSS were included.^{7,8} This difference in patient selection could be the explanation for the discrepancy in reported correlation.

As Alegria et al.¹ stated, Chisholm et al.⁹ found that minor salivary gland biopsies have a greater specificity compared with major salivary gland biopsies. Whether this study is representative for current medical practice is discussable. First, Chisholm et al.⁹ evaluated major salivary gland biopsies from the submandibular gland, whereas in our study,² parotid gland biopsies were evaluated. Whether the histopathology of the submandibular glands is similar to the parotid gland is unknown, since submandibular salivary glands are not easily accessible for taking incisional biopsies. Second, another important difference is that Chisholm et al.⁹ performed a postmortem study. Thus, the study of Chisholm et al.⁹ cannot be used as a direct comparison. On the other hand, two other studies showed that major salivary gland biopsies performed at least as well as minor salivary gland biopsies in the diagnosis of primary Sjögren's syndrome.^{6,10}

Currently, most groups still prefer to perform labial gland instead of parotid gland biopsies, as taking parotid gland biopsies requires specific surgical expertise.⁶ Therefore, labial gland biopsies are included in the various classification criteria sets and were the main focus of recent guidelines, regarding the histopathological evaluation of Sjögren's syndrome.^{8,11-13} However, there are several advantages of parotid gland biopsies in comparison to labial gland biopsies, including the possibility of identifying mucosa-associated lymphoid tissue lymphoma at an earlier stage, the possibility of performing repeated biopsies of the same gland and allowing a direct gland-specific comparison with other diagnostic methods, like glandular specific saliva and ultrasound.^{6,14,15} Furthermore, in the study of Marx et al.,¹⁰ in five Sjögren's syndrome patients, a diagnosis of lymphoma was made based on the parotid gland biopsies, while the labial gland biopsies all lacked lymphomatous changes. Interestingly, we have shown that the baseline number of CD20⁺B-cells/mm² of parenchyma could serve as a prognostic biomarker to predict rituximab response. Hence, baseline characteristics of the parotid gland biopsy might be our guide to personalized treatment.¹⁶

To conclude, the parotid gland should not be ignored and is an important organ in the evaluation of Sjögren's syndrome.

References

1. Alegria GC, Costa S, Jousse-Joulin S, *et al.* What is the agreement between pathological features of parotid gland and labial salivary gland biopsies? *Ann Rheum Dis* 2018;77:e37.
2. Mossel E, Delli K, van Nimwegen JF, *et al.* Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1883-89.
3. Daniels TE. Labial salivary gland biopsy in Sjögren's syndrome. Assessment as a diagnostic criterion in 362 suspected cases. *Arthritis Rheum* 1984;27:147-56.
4. Segerberg-Konttinen M, Konttinen YT and Bergroth V. Focus score in the diagnosis of Sjögren's syndrome. *Scand J Rheumatol Suppl* 1986;61:47-51.
5. Radfar L, Kleiner DE, Fox PC, *et al.* Prevalence and clinical significance of lymphocytic foci in minor salivary glands of healthy volunteers. *Arthritis Rheum* 2002;47:520-4.
6. Pijpe J, Kalk WWI, van der Wal JE, *et al.* Parotid gland biopsy compared with labial biopsy in the diagnosis of patients with primary Sjögren's syndrome. *Rheumatology (Oxford)* 2007;46:335-41.
7. Cornec D, Jousse-Joulin S, Costa S, *et al.* High-grade salivary-gland involvement, assessed by histology or ultrasonography, is associated with a poor response to a single rituximab course in primary Sjögren's syndrome: Data from the TEARS randomized trial. *PLoS One* 2016;11:1-11.
8. 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.
9. Chisholm DM, Waterhouse JP and Mason DK. Lymphocytic sialadenitis in the major and minor glands: a correlation in postmortem subjects. *J Clin Pathol* 1970;23:690-4.
10. Marx RE, Hartman KS and Rethman KV. A prospective study comparing incisional labial to incisional parotid biopsies in the detection and confirmation of sarcoidosis, Sjögren's disease, sialosis and lymphoma. *J Rheumatol* 1988;15:621-9.
11. Shiboski SC, Shiboski CH, Criswell LA, *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.

12. Shiboski CH, Shiboski SC, Seror R, *et al.* 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:9-16.
13. Fisher BA, Jonsson R, Daniels T, *et al.* Standardisation of labial salivary gland histopathology in clinical trials in primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1161-8.
14. Spijkervet FKL, Haacke E, Kroese FGM, *et al.* Parotid gland biopsy, the alternative way to diagnose Sjögren's syndrome. *Rheum Dis Clin North Am* 2016;42:485-99.
15. Delli K, Vissink A and Spijkervet FKL. Salivary gland biopsy for Sjögren's syndrome. *Oral Maxillofac Surg Clin North Am* 2014;26:22-33.
16. Delli K, Haacke EA, Kroese FGM, *et al.* Towards personalised treatment in primary Sjögren's syndrome; baseline parotid histopathology predicts responsiveness to rituximab treatment. *Ann Rheum Dis* 2016;75:1933-8.

CHAPTER 3

Scoring hypoechogenic areas in one parotid and one submandibular gland increases feasibility of ultrasound in primary Sjögren's syndrome

Esther Mossel¹, Suzanne Arends¹, Jolien F van Nimwegen¹, Konstantina Delli², Alja J Stel¹, Frans G M Kroese¹, Fred KL Spijkervet², Arjan Vissink², Hendrika Bootsma¹, on behalf of the EULAR US-pSS study group

¹ Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

² Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Edited version of Ann Rheum Dis 2018;77:556-562.

Abstract

Objective

To assess whether ultrasonographic scoring of (i) both parotid and submandibular salivary glands and (ii) all individual components of the Hocevar scoring system, is needed for classifying patients as primary Sjögren's syndrome (pSS).

Methods

Ultrasound examination of the major salivary glands (SGUS) was performed in 204 consecutive patients clinically suspected (n=171) or diagnosed (n=33) with pSS. Parenchymal echogenicity, homogeneity, hypoechogenic areas, hyperechogenic reflections and salivary gland posterior border were scored in left and right parotid and submandibular glands. Logistic regression analyses were performed to assess which glands and SGUS components contributed significantly to classification as pSS or non-pSS according to the 2016 American College of Rheumatology – European League Against Rheumatism (ACR-EULAR) criteria.

Results

116 (57%) patients were classified as pSS, the remaining as non-pSS. Instead of scoring both sides (area under the curve; AUC=0.856, Nagelkerke $R^2=0.526$), multivariate analysis showed that SGUS scoring of only right (AUC=0.850; $R^2=0.518$) or left (AUC=0.852; $R^2=0.511$) parotid and submandibular glands is sufficient to predict ACR-EULAR classification. Moreover, all individual components of the Hocevar scoring system significantly predicted classification. Multivariate analysis showed that parenchymal echogenicity and hypoechogenic areas contributed independently to ACR-EULAR classification (AUC=0.857; $R^2=0.539$). Scoring these components in one parotid and one submandibular gland highly predicted ACR-EULAR classification (AUC=0.855; $R^2=0.539$). Scoring only hypoechogenic areas on one side showed almost similar results (AUC=0.846; $R^2=0.498$).

Conclusion

SGUS examination of parotid and submandibular glands on one side is sufficient to predict classification of patients according to the ACR-EULAR criteria. To further increase feasibility of SGUS in outpatient clinics worldwide, only hypoechogenic areas can be scored.

Introduction

Primary Sjögren's syndrome (pSS) is a chronic, systemic autoimmune disease with inflammation of the exocrine glands being the most common feature, leading to keratoconjunctivitis sicca and a sensation of oral dryness (xerostomia). Extreme fatigue and several extraglandular manifestations are also common.^{1,2}

Ultrasound of major salivary glands (SGUS) is an upcoming imaging technique to assess the involvement of major salivary glands in pSS. It is gaining more acceptance in daily clinical practice, as it is non-irradiating, non-invasive and easy to perform.³⁻⁶

A meta-analysis has recently identified 33 scoring systems used to evaluate the major salivary glands in patients with SS.³ This heterogeneity within the scoring systems was related to the type of salivary glands examined, the ultrasonographic characteristics evaluated and the cut-off points applied. The scoring system of Hocevar et al.⁷ is the most extensive and has previously been applied in several studies.⁶⁻¹⁰ However, this scoring system is time consuming and therefore less feasible in daily clinical practice. In order to become part of the routine in daily clinical practice in the outpatient rheumatologic clinics worldwide, SGUS should be as easy, fast and efficient as possible. Previously, more simple scoring systems have been used,¹¹⁻¹⁴ but none of these systems were constructed on the basis of a multivariate regression model.

According to the Outcome Measures in Rheumatology Clinical Trials (OMERACT) filter, three aspects are considered important when it comes to the applicability of a new diagnostic method, that is, truth, discrimination and feasibility.¹⁵ Truth questions whether SGUS measures what is intended. Discrimination evaluates whether SGUS differentiates between situations of interest. Both aspects were previously investigated in our validity and reliability studies.^{6,10} Feasibility examines whether SGUS can be applied easily, given constraints in time, money, and interpretability. Since time is precious in daily clinical practice, the feasibility of SGUS could be increased by reducing the number of glands and/or the number of components scored. Therefore, the aim of this study was to assess whether ultrasonographic scoring of (i) both parotid and submandibular glands and (ii) all five components of the Hocevar scoring system,⁷ is needed for classifying patients as pSS.

Patients and Methods

Patients

In this cross-sectional study derived from our daily clinical practice, consecutive patients clinically suspected or diagnosed with pSS were included. The largest group consisted of tertiary referrals with clinical suspicion of pSS, referred to the University Medical Center Groningen (UMCG) for evaluation at our multidisciplinary Sjögren's expertise centre. The remaining part were participants of our REgistry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) cohort, in which consecutive patients with pSS are included. All patients were ≥ 18 years and underwent a full (diagnostic) work-up, including an ultrasound of the major salivary glands. Patients with an incomplete work-up or an associated systemic autoimmune disease (eg., rheumatoid arthritis, systemic lupus erythematosus) were excluded.

Ultrasonography

The ultrasonographic scanner (Esaote MyLabSeven, Genova, Italy), equipped with a high-resolution linear scanner (4-13 MHz) was used. During the examination, patients were lying in supine position, with their neck slightly extended and head turned towards the opposite side. The scoring system by Hocevar et al.⁷ was used and the following components were investigated in both parotid and submandibular salivary glands: (i) parenchymal echogenicity compared with thyroid gland or when there was co-incident thyroid gland disease compared with surrounding anatomical structure, graded 0-1; (ii) homogeneity, graded 0-3; (iii) presence of hypoechogenic areas, graded 0-3; (iv) hyperechogenic reflections, graded 0-3 in the parotid glands and 0-1 in the submandibular glands and (v) clearness of salivary gland border, graded 0-3. Ultrasonographic examination was performed by three well-trained ultrasonographers (AJS, KD and JFN). These ultrasonographers previously showed good to excellent interobserver reliability and excellent intraobserver reliability on the total Hocevar score.⁶

ACR-EULAR criteria

Patients were classified as pSS or non-pSS according to the recently published 2016 American College of Rheumatology – European League Against Rheumatism (ACR-EULAR) classification criteria.¹⁶ Routinely, all items of these criteria were assessed, including serology and histopathology. To assess fulfilment of the salivary gland histology item, labial as well as parotid glands biopsy results were considered.^{17,18}

Statistical analysis

Descriptive parameters were expressed as number of patients (%) for categorical data and mean (SD) for continuous data. Mann-Whitney U test was used to evaluate the differences in ultrasound scores between pSS and non-pSS patients according to the ACR-EULAR criteria. Associations between different ultrasound scores were analysed using Spearman's correlation coefficient (ρ); ρ was interpreted as poor (0.0-0.2), fair (0.2-0.4), moderate (0.4-0.6), good (0.6-0.8) or excellent (0.8-1.0) association.¹⁹ Receiver operating characteristic (ROC) analysis was performed to determine the accuracy of SGUS to predict classification according to the ACR-EULAR criteria. Area under the curve (AUC) was interpreted as no discrimination (0-0.5), poor (0.5-0.7), fair (0.7-0.8), good (0.8-0.9) or excellent (0.9-1.0) accuracy.²⁰

The Hocevar total score (four glands, five components) was evaluated using univariate logistic regression analysis with ACR-EULAR classification as dependent variable. Univariate and multivariate logistic regression analyses applying the Forward Wald method were performed for ultrasound total score for (i) all separate glands, (ii) all individual components of the Hocevar scoring system and (iii) combining i and ii.⁷ The explained variance was evaluated using Nagelkerke R^2 . IBM SPSS Statistics V.23 (SPSS, Chicago, Illinois, USA) was used to perform the statistical analyses. P values <0.05 were considered to be statistically significant.

Results

Table 1.

Characteristics of the study population.

Characteristics	Total group (n=204)	pSS (n=116)	non-pSS (n=88)
Age, years	52.2 (13.8)	53.5 (12.8)	50.4 (14.8)
Females	174 (85.3%)	100 (86.2%)	74 (84.1%)
Parotid gland biopsy, focus score $\geq 1^*$	64 (40.3%)	61 (68.5%)	3 (4.3%)
Labial gland biopsy, focus score $\geq 1^*$	52 (52.0%)	47 (81.0%)	5 (11.9%)
Anti-SSA/Ro antibodies	106 (52.0%)	104 (89.7%)	2 (2.3%)
UWS ≤ 0.1 ml/min	113 (55.4%)	87 (75.0%)	26 (29.5%)
OSS $\geq 5^\dagger$	63 (31.0%)	55 (47.4%)	8 (9.2%)
Schirmer's test ≤ 5 mm/5 min †	144 (71.3%)	98 (84.5%)	46 (53.4%)
Ultrasound total score	16.9 (11.6)	23.2 (11.5)	8.7 (4.5)
Right parotid and submandibular glands	8.6 (6.1)	11.9 (6.1)	4.4 (2.4)
Left parotid and submandibular glands	8.3 (5.7)	11.3 (5.6)	4.3 (2.5)

Data are expressed as number of patients (%) or mean (SD).

OSS was missing in two patients and Schirmer's test was missing in one patients.

Data were complete for anti-SSA/Ro antibody status, UWS and the ultrasonographic examination.

* A parotid gland biopsy was performed in 159 out of the 204 patients and a labial gland biopsy was performed in 100 out of the 204 patients. In 75 patients, both biopsies were performed. In 20 patients, there was no salivary gland biopsy performed. In these patients, the outcome of a biopsy would not influence classification as pSS or non-pSS according to the ACR-EULAR criteria.

† In at least one eye.

UWS, unstimulated whole saliva; OSS, ocular staining score.

Between October 2014 and February 2017, 224 consecutive patients were evaluated in our multidisciplinary Sjögren's expertise centre. Twenty patients had to be excluded because of incomplete work-up ($n=14$) or presence of an associated systemic autoimmune disease ($n=6$). Of the 204 included patients, mean age was 52 (13.8) years and 85% were female (table 1). This cohort consisted of tertiary referrals with clinical suspicion of pSS ($n=171$), and patients previously diagnosed with pSS from our RESULT cohort ($n=33$). A salivary gland biopsy was available in 184 patients. In the remaining 20 patients, the result of the biopsy would not influence classification as pSS or non-pSS and therefore these patients were not willing to undergo a salivary gland biopsy as part of the routine diagnostic workup. In this cohort, 116 (57%) patients were classified as pSS according to the ACR-EULAR criteria.

Supplementary figure 1 shows ultrasonographic images of normal submandibular and parotid glands and images of submandibular and parotid glands suggestive of pSS. Ultrasound scores of all four glands differed significantly between pSS and non-pSS patients (supplementary figure 2).

Scoring only one parotid and one submandibular gland is sufficient.

In the total group, there were very high correlations between the ultrasound scores of both parotid glands ($p=0.909$), both submandibular glands ($p=0.868$) and between the combination of left parotid and submandibular glands versus the right parotid and submandibular glands ($p=0.926$; figure 1). In other words, patients who scored high on their left parotid gland also scored high on their right parotid gland and the same is true for the submandibular glands.

Correlations between the ultrasound scores of the left parotid and left submandibular gland ($p=0.731$) as well as the right parotid and right submandibular gland ($p=0.734$) were slightly lower (figure 1). In summary, the ultrasound scores of the same type of salivary glands are interchangeable (i.e., right vs. left parotid gland; right vs. left submandibular gland), but this is not always true for the score of the different types of glands (i.e., parotid vs. submandibular gland).

Figure 1 Association of ultrasound total scores between (A) right vs left parotid glands; (B) right vs left submandibular glands; (C) left parotid vs left submandibular glands; (D) right parotid vs right submandibular glands and (E) right parotid and submandibular glands vs left parotid and submandibular glands. Spearman's ρ is shown for (i) total group, (ii) American College of Rheumatology – European League Against Rheumatism (ACR-EULAR) positive group and (iii) ACR-EULAR negative group.

Ultrasound;	Area under the curve;	95% CI
A-D. 4 glands, 5 components	0.856	0.805 -- 0.907
A. 2 glands, 5 components: right	0.850	0.798 -- 0.902
A. 2 glands, 5 components: left	0.852	0.800 -- 0.903
B. 4 glands, 2 components	0.857	0.805 -- 0.909
C. 2 glands, 2 components: right	0.855	0.803 -- 0.906
C. 2 glands, 2 components: left	0.847	0.793 -- 0.900
D. 2 glands, 1 component: right	0.846	0.794 -- 0.899
D. 2 glands, 1 component: left	0.841	0.787 -- 0.895

Figure 2 Receiver operating curves (ROC) (A-D) 4 glands, 5 components* compared with: (A) 2 glands, 5 components*, right and left; (B) 4 glands, 2 components †; (C) 2 glands, 2 components †, right and left and (D) 2 glands, 1 component ‡ right and left.

* 5 components: parenchymal echogenicity, homogeneity, hypoechogenic areas, hyperechogenic reflections and salivary gland posterior border.

† 2 components: parenchymal echogenicity and hypoechogenic areas.

‡ 1 component: hypoechogenic areas.

The accuracy of SGUS to predict ACR-EULAR classification based on the original Hocevar score (four glands, five components) was good (AUC=0.856; figure 2). Accuracy was comparable when only the scores of the right parotid and right submandibular gland (AUC=0.850), or only the scores of the left parotid and left submandibular gland (AUC=0.852) were included (figure 2).

Table 2.

Logistic regression analyses of separate major salivary glands and Hocevar components to predict ACR-EULAR criteria classification as pSS or non-pSS.

	Univariate analysis		Multivariate analysis		
	OR (95% CI)	p-value	R ²	OR (95% CI)	p-value
Overall Hocevar score					
Total Hocevar score (0-48)	1.216 (1.148–1.288)	<0.001*	0.526	Not applicable	
Separate major salivary glands					
Right parotid gland (0-13)	1.765 (1.488–2.094)	<0.001*	0.449	1.415 (1.156–1.732)	0.001*
Left parotid gland (0-13)	1.718 (1.467–2.013)	<0.001*	0.419	-	†
Right submandibular gland (0-11)	1.854 (1.554–2.211)	<0.001*	0.472	-	†
Left submandibular gland (0-11)	1.913 (1.587–2.307)	<0.001*	0.473	1.538 (1.241–1.906)	<0.001*
Individual components of Hocevar score					
Parenchymal echogenicity (0-4)	4.350 (2.402–7.878)	<0.001*	0.433	2.279 (1.173–4.428)	0.015*
Homogeneity (0-12)	1.592 (1.395–1.816)	<0.001*	0.451	-	†
Hypoechogenic areas (0-12)	1.683 (1.456–1.946)	<0.001*	0.505	1.453 (1.229–1.718)	<0.001*
Hyperechogenic reflections (0-8)	2.101 (1.600–2.760)	<0.001*	0.236	-	†
Salivary gland posterior border (0-12)	1.793 (1.462–2.200)	<0.001*	0.315	-	†

* Statistically significant

† Not included in the multivariate model (based on Forward Wald method).

In the univariate analysis, the explained variance (R^2) for the total Hocevar score of all four glands was 0.526 (table 2 and table 3). SGUS scores of all individual glands significantly predicted ACR-EULAR classification, with R^2 ranging from 0.419 for the left parotid gland to 0.473 for the left submandibular gland (table 2B).

In the multivariate analysis, only the SGUS scores of the right parotid and left submandibular gland independently predicted ACR-EULAR classification ($R^2=0.527$), due to high correlation between both parotid glands and between both submandibular glands. When only the right parotid and right submandibular gland ($R^2=0.518$) or only the left parotid and left submandibular gland ($R^2=0.511$) were included, this resulted in models with equally high explained variance. In summary, scoring of only one side is enough to predict ACR-EULAR classification. Whether left or right side is scored does not make any difference.

Scoring only parenchymal echogenicity and hypoechogenic areas is sufficient.

The distribution of scores for the individual Hocevar components differed significantly between pSS and non-pSS patients (supplementary figure 3). In univariate analysis, all individual components significantly predicted ACR-EULAR classification. ACR-EULAR classification was predicted best by hypoechogenic areas ($R^2=0.505$). The correlation between the components homogeneity and hypoechogenic areas was very high ($p=0.927$; table 3), suggesting that these two components measure the same and that scoring both is not necessary. The multivariate analysis, including all five components measured in all four glands, showed that only parenchymal echogenicity and hypoechogenic areas independently predicted ACR-EULAR classification ($R^2=0.539$), with a good AUC of 0.857.

Table 3.

Spearman's correlation coefficients for individual Hocevar components.

A. Right and left parotid and submandibular glands

	Homogeneity	Hypoechogenic areas	Hyperechogenic reflections	Salivary gland posterior border
Parenchymal echogenicity	0.736	0.729	0.508	0.582
Homogeneity	-	0.927	0.574	0.624
Hypoechogenic areas	-	-	0.603	0.593
Hyperechogenic reflections	-	-	-	0.511

B. Right parotid and submandibular glands

	Homogeneity	Hypoechogenic areas	Hyperechogenic reflections	Salivary gland posterior border
Parenchymal echogenicity	0.725	0.718	0.494	0.573
Homogeneity	-	0.921	0.540	0.602
Hypoechogenic areas	-	-	0.556	0.568
Hyperechogenic reflections	-	-	-	0.503

C. Left parotid and submandibular glands

	Homogeneity	Hypoechogenic areas	Hyperechogenic reflections	Salivary gland posterior border
Parenchymal echogenicity	0.728	0.730	0.427	0.552
Homogeneity	-	0.921	0.517	0.613
Hypoechogenic areas	-	-	0.520	0.589
Hyperechogenic reflections	-	-	-	0.413

Scoring two glands and one or two components increases feasibility of ultrasound in PSS.

Taking into account only the right parotid and right submandibular glands, the results of the univariate analysis of the individual components of the Hocevar scoring system were similar compared with the model including all four glands (table 4B). In the multivariate analysis, only parenchymal echogenicity and hypoechogenic areas independently predicted ACR-EULAR classification ($R^2=0.539$), with a good AUC of 0.855 (figure 2). Taking into account only the left parotid and left submandibular glands, similar results were found (table 4C; figure 2).

Thus, SGUS examination of parotid and submandibular glands on one side and scoring parenchymal echogenicity and hypoechogenic areas, is sufficient to predict classification of patients according to the ACR-EULAR criteria.

However, this system can be simplified even further by only scoring the hypoechogenic component of the Hocevar system, without decreasing the accuracy. When only scoring the hypoechogenic areas of the right parotid and right submandibular glands, the explained variance was 0.498, with a good AUC of 0.846. The explained variance and accuracy were again similar for the left parotid and left submandibular glands (figure 2).

Table 4.

Logistic regression analyses of Hocevar components in the right and left glands to predict ACR-EULAR classification as pSS or non-pSS.

	Univariate analysis			Multivariate analysis	
	OR (95% CI)	p-value	R ²	OR (95% CI)	p-value
Overall Hocevar score					
Total Hocevar score (0-48)	1.216 (1.148–1.288)	<0.001*	0.526	Not applicable	
Individual components of Hocevar system for right parotid and submandibular glands					
Parenchymal echogenicity (0-2)	15.024 (4.988-45.249)	<0.001*	0.427	5.188 (1.533–17.558)	0.008*
Homogeneity (0-6)	2.400 (1.872–3.077)	<0.001*	0.444	-	†
Hypoechogenic areas (0-6)	2.694 (2.034–3.552)	<0.001*	0.498	2.031 (1.490–2.770)	<0.001*
Hyperechogenic reflections (0-4)	3.242 (2.006–5.238)	<0.001*	0.188	-	†
Salivary gland posterior border (0-6)	2.669 (1.872–3.804)	<0.001*	0.292	-	†
Individual components of Hocevar system for left parotid and submandibular glands					
Parenchymal echogenicity (0-2)	13.779 (5.088–37.314)	<0.001*	0.404	4.071 (1.369–12.103)	0.012*
Homogeneity (0-6)	2.443 (1.894–3.152)	<0.001*	0.430	-	†
Hypoechogenic areas (0-6)	2.589 (1.984–3.379)	<0.001*	0.471	1.989 (1.458–2.713)	<0.001*
Hyperechogenic reflections (0-4)	3.552 (2.206–5.720)	<0.001*	0.214	-	†
Salivary gland posterior border (0-6)	2.996 (2.039–4.402)	<0.001*	0.291	-	†

* Statistically significant

† Not included in the multivariate model (based on Forward Wald method).

Discussion

We proved that ultrasound scoring of hypoechogenic areas in one parotid and one submandibular gland is enough to predict ACR-EULAR classification of patients with or without pSS.

Thus far, there is no agreement regarding which ultrasonographic scoring system should be used in the evaluation of major salivary glands of patients suspected or clinically diagnosed with pSS. Based on a recent meta-analysis,³ the Hocevar scoring system⁷ was incorporated in our daily clinical practice to address all relevant aspects of SGUS, since this system appeared to most comprehensive and was therefore also recently applied in our reliability and validity studies.^{6,10} As mentioned before, more simple scoring systems have been used previously,¹¹⁻¹⁴ but, to the best of our knowledge, none of these simpler scoring systems were constructed applying a multivariate regression model. Ideally, in every rheumatologic outpatient clinic around the world the same scoring system for evaluation of salivary glands in the diagnostic workup of pSS should be applied. In order to reach such a goal, an essential first step was to perform a multivariate logistic regression analysis using the results of the most comprehensive scoring system commonly applied.⁷ Having access to a reliable, simpler SGUS scoring system would greatly increase the feasibility of SGUS as a routine procedure to be applied in the diagnostic workup of pSS patients.

There was an excellent correlation between the ultrasound scores of the right and left parotid glands and the ultrasound scores of the right and left submandibular glands. Correlation between ultrasound scores of parotid and submandibular glands was somewhat lower, in agreement with previous reports.²¹ These results indicate that scoring of the two parotid glands is interchangeable and the same applies to both submandibular glands, but that ultrasound scores of the parotid compared with the submandibular glands can differ within patients. In some patients, the parotid glands are more affected, whereas in others the submandibular glands are more affected. Therefore, ultrasound abnormalities might be underdiagnosed in a subgroup of patients when only parotid glands or only submandibular glands are taken into account. This is confirmed by the multivariate regression model, in which only one parotid gland and one submandibular gland were found to independently contribute to ACR-EULAR classification as pSS or non-pSS. Our results therefore show that scoring of one side is adequate, but that the parotid as well as the submandibular gland should be scored.

Previously, studies found more pathological changes in the submandibular glands compared with the parotid glands in patients suspected with pSS.^{22,23} Takagi

et al. therefore suggested to reduce the SGUS examination to only the submandibular glands in order to save time.²² On the other hand, very recently it was suggested to measure only the parotid glands with a simplified scoring system.¹⁴ Based on our results, we cannot confirm either suggestion. Furthermore, neither of these studies performed multivariate regression analyses to actually proof their statements.

Another important reason to include SGUS examination of the parotid glands, besides that it was included in our model, is that suspicion of the presence of a mucosa-associated lymphoid tissue lymphoma in the parotid gland might rise at an earlier stage. We therefore recommend to compile a global view of all four glands, especially when patients experience unilateral complaints, but to actually score only one side. With this global view, unilateral and bilateral morbidities of the salivary glands can be distinguished, increasing the specificity of the examination.

At group level, all five individual components of the Hocevar scoring system were significant predictors of ACR-EULAR classification. However, the differences in explained variance were quite large, that is, with the lowest value for hyperechoic reflections and the highest value for hypoechogenic areas, indicating that not every component is equally highly predictive of ACR-EULAR classification. In the multivariate analysis, parenchymal echogenicity and hypoechogenic areas were found to independently predict ACR-EULAR classification. It might even be considered to measure only hypoechogenic areas of one parotid and one submandibular gland, since parenchymal echogenicity added only slightly to the explained variance and diagnostic accuracy. Taking into account the reliability of the individual components of the Hocevar score, reducing the scoring system to only hypoechogenic areas increases the applicability even more. Recently, we showed that the interobserver reliability of hypoechogenic areas was good ($\kappa > 0.7$). In contrast, the reliability of parenchymal echogenicity was only fair.⁶ Two other studies showed similar results even though slightly different scoring systems were applied.^{24,25} Both articles also contain ultrasonographic images of major salivary glands with different stages of ultrasonographic abnormalities suggestive of pSS.^{24,25}

Previously, it was assumed that the presence of hypoechoic rounded multiple parenchymal lesions is a pathognomonic feature of pSS and that only evident parenchymal inhomogeneity is of true diagnostic value.^{12,13} We agree with these assumptions. In the Hocevar scoring system, homogeneity and hypoechogenic areas are scored separately. On the other hand, in the scoring system used by Theander and Mandl,¹² only parenchymal homogeneity is measured. However, when taking a closer look at their scoring system, in fact a combination of homogeneity and hypoechogenic areas is scored.¹² Approximately the same component was measured in a German cohort.¹³ In our study, the correlation between homogeneity and

hypoechoogenic areas was excellent. Therefore, it can indeed be interpreted that both components measure the same concept in salivary glands of patients suspected or diagnosed with pSS. When there is a homogenous gland, there are no hypoechoogenic areas present and when the gland is very inhomogeneous, this is usually due to the presence of hypoechoogenic areas and less due to the presence of hyperechogenic reflections.

Reduction of the examination duration has the advantage that the SGUS examination is easier to learn and takes less time to perform in clinical practice. By increasing the feasibility of SGUS, it will facilitate outpatient rheumatologic clinics worldwide to integrate the ultrasound in the daily clinical practice when it comes to evaluating patients clinically suspected or diagnosed with pSS. We demonstrated that it is adequate to measure only one out of five components of the Hoyer scoring system, in only one side, without recommending a particular side.

The next step, before incorporation of SGUS in the classification criteria is possible, would be to define exactly how to score hypoechoogenic areas and decide the optimal cut-off point based on data from multiple centers.

In conclusion, in our cohort study derived from daily clinical practice, we showed that it is sufficient to measure parenchymal echogenicity and hypoechoogenic areas in one parotid and one submandibular gland to predict classification as pSS or non-pSS according to the ACR-EULAR criteria. Taking into account the relatively small difference in diagnostic accuracy as well as the reliability of the individual ultrasound components, we pose that only measuring hypoechoogenic areas in one parotid and one submandibular gland also is sufficient. As the severity of SGUS findings in the parotid and submandibular glands may differ, one parotid and one submandibular gland of either the left or the right side should be scored. Simplification of SGUS scoring will increase the feasibility of SGUS in daily clinical practice and will facilitate worldwide implementation of the ultrasound in the outpatient rheumatologic clinics.

Acknowledgement:

This study was conducted on behalf of the EULAR US-pSS study group. Support was obtained from the Dutch Arthritis Foundation and Horizon 2020.

Part of this study has been presented at the ACR 2017 in San Diego and the corresponding abstract has been published as a conference abstract.²⁶

Funding info:

The Dutch Arthritis Foundation, research grant for the REgistry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) cohort; METc 2014/491.

Bristol-Myers Squibb, unrestricted research grant for the RESULT cohort.

Horizon 2020, a research project supported by European Commission: HarmonicSS primary Sjögren's Syndrome (pSS) towards improved satisfaction, treatment and health policy making (H2020-SC1-2016-RTD, proposal 731944).

No funding has been received for the conduct and/or interpretation of this article.

References

1. Fox RI. Sjögren's syndrome. *Lancet* 2005;366:321-31.
2. Kassan SS, Moutsopoulos HM. Clinical manifestations and early diagnosis of Sjögren's syndrome. *Arch Intern Med* 2004;164:1275-84.
3. Delli K, Dijkstra PU, Stel AJ, *et al.* Diagnostic properties of ultrasound of major salivary glands in Sjögren's syndrome: a meta-analysis. *Oral Dis* 2015;21:792-800.
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 (Oxford)* 2016;55:789-800.
5. Bootsma H, Spijkervet FKL, Kroese FKL, *et al.* Toward new classification criteria for Sjögren's syndrome? *Arthritis Rheum* 2013;65:21-3.
6. Delli K, Arends S, van Nimwegen JF, *et al.* Ultrasound of the major salivary glands is a reliable imaging technique in patients with clinically suspected primary Sjögren's syndrome. *Ultraschall Med* 2018;39:328-33.
7. Hocevar A, Ambrozic A, Rozman B, *et al.* Ultrasonographic changes of major salivary glands in primary Sjögren's syndrome. Diagnostic value of a novel scoring system. *Rheumatology (Oxford)* 2005;44:768-72.
8. 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.
9. 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.
10. Mossel E, Delli K, van Nimwegen JF, *et al.* Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1883-89.
11. Cornec D, Jousse-Joulin S, Pers J-O, *et al.* Contribution of salivary gland ultrasonography to the diagnosis of Sjögren's syndrome: Toward new diagnostic criteria? *Arthritis Rheum* 2013;65:216-25.
12. Theander E, Mandl T. Primary Sjögren's syndrome: diagnostic and prognostic value of salivary gland ultrasonography using a simplified scoring system. *Arthritis Care Res* 2014;66:1102-7.
13. Wernicke D, Hess H, Gromnica-Ihle E, *et al.* Ultrasonography of salivary glands – A highly specific imaging procedure for diagnosis of Sjögren's syndrome. *J Rheumatol* 2008;35:285-93.

14. 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.
15. Boers M, Brooks P, Strand CV, *et al.* The OMERACT filter for Outcome Measures in Rheumatology. *J Rheumatol* 1998;25:198-9.
16. Shiboski CH, Shiboski SC, Seror R, *et al.* American College of Rheumatology/ European League Against Rheumatism classification criteria for primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:9-16.
17. Pijpe J, Kalk WWI, van der Wal JE, *et al.* Parotid gland biopsy compared with labial biopsy in the diagnosis of patients with primary Sjögren's syndrome. *Rheumatology (Oxford)* 2007;46:335-41.
18. Chisholm DM, Mason DK. Labial salivary gland biopsy in Sjögren's disease. *J Clin Pathol* 1968;21:656-60.
19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
20. Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. Third edition 2013.
21. Baldini C, Luciano N, Tarantini G, *et al.* Salivary gland ultrasonography: a highly specific tool for the early diagnosis of primary Sjögren's syndrome. *Arthritis Res Ther* 2015;17:146
22. Takagi Y, Kimura Y, Nakamura H, *et al.* Salivary gland ultrasonography: can it be an alternative to sialography as an imaging modality for Sjögren's syndrome? *Ann Rheum Dis* 2010;69:1321-4.
23. Hammenfors DS, Brun JG, Jonsson R, *et al.* Diagnostic utility of major salivary gland ultrasonography in primary Sjögren's syndrome. *Clin Exp Rheumatol* 2015;33:56-62.
24. Damjanov N, Milic V, Nieto-Gonzalez JC, *et al.* Multiobserver reliability of ultrasound assessment of salivary glands in patients with established primary Sjögren's syndrome. *J Rheumatol* 2016;43:1858-63.
25. 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.
26. Mossel E, Arends S, van Nimwegen JF, *et al.* Ultrasound scoring of one parotid and one submandibular gland in primary Sjögren's syndrome – further increasing feasibility in outpatient clinics [abstract]. *Arthritis Rheumatol* 2017;69:(suppl 10).

Supplementary Figure 1

- A. Normal submandibular gland;
- B. Submandibular gland suggestive of primary Sjögren's syndrome;
- C. Normal parotid gland and
- D. Parotid gland suggestive of primary Sjögren's syndrome.

Supplementary Figure 2

Ultrasound scores of major salivary glands in pSS and non-pSS patients according to the ACR-EULAR classification criteria. A. Total ultrasound score for right and left parotid and submandibular glands; Total ultrasound score for the left B. and right C. parotid; and Total ultrasound score for the left D. and right E. submandibular glands. Ultrasound scores of all four glands differed significantly between pSS and non-pSS patients.

ACR-EULAR, American College of Rheumatology – European League Against Rheumatism; pSS, primary Sjögren's syndrome.

Supplementary Figure 3

Distribution of scores on the individual components of Hoyer scoring system for ACR-EULAR positive and ACR-EULAR negative patients. A. Parenchymal echogenicity; B. Homogeneity; C. Hypoechoic areas; D. Hyperechoic reflections and E. Salivary gland posterior border. Ultrasound scores of all five components differed significantly between pSS and non-pSS patients. ACR-EULAR, American College of Rheumatology – European League Against Rheumatism; pSS, primary Sjögren’s syndrome.

CHAPTER 4

Histopathology, salivary flow and ultrasonography of the parotid gland: three complementary measurements in primary Sjögren's syndrome

Esther Mossel^{1*}, Martha S van Ginkel^{1*}, Erlin A Haacke^{1,2}, Suzanne Arends¹, Silvia Liefers¹, Konstantina Delli³, Jolien F van Nimwegen¹, Alja J Stel¹, Fred KL Spijkervet³, Bert van der Vegt², Arjan Vissink³, Frans GM Kroese¹, Hendrika Bootsma¹

* both authors contributed equally

¹ Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

² Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

³ Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Edited version of *Submitted Manuscript*

Abstract

Objective

The impact of primary Sjögren's syndrome (pSS) on the salivary gland can be assessed in three different ways: histopathology, salivary flow and ultrasonography. To understand the relative importance of these different approaches, it is crucial to understand the relationship between them. In general practice these measurements are not performed on one and the same salivary gland. As we routinely perform these three modalities in the diagnostic work-up (all in the parotid gland), our aim was to investigate the construct validity between these modalities in the same gland.

Methods

Consecutive patients were evaluated in a multidisciplinary diagnostic setting and underwent a parotid gland biopsy, collection of parotid gland specific saliva and parotid ultrasonography. Patients who were classified as pSS according to the ACR – EULAR criteria were evaluated for eligibility. Construct validity was assessed using Spearman's correlation coefficients.

Results

The 41 included pSS patients completed a full work-up within mean time interval of 2.6 months. Correlations between histopathological features and unstimulated and stimulated parotid salivary flow were poor to fair ($p=0.095$ and $p=-0.123$ for focus score and $p=-0.226$ and $p=-0.259$ for percentage of CD45⁺ infiltrate). Poor correlations were observed between (unstimulated and stimulated) parotid salivary flow and parotid ultrasonography ($p=0.069$ and $p=-0.196$). Moderate to good associations were found between the histopathological items focus score and percentage of CD45⁺ infiltrate, with parotid ultrasound scores (total score: $p=0.510$ and $p=0.560$; highest for homogeneity: $p=0.574$ and $p=0.633$).

Conclusion

The three modalities that evaluate salivary gland involvement, assess different or at best partly related constructs. Therefore, histopathology, salivary flow and ultrasonography are complementary measurements and cannot directly replace each other in the work-up of pSS.

Introduction

Primary Sjögren's syndrome (pSS) is a chronic, systemic auto-immune disease characterized by oral and ocular sicca complaints, besides a wide variety of systemic signs and symptoms.^{1,2} Inflammation of the salivary and lacrimal glands is a hallmark of the disease and plays a central role in the current classification criteria.^{3,4} The impact of pSS on salivary glands can be assessed in three different ways: by performing a salivary gland biopsy, by collecting saliva and by salivary gland ultrasonography.⁵ The histopathological assessment of salivary glands and collection of salivary secretion are included in the current classification criteria as separate items, and it has recently been proposed to add ultrasonography as an item.^{3,6,7}

Histopathological analysis of salivary gland biopsies consists of the detection of features that are characteristic for pSS. The inflammatory process in salivary glands of pSS patients is characterized by periductal lymphocytic infiltrates called foci (clusters of ≥ 50 lymphocytes). The number of foci per 4 mm² glandular parenchyma can be calculated into a focus score (FS).^{8,9} For classification of pSS, in which histopathology of salivary glands is a cornerstone, biopsies with FS ≥ 1 are considered positive.³ Other characteristic histopathological features of pSS are presence of lymphoepithelial lesions (LELs) and germinal centers (GCs), and a relative increase of IgG and IgM plasma cells, compared to the number of IgA plasma cells (so-called plasma cell shift).¹⁰⁻¹⁴ Furthermore, salivary glands of pSS patients contain higher proportions of fibrosis and acinar atrophy than controls.^{15,16} Whether fatty infiltration is specific for pSS or an age-associated condition remains to be shown.^{17,18} Salivary gland biopsies are mostly taken from labial glands during the diagnostic work-up of pSS. However, taking biopsies from parotid glands offers several advantages, such as increased specificity for diagnosing pSS, fewer co-morbidities, the possibility to take multiple biopsies of the same gland and the advantage of early detection of pSS associated lymphomas localized in the parotid gland.^{19,20} For these reasons, we routinely perform parotid gland biopsies instead of labial gland biopsies from patients suspected of pSS.

The second tool to evaluate involvement of salivary glands in pSS is the assessment of salivary secretion. Saliva can be collected in different ways, and unstimulated whole salivary flow (UWSF) is currently included in the ACR-EULAR criteria as an item.^{3,4} UWS is composed of contributions of all minor and major glands, albeit that UWS only contains minor amounts of saliva from the parotid glands, since the parotid glands only contribute significantly to whole saliva after stimulation.²¹ Saliva can also be collected under stimulated conditions, such as masticatory (chewing)

or gustatory (citric acid solution) stimulation. Stimulated whole salivary flow (SWS) is also a mixture of saliva produced by all minor and major salivary glands, but is not predominantly composed of saliva from parotid and submandibular glands. It is, however, also possible to collect saliva exclusively from the parotid glands.²¹ Herewith, parotid flow rates can be used for direct functional assessment of the parotid glands.

The third tool to assess salivary gland involvement in pSS is salivary gland ultrasonography (SGUS). SGUS evaluates major salivary glands, i.e., parotid and submandibular glands. This upcoming tool is non-invasive, well-tolerated, inexpensive, more and more available in the outpatient clinic, and can be repeatedly performed.^{22,23} Multiple scoring systems of SGUS are in use, of which the Hocevar scoring system is most commonly applied.²⁴ We previously showed that scoring hypoechogenic areas suffices for classification and herewith further increases the feasibility of SGUS in the outpatient clinic.²⁵

Together, there are three modalities to assess salivary gland involvement in pSS: histopathology, collection of saliva and ultrasonography. To understand the relative importance of these different approaches, it is crucial to understand the relationship between them. Although patients, secreting less than 0.1 ml/min UWS were two times more likely to also have a $FS \geq 1$,²⁶ correlations between histopathological findings, such as FS and fibrosis, and UWS as continuous parameters are poor.²⁷⁻²⁹ When comparing SGUS scores (from major salivary glands) with labial gland FS and salivary secretion, correlation coefficients seem to be higher.³⁰⁻³² However, interpretation of these findings is complicated since histopathology, collection of saliva and ultrasonography were not assessed in the same salivary gland. Furthermore, different methods of saliva collection and SGUS scoring systems were used. Since we routinely perform parotid gland biopsies, collect parotid gland specific saliva, and perform parotid gland ultrasonography in the diagnostic work-up of pSS, we are in the unique position to analyze the relationship between these three modalities in the same gland. In the work presented here, we aim to explore this relationship by assessing construct validity between histopathology, salivary secretion, and ultrasonography of the same parotid gland in a group of pSS patients.

Patients and Methods

Patients

This cross-sectional study is based on prospective data from the multidisciplinary Sjögren's expertise center at the University Medical Center Groningen (UMCG), a tertiary referral center. Consecutive patients were evaluated in a diagnostic setting and underwent a full work-up including parotid gland biopsy, collection of parotid gland specific saliva and ultrasonography, all within a maximum timeframe of one year. Patients who were over 18 years of age and classified as pSS according to the American College of Rheumatology (ACR) – European League Against Rheumatism (EULAR) criteria were included.^{3,4}

Histochemical and immunohistochemical staining

Parotid gland biopsies were formalin fixed, paraffin embedded and sectioned at 3 µm thickness. Consecutive slides were stained with hematoxylin and eosin (H&E) immunohistochemically for CD45, Bcl-6, IgA/IgG and IgM and by a modified Masson stain. Immunohistochemical staining for CD45, IgA/IgG and IgM was performed on an automated staining platform (Benchmark XT, Ventana Medical Systems, Inc.), following the manufacturer's protocols. Immunohistochemical staining was performed with monoclonal anti-CD45 antibodies (clones 2B11+PD7/26, Ventana Roche), and IgA/IgG dual staining and an IgM single staining with polyclonal rabbit-anti-human antibodies (Ventana Roche). All antibodies were pre-diluted by the supplier. Tissue samples were manually stained with a modified Masson stain and for Bcl-6. For the modified Masson staining, tissue sections were deparaffinized and incubated for 5 minutes with acid fuchsin and Ponceau xylin in the proportion of 1:2. After rinsing, sections were incubated with 1% phosphomolybdic acid for 5 minutes. Sections were rinsed again and incubated for 45 seconds with aniline blue. After the final rinsing step, slides were dehydrated and covered with a xylene based mounting medium. For the Bcl-6 staining, tissue sections were deparaffinized and antigen retrieval was performed (EDTA pH 8, 15 minutes 98°C). After blocking of endogenous peroxidase, slides were incubated with monoclonal anti-Bcl6 (clone GL19E/A8, Ventana Roche, pre-diluted by supplier) for 75 min. After rinsing, sections were incubated with anti-mouse Ig HRP-polymer (TL-125-HL, Thermofisher) for 40 minutes. Staining was visualized with diaminobenzidine and slides were counterstained with hematoxylin.

Figure 1. (Immuno-)histological analysis of parotid gland tissue from pSS patients

Parotid gland biopsies from pSS patients showing A. A periductal focus on H&E staining with a centrally located lymphoepithelial lesion (LEL). B. High resolution image of the same LEL, showing hyperplastic epithelium with intra-epithelial lymphocytes. C. Presence of a germinal center, as shown by a cluster of ≥ 5 Bcl6⁺ cells. D. Dual staining for IgA (red) and IgG (brown) plasma cells showing influx of IgG⁺ plasma cells. E. CD45 staining of parotid gland tissue. F. Digital image analysis of the relative area of CD45⁺ infiltrate by using Qupath. G. Parotid gland tissue stained with a modified Masson stain in which dense connective tissue is colored blue. H. Digital image analysis of the modified Masson staining by using Qupath, in which red represents fibrotic tissue, yellow represents fat cells and green represents glandular parenchyma.

Histological analysis

FS and presence of LELs were assessed on H&E stained sections (figure 1A). FS was defined as the number of periductal foci per 4 mm² glandular parenchyma (figure 1A). An arbitrary maximum FS of 12 was given to confluent foci. LELs were determined by the presence of lymphocytes within the ductal epithelium with concurrent epithelial hyperplasia (figure 1A). Presence of a plasma cell shift was assessed using IgA/IgG double staining and IgM staining (figure 1D). A relative decrease of $\leq 70\%$ IgA⁺ plasma cells in relation to the total plasma cell population (foci and unaffected parenchyma) was considered as abnormal.¹⁴ A cluster of ≥ 5 Bcl6⁺ cells, lying next to each other, within a focus was classified as GC (Figure 1G).^{33,34} Biopsies were independently scored by two pathologists (EAH and BV). Inconsistencies were solved during a consensus meeting.

Relative area of CD45⁺ infiltrate in relation to the total area of glandular parenchyma was evaluated on CD45 stained slides, and relative areas of fatty and fibrotic tissue were evaluated on the modified Masson staining. Digital image analyses (DIA) of both slides were performed by using the open software QuPath v0.2.3 (figure 1H).

Salivary secretion

Unstimulated whole saliva (UWS) and chewing stimulated whole saliva (SWS) were collected. UWS secretion rates of $\leq 1.5\text{mL}/15\text{min}$ and SWS secretion rates of $\leq 3.5\text{mL}/5\text{ min}$ were considered abnormal. Unstimulated parotid salivary flow (UPF) and 2% citric acid stimulated parotid salivary flow (SPF) were measured using Lashley cups.²¹

Figure 2. Presence of various amounts of hypoechogenic areas on parotid gland ultrasonography. Scores range from 0-3 on the Hoyer scoring system; A. 0 = absent; B. 1 = a few; C. 2 = several, and D. 3 = numerous.

Ultrasonography

SGUS was performed using an ultrasonographic scanner (Esaote MyLabSeven, Genova, Italy), equipped with a high-resolution probe (4-13MHz), by three well-trained ultrasonographers (AJS, KD, JFN), with good to excellent inter- and intra-observer reliability.³⁵ Patients were examined in a supine position with their head turned towards the opposite side and neck slightly extended. According to the scoring system of Hocevar et al.,²⁴ the following components were scored in both parotid and submandibular glands: (i) parenchymal echogenicity; (ii) homogeneity; (iii) presence of hypoechoic areas (figure 2); (iv) hyperechoic reflections and (iv) clearness of the salivary gland border. For the total Hocevar score, a score of ≥ 15 was considered positive for diagnosing Sjögren's syndrome.³⁶ Separate analyses were done for parotid gland ultrasonography (PGUS), with emphasis on the components 'hypoechoic areas' and 'hyperechoic reflections' (PGUS-hypo and PGUS-hyper).²⁵

Statistical analysis

Descriptive parameters were expressed as number (%), mean (SD) or median (IQR) as appropriate.

Construct validity was assessed using Spearman's correlation coefficients (ρ) between histopathological parameters and parotid salivary secretion, between histopathological parameters and PGUS scores, and between parotid salivary flow and PGUS scores. Spearman's correlation coefficient (ρ) was interpreted as poor agreement (0.0-0.2), fair (0.2-0.4), moderate (0.4-0.6), good (0.6-0.8) or excellent agreement (0.8-1.0).³⁷

It was hypothesized that there would be fair to moderate correlations between parotid histopathological parameters and parotid salivary secretion, as they assess related, but different types of outcomes. Fair correlations between both measurements were also found in previous studies.²⁷⁻²⁹ Furthermore, it is known that intra-individual variation of parotid salivary flow is around 24%,³⁸ and that salivary flow fluctuates with circadian rhythm and age.^{39,40} This variability will lower the level of correlation. Correlations between histopathological parameters and PGUS scores were hypothesized to be moderate as they partly assess the same construct; the parotid gland biopsy represents a small part of the whole gland that is analyzed by ultrasound. Finally, it was hypothesized that there would be fair correlations between parotid salivary flow and PGUS, as they assess different outcomes within the same construct, and the measurement variation in parotid salivary flow was also taken into account.

As exploratory analyses, Kruskal Wallis tests were used to compare histopathological parameters between various groups based on PGUS-hypo and PGUS-hyper scores. Mann Whitney U tests were used to compare variables between groups based on presence vs. absence of histopathological features and between groups based on presence of multiple hypoechogenic areas (PGUS-hypo $\geq$ 2) vs. limited or no hypoechogenic areas (PGUS-hypo $\leq$ 1).

As the histopathological and salivary flow items that are included in the current 2016 ACR-EULAR criteria of pSS are not salivary gland specific, we also calculated the percentage of absolute agreement between the different methods based on the cut-off values that are used for classification of pSS: FS $\geq$ 1, UWSF $\leq$ 0.1 ml/min, SGUS $\geq$ 15.36 Furthermore, the associations between FS, UWS and SGUS were explored using Spearman's correlation coefficients (ρ). Statistical analyses were executed using IBM SPSS Statistics 23 (SPSS, Chicago, IL, USA). P-values <0.05 were considered statistically significant.

Results

Table 1.

Patient characteristics of the included pSS patients.

	pSS patients (n=41)
Age (years)	50.4 ($\pm$ 13.6)
Females (%)	40 (97.6)
ACR-EULAR items	
Anti-SSA/Ro antibodies	37 (90.2)
FS $\geq$ 1, parotid gland	25 (61.0)
UWSF $\leq$ 0.1 mL/min*	27 (65.9)
SWSF (mL/min)*	0.29 (0.09-0.45)
Schirmer's test $\leq$ 5mm/5min	32 (78.0)
OSS $\geq$ 5*	23 (56.1)
SGUS $\geq$ 15#	25 (61.0)
Parotid histology	
FS	1.1 (0.4-2.3)
% CD45 ⁺ infiltrate	5.3 (2.2-18.9)
% Fat	37.9 (24.2-45.8)
% Fibrosis	19.8 (6.8-26.1)
Presence of LELs	20 (48.8)
$\leq$ 70% IgA plasma cells	15 (36.6)
Presence of GCs	14 (34.1)
Parotid saliva	
UPF (mL/min)	0.00 (0.00-0.03)
SPF (mL/min)	0.06 (0.02-0.11)
Parotid ultrasonography	
PGUS total score (range 0-10)	4.0 (2.0-7.5)
PGUS-homogeneity (range 0-3)	1.0 (0.0-2.0)
PGUS-hypo (range 0-3)	1.0 (1.0-3.0)
PGUS-hyper (range 0-2)	1.0 (1.0-1.0)

Data are expressed as number (%), mean ($\pm$ SD), or median (IQR). * <5% missing data. # Salivary gland ultrasonography (SGUS) is not (yet) included in the ACR-EULAR criteria, but a score of ≥ 15 has been proposed. FS: focus score; UWSF: unstimulated whole saliva flow; SWSF: stimulated whole saliva flow; OSS: ocular staining score; SGUS: salivary gland ultrasonography, UPF: unstimulated parotid flow (biopsy side); SPF: stimulated parotid flow (biopsy side); LELs: lymphoepithelial lesions; GCs: germinal centers; PGUS: parotid gland ultrasonography (biopsy side); PGUS-hypo: hypoechogenic areas; PGUS-hyper: hyperechogenic reflections.

Patient characteristics

Forty-six consecutive pSS patients were evaluated for eligibility. Four patients had to be excluded because of insufficient biopsy material (surface area $<4\text{mm}^2$) and one patient because of presence of a parotid gland MALT lymphoma. Patient characteristics are presented in table 1. The mean time interval between the biopsy and saliva collection was 2.8 months (± 2.8), between the biopsy and SGUS 2.7 months (± 2.9), and between SGUS and saliva collection 2.3 months (± 3.2).

Parotid salivary flow versus parotid histology

Unstimulated parotid flow (UPF) showed poor to fair associations with FS, percentages of CD45⁺ infiltrate, and percentages of fat and fibrosis within the parotid gland (table 2, figure 3A, 3C). These limited associations are not surprising, because parotid glands secrete saliva only after stimulation. Correlation coefficients between stimulated parotid flow (SPF) and histological parameters were slightly higher, but associations were still fair (table 2, figure 3B, 3D). No significant differences in parotid flow (UPF or SPF) were found between pSS patients with presence of LELs (UPF $p=0.27$ and SPF $p=0.39$) or presence of a plasma cell shift (UPF $p=0.83$, and SPF $p=0.75$) compared to pSS patients without these histopathological features. Although SPF levels seemed to be lower in pSS patients with presence of GCs compared to pSS patients without GCs, no significant differences were found (median 0.02 vs. 0.07, $p=0.08$). Overall, as we expected, there is no clear relation between histological features within parotid gland biopsies and parotid salivary flow rates in our diagnostic cohort.

Table 2.

Spearman correlations between parotid histological parameters, parotid salivary flow and parotid gland ultrasonography.

	Parotid gland histopathology				Parotid salivary flow	
	FS	% CD45	% Fat	% Fibrosis	UPF (ml/min)	SPF (ml/min)
Parotid salivary flow						
UPF (ml/min)	0.095	-0.226	0.228	-0.045	-	-
SPF (ml/min)	-0.123	-0.259	0.239	-0.133	-	-
Parotid gland ultrasonography						
PGUS total score (biopsy side) (0-10)	0.510*	0.560*	-0.110	0.077	0.069	-0.196
PGUS subscores of individual ultrasound components						
Parenchymal echogenicity (0-1)	0.219	0.346*	-0.186	0.177	0.134	0.015
Homogeneity (0-3)	0.574*	0.633*	0.197	0.156	-0.002	-0.224
Hypoechoogenic areas (0-3)	0.523*	0.540*	-0.114	0.034	0.042	-0.227
Hyperechoogenic reflections (0-2)	0.274	0.170	0.173	0.061	0.226	0.157
Salivary gland posterior border (0-2)	0.091	0.231	0.028	-0.069	-0.027	-0.159

FS: focus score; UPF: unstimulated parotid flow; SPF: stimulated parotid flow; PGUS: parotid gland ultrasonography. The range of individual ultrasound components is shown between brackets.

ρ: Spearman correlation coefficient. *p<0.05

Figure 3: Parotid flow in comparison with focus score, percentage of CD45⁺ infiltrate and parotid gland ultrasonography.

- A. Unstimulated parotid flow (UPF) and parotid gland focus score (FS).
- B. Stimulated parotid flow (SPF) and parotid gland FS.
- C. UPF and percentage of CD45⁺ infiltrate (%CD45).
- D. SPF and %CD45.
- E. UPF and parotid gland ultrasonography (PGUS);
- F. SPF and PGUS.

Parotid histology versus parotid ultrasonography

Correlation coefficients of histological parameters with PGUS total and PGUS subscores are shown in table 2. Moderate associations were found between FS, percentage of CD45⁺ infiltrate and PGUS total scores. For the individual ultrasound components PGUS-homogeneity and PGUS-hypo, moderate to good associations were found with FS and percentage of CD45⁺ infiltrate (table 2). Remarkably, three patients with a positive FS (i.e. FS $\geq$ 1) had no hypoechogenic areas within the parotid gland (i.e. PGUS-hypo=0), and vice versa three patients with a negative FS did have hypoechogenic areas (i.e. PGUS-hypo $\geq$ 2) (figure 4A). In other words, FS $\geq$ 1 was not always reflected by presence of hypoechogenic areas, and presence of hypoechogenic areas was not always reflected by a positive FS. The percentages of fat and fibrosis did not differ significantly between the different PGUS-hypo and PGUS-hyper scores (figure 4C-D, 4G-H), which was also reflected by poor associations between these parameters (table 2).

Patients with presence of LELs and patients with presence of GCs had significantly higher PGUS total scores compared to patients without these features (figure 5A-B). The two individual components that contributed most to this difference in PGUS total scores were homogeneity and hypoechogenic areas. Scores in these two components were significantly higher in patients with LELs and GCs (p=0.014 for PGUS-homogeneity and p=0.004 for PGUS-hypo). However, patients with LELs and GCs did not always show hypoechogenic areas on ultrasound, as four patients with LELs (20%) and two patients with GCs (14%) had a PGUS-hypo score of 0 (figure 5D-E). In conclusion, pSS associated ultrasonographic findings (PGUS-hypo and PGUS-homogeneity) correlated with glandular histological parameters of inflammation and were also associated with typical histopathological features.

Figure 4: Comparison between different ultrasound and biopsy features.

Hypoechogenic areas in the parotid gland compared with

- A. focus score;
- B. percentage of CD45⁺ infiltrate;
- C. percentage of fat; and
- D. percentage of fibrosis.

Hyperechogenic reflections in the parotid gland compared with

- E. focus score;
- F. percentage of CD45⁺ infiltrate;
- G. percentage of fat; and
- H. percentage of fibrosis.

Hypoechogenic areas were scored as follows: 0 = absent; 1 = a few (less than 25%); 2 = several (between 25 and 50%) and 3 = numerous (more than 50%).

Hyperechogenic reflections were scored as follows: 0 = absent; 1 = a few, scattered (less than 25%); 2 = several (between 25 and 50%) and 3 = numerous (more than 50%). Dotted lines represent FS=1

Figure 5: PGUS total and PGUS-hypo scores in comparotison with presence of LELs, GCs and plasma cell shift

PGUS total scores in pSS patients with and without presence of A. lymphoepithelial lesions (LELs); B. germinal centers (GCs); and C. plasmacell shift. PGUS-hypo scores in pSS patients with and without presence of D. lymphoepithelial lesions (LELs); E. germinal centers (GCs); and F. plasmacell shift. Horizontal bars respresent medians.

Parotid salivary flow versus parotid ultrasonography

Reduced parotid flow was poorly reflected by abnormal PGUS scores, as UPF and SPF both showed poor associations with PGUS total scores and PGUS components (table 2, figure 3E-F). UPF and SPF levels were not significantly lower in pSS patients with multiple hypoechogenic areas in the parotid gland (PGUS-hypo ≥ 2), compared to patients with limited or no hypoechogenic areas (PGUS-hypo ≤ 1) ($p=0.62$ and $p=0.29$, respectively). Apparently, typical PGUS abnormalities and parotid salivary secretion are not directly related to each other.

Agreement of the three modalities used for pSS classification

The histopathological and salivary flow items that are included in the current 2016 ACR-EULAR classification criteria of pSS are not salivary gland specific. Also, the item ultrasonography, that has been proposed to be included in these criteria, is not specific for the parotid gland. To compare the three different modalities as used in the classification criteria, we analysed the agreement between the different modalities based on their cut-off values (FS ≥ 1 , UWSF ≤ 0.1 ml/min, SGUS-score ≥ 15) (figure 6). Absolute agreement between FS and UWSF was 61%, between FS and SGUS 80% and between UWSF and SGUS 56%. The highest correlation was found between FS and SGUS ($\rho=0.607$, $P<0.001$). Correlation between FS and UWSF was poor ($\rho=-0.140$, $P=0.38$), and between UWSF and SGUS fair ($\rho=-0.320$, $P=0.042$) (figure 6). As shown in figure 6D, all SGUS-positive patients had foci in the parotid gland (i.e. FS >0), and $>80\%$ had a FS ≥ 1 . Furthermore, almost all SGUS-positive patients (89%) showed at least one histological feature associated with pSS (FS ≥ 1 , presence of LELs, GCs, plasma cell shift) within their biopsy. However, 38% of SGUS-negative patients also showed one or more histopathological features within their parotid gland biopsy, indicating that presence of specific histopathological features of pSS is not always associated with a positive ultrasound.

Figure 6. Comparison of parotid gland items as used in the ACR-EULAR classification criteria
A. Unstimulated whole saliva flow (UWSF) and focus score (FS); B. FS and salivary gland ultrasonography (SGUS); and C. UWSF and SGUS. D. Percentages of SGUS positive and SGUS negative patients with presence of histopathological features within the parotid gland biopsy. * $p < 0.05$.

Discussion

The present unique study showed that evaluations of the same parotid gland by histopathology, salivary flow and ultrasonography are complementary to each other. We showed that in patients with pSS, correlations between parotid gland salivary flow and parotid gland histopathological parameters were poor to fair. Likewise, no clear relation was found between parotid salivary flow and abnormalities found by parotid gland ultrasonography. Apparently, salivary flow does not assess the same construct as parotid gland histopathology, nor as parotid gland ultrasonography. On the other hand, we showed moderate to good correlations between the histopathological parameters FS and amount of infiltrate and the ultrasound parameters PGUS-total, PGUS-homogeneity and PGUS-hypo (table 2). Thus, typical ultrasound findings in pSS patients seem to be related to the amount of inflammatory infiltrate in the salivary gland biopsy. However, this was not seen in all pSS patients, as there are still some patients with a positive ultrasound and $FS < 1$, as well as patients with lymphocytic foci within the biopsy without typical abnormalities found on ultrasound (figure 4).

Recently, Baldini et al.⁴¹ compared ultrasound and histopathology of labial salivary glands. They found a moderate correlation ($p=0.407$) between labial gland FS and labial gland ultrasound scores in patients suspected of pSS.⁴¹ We found a comparable, slightly higher correlation ($p=0.510$) between FS and PGUS total score of the parotid glands and even a better correlation between the amount of infiltrate in terms of amount of infiltrate and PGUS total scores. From these results we can conclude that the inflammatory infiltrates are somehow related to ultrasound scores, both in the labial and in the parotid salivary gland. In the present study, the highest correlations were found between FS and amount of CD45 infiltrate and the ultrasound scores PGUS-hypo and PGUS-homogeneity. This implies that hypoechogenic areas and inhomogeneity are linked to histopathological lymphocytic infiltrates. Previously, it has been suggested that hypoechogenic areas represent lymphocytic infiltrates.^{42,43} However, as hypoechogenic areas are significantly larger in size compared to foci seen in tissue sections of biopsies, these hypoechogenic areas cannot directly represent foci. Therefore, an indirect association between hypoechogenic areas and infiltrates is more likely. Recently, Wang et al.⁴⁴ hypothesized that presence of hypoechogenic areas could be caused by leakage of saliva into the periductal infiltrate and eventually into the salivary gland parenchyma, due to dysfunction of tight junctions between striated ductal cells.⁴⁴ Although this hypothesis may explain why hypoechogenic areas are much larger than infiltrates, large amounts of interstitial fluid have not been observed in glandular tissue.

Another explanation would be that hypoechogenic areas represent areas with a higher density of periductal foci with relatively little unaffected glandular tissue in between these foci, which is also supported by the good correlation between the amount of infiltrate and hypoechogenic areas ($p=0.540$) in this study.

Although it has previously been hypothesized that hyperechogenic reflections consist of areas of fibrosis,⁴² we did not find a correlation of hyperechogenic reflections with any of the histological parameters (including fat and fibrosis). The analysis of ultrasonographic images, however, is based on a categorical scale and may also suffer from inter- and intra-observer differences. Several initiatives, like the European Union funded Horizon 2020 project called HarmonicSS, and a sub-taskforce of the OMERACT working group, aim to develop more objective ultrasound scoring systems, to overcome the risk of inter- and intra-observer differences.⁴⁵⁻⁴⁷

Furthermore, as described earlier, a parotid gland biopsy represents only a small part of the gland, while the whole gland is analysed by ultrasonography, and therefore a biopsy can be taken from non-affected salivary gland tissue and result in a false negative outcome (sampling error). A technique that could help to better understand the relationship between hyper- and hypoechogenic areas and histopathological features, is ultrasound-guided core needle biopsy (CNB). Recently, this technique has proven to be useful in the evaluation of suspected salivary gland lymphoma in pSS.⁴⁸ Although less tissue is collected in comparison with an open biopsy, CNB also provides biopsies with a preserved architecture. Herewith, a biopsy can be taken specifically from the location of a hyper- or hypoechogenic areas.

We showed that pSS patients with histopathological features as LELs, GCs and a plasma cell shift had higher PGUS-hypo scores. Nevertheless, presence of hypoechogenic areas does not predict whether the biopsy contains any of these features (figure 5D-F). Also, not all pSS patients with these histopathological features show hypoechogenic areas on ultrasound. This is in line with previous findings, as Hammenfors et al. showed that only 67% of pSS patients with GC-like structures in the labial gland had an abnormal ultrasound score.⁴⁹ Detection of GCs can be difficult, but usage of Bcl-6 staining results in a more robust detection of ectopic GCs in the salivary gland tissue.³⁴ In our study, the amount of infiltrate, fat and fibrosis were analysed by an objective digital image analysis (DIA) algorithm. DIA results in more reliable, objective results, and previously showed to achieve superior inter-rater agreement in the histological assessment of salivary glands of pSS patients compared to manual analyses.⁵⁰

Despite the notion that there is (some) relationship between parotid histology and ultrasonography, parotid salivary flow rates do not correlate well with histopathological features, nor with ultrasonography. In line with our results, Bookman et

al. also found poor agreement between UWSF and labial gland FS and fair agreement between SWSF and labial gland FS.²⁶ Likewise, Daniels et al. found an absolute agreement of 60% between UWSF and labial gland FS in the large Sjögren's International Collaborative Clinical Alliance (SICCA) cohort.²⁵

Yalcinkaya et al.⁵¹ recently showed that patients with $UWSF \leq 0.1 \text{ mL/min}$ had significantly higher PGUS scores compared with patients with a UWSF above the threshold. In our study, UPF did not differ significantly between patients with high (≥ 2) and low (< 2) PGUS-hypo scores. However, in unstimulated conditions, whole saliva mainly consists of the contributions of submandibular and sublingual glands,²¹ and barely consists of contributions of the parotid glands, and herewith, analyses containing either UWSF or UPF cannot be equally compared to each other. Furthermore, the parotid glands appear to be the last to show a functional decline in patients with pSS.⁴³⁻⁴⁴

In this study investigating a diagnostic cohort, parotid gland salivary flow was not related to parotid gland histopathology nor to parotid gland ultrasound scores. Parotid gland histopathology was partly related to parotid gland ultrasound, as we found moderate to good correlations between histopathological inflammatory parameters and PGUS scores. In conclusion, histopathology, salivary flow and ultrasonography are three complementary measurements and cannot directly replace each other in the workup of pSS.

References

1. Fox RI. Sjögren's syndrome. *Lancet* 2005;366:321-31.
2. Brito-Zerón P, Baldini C, Bootsma H, *et al.* Sjögren syndrome. *Nat Rev Dis Prim* 2016;2:16047.
3. 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 patients cohorts. *Ann Rheum Dis* 2017;76:9-16.
4. 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 patients cohorts. *Arthritis Rheumatol* 2017;69:35-45.
5. Spijkervet FKL, Haacke E, Kroese FGM, *et al.* Parotid gland biopsy, the alternative way to diagnose Sjögren's syndrome. *Rheum Dis Clin North Am* 2016;42 :485-99.
6. 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* 2020;72:583-90.
7. Jousse-Joulin S, Gatineau F, Baldini C, *et al.* Weight of salivary gland ultrasonography compared to other items of the 2016 ACR/EULAR classification criteria for primary Sjögren's syndrome. *J Intern Med* 2020;287:180-8.
8. Chisholm DM, Mason DK. Labial salivary gland biopsy in Sjögren's disease. *J Clin Pathol* 1968;21:656-60.
9. Fisher BA, Jonsson R, Daniels T, *et al.* Standardisation of labial salivary gland histopathology in clinical trials in primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1161-8.
10. Kroese FGM, Haacke EA, Bombardieri M. The role of salivary gland histopathology in primary Sjögren's syndrome: promises and pitfalls. *Clin Exp Rheumatol* 2018;36:222-33.
11. Risselada AP, Looije MF, Kruize AA, *et al.* The role of ectopic germinal centers in the immunopathology of primary Sjögren's syndrome: a systematic review. *Semin Arthritis Rheum* 2013;42:368-76.
12. Ihrler S, Zietz C, Sendelhofert A, *et al.* Lymphoepithelial duct lesions in Sjögren-type sialadenitis. *Virchows Arch* 1999;434:315-23.
13. Zandbelt M, Wentink J, De Wilde P, *et al.* The synergistic value of focus score and IgA% score of sublabial salivary gland biopsy for the accuracy of the diagnosis of Sjögren's syndrome: a 10-year comparison. *Rheumatology (Oxford)* 2002;41:819-23.

14. Bodeutsch C, De Wilde P, Kater L, *et al.* Quantitative immunohistologic criteria are superior to the lymphocytic focus score criterion for the diagnosis of Sjögren's syndrome. *Arthritis Rheum* 1992;35:1075-87.
15. Llamas-Gutierrez FJ, Reyes E, Martínez B, *et al.* Histopathological environment besides focus score in Sjögren's syndrome. *International Journal of Rheumatic Diseases* 2014;17:898-903.
16. Leehan KM, Pezant NP, Rasmussen A, *et al.* Minor salivary gland fibrosis in Sjögren's syndrome is elevated, associated with focus score and not solely a consequence of aging. *Clin Exp Rheumatology* 2018;36Suppl112:80-8.
17. Leehan KM, Pezant NP, Rasmussen A, *et al.* Fatty infiltration of the minor salivary gland is a selective feature of aging but not Sjögren's syndrome. *Autoimmunity* 2017;50:451-7.
18. Skarstein K, Agrawi L, Øijordsbakken G, *et al.* Adipose tissue is prominent in salivary glands of Sjögren's syndrome patients and appears to influence the microenvironment in these organs. *Autoimmunity* 2016;49:338-46.
19. Pijpe J, Kalk WWI, van der Wal JE, *et al.* Parotid gland biopsy compared with labial biopsy in the diagnosis of patients with primary Sjögren's syndrome. *Rheumatology (Oxford)* 2007;46:335-41.
20. Delli K, Vissink A, Spijkervet FKL. Salivary gland biopsy for Sjögren's syndrome. *Oral and Maxillofacial Surgery Clinics of North America* 2014;26:22-33.
21. Jensen SB, Vissink A. Salivary Gland Dysfunction and Xerostomia in Sjögren's Syndrome. *Oral Maxillofac Surg Clin NA* 2014;26:35-53.
22. Delli K, Dijkstra PU, Stel AJ, *et al.* Diagnostic properties of ultrasound of major salivary glands in Sjögren's syndrome: a meta-analysis. *Oral Dis* 2015;21:792-800.
23. Jousse-Joulin S, Milic V, Jonsson M, *et al.* Is salivary gland ultrasonography a useful tool in Sjögren's syndrome? A systematic review. *Rheumatology (Oxford)* 2015;55:1-12.
24. Hocevar A, Ambrozic A, Rozman B, *et al.* Ultrasonographic changes of major salivary glands in primary Sjögren's syndrome. Diagnostic value of a novel scoring system. *Rheumatology (Oxford)* 2005;44:768-72.
25. 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.
26. Daniels TE, Cox D, Shiboski CH, *et al.* Associations between salivary gland histopathologic diagnoses and phenotypic features of Sjögren's syndrome among 1,726 registry participants. *Arthritis Rheum* 2011;63:2021-30.
27. Bookman AAM, Shen H, Cook RJ, *et al.* Whole stimulated salivary flow: correlation with the pathology of inflammation and damage in minor salivary gland biopsy specimens from patients with primary Sjögren's syndrome but not patients with sicca. *Arthritis Rheum* 2011;63:2014-20.

28. Fisher BA, Brown RM, Bowman SJ, *et al.* A review of salivary gland histopathology in primary Sjögren's syndrome with a focus on its potential as a clinical trials biomarker. *Ann Rheum Dis* 2015;74:1645-50.
29. Atkinson JC, Travis WD, Pillemer SR, *et al.* Major salivary gland function in primary Sjögren's syndrome and its relationship to clinical features. *J Rheumatol* 1990;17:318-22.
30. Kim JW, Lee H, Park SH, *et al.* Salivary gland ultrasonography findings are associated with clinical, histological, and serologic features of Sjögren's syndrome. *Scand J Rheumatol* 2018;47:303-10.
31. Caraba A, Babalic FC, Lurciuc S, *et al.* The utility of major salivary gland ultrasonographic parameters in the diagnosis of Sjögren's syndrome. *Dis Markers* 2019;1716848
32. Cornec D, Jousse-Joulin S, Costa S, *et al.* High-grade salivary-gland involvement, assessed by histology or ultrasonography, is associated with a poor response to a single rituximab course in primary Sjögren's syndrome: Data from the TEARS randomized trial. *PLoS One* 2016;11:1-11.
33. Haacke EA, Van der Vegt B, Vissink A, *et al.* Germinal centres in diagnostic labial gland biopsies of patients with primary Sjögren's syndrome are not predictive for parotid MALT lymphoma development. *Ann Rheum Dis* 2017;76:1783-86.
34. Nakshbandi U, Haacke EA, Bootsma H, *et al.* Bcl6 for identification of germinal centres in salivary gland biopsies in primary Sjögren's syndrome. *Oral Dis* 2020;26:707-10.
35. Delli K, Arends S, Van Nimwegen JF, *et al.* Ultrasound of the major salivary glands is a reliable imaging technique in patients with clinically suspected primary Sjögren's syndrome. *Ultraschall Med* 2018;39:328-33.
36. Mossel E, Delli K, Van Nimwegen JF, *et al.* Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1883-9.
37. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
38. Burlage FR, Pijpe J, Coppes RP, *et al.* Variability of flow rate when collecting stimulated human parotid saliva. *Eur J Oral Sci* 2005;113:386-90.
39. Baum BJ. Salivary gland fluid secretion during aging. *J Am Geriatr Soc* 1989;37:453-8.
40. Kariyawasam AP, Dawes C. A circannual rhythm in unstimulated salivary flow rate when the ambient temperature varies by only about 2 degrees C. *Arch Oral Biol* 2005;50:919-22.

41. Baldini C, Ferro F, Izzetti R, *et al.* Ultra high-resolution ultrasound (UHFUS) of labial salivary glands might help avoid unnecessary lip biopsy in patients with sicca syndrome [abstract]. *Arthritis Rheumatol* 2020;72suppl10
42. Wernicke D, Hess H, Gromnica-Ihle E, *et al.* Ultrasonography of salivary glands – a highly specific imaging procedure for diagnosis of Sjögren's syndrome. *J Rheumatol* 2008;35:285-93.
43. Takashima S, Morimoto S, Tomiyama N, *et al.* Sjögren syndrome: comparison of sialography and ultrasonography. *J Clin Ultrasound* 1992;20:99-109.
44. Wang X, Bootsma H, Terpstra J, *et al.* Progenitor cell niche senescence reflects pathology of the parotid salivary gland in primary Sjögren's syndrome. *Rheumatology (Oxford)* 2020;59:3003-13.
45. 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.
46. Fisher BA, Emery P, Pitzalis C, *et al.* Response to: Can ultrasound of the major salivary glands assess histopathological changes induced by treatment with rituximab in primary Sjögren's syndrome? *Ann Rheum Dis* 2019;78:e28.
47. Harmonization and integrative analysis of regional, national and international cohorts on primary Sjögren's syndrome (pSS) towards improved stratification, treatment and health policy making (HarmonicSS). *Horizon2020 project*. Grant agreement number 731944.
48. Baer A, Grader-Beck T, Antiochos B, *et al.* Ultrasound-guided biopsy of suspected salivary gland lymphoma in Sjögren's syndrome. *Arthritis Care Res* 2020; published online first: 5 April 2020.
49. Hammenfors DS, Brun, JG, Jonsson R, *et al.* Diagnostic utility of major salivary gland ultrasonography in primary Sjögren's syndrome. *Clin Exp Rheumatol* 2015;33:56-62.
50. Lucchesi D, Pontarini E, Donati V, *et al.* The use of digital image analysis in the histological assessment of Sjögren's syndrome salivary glands improves inter-rater agreement and facilitates multicentre data harmonisation. *Clin Exp Rheumatol* 2020;38Suppl126:180-8.
51. Yalcinkaya Y, Mumcu G, Özdemir FT, *et al.* Are salivary gland ultrasonography scores associated with salivary flow rates and oral health-related quality of life in Sjögren's syndrome? *J Rheumatol* 2020;47:1774-9.

CHAPTER 5

Incorporation of salivary gland ultrasonography into the ACR-EULAR criteria for primary Sjögren's syndrome

Esther Mossel^{1*}, Jolien F van Nimwegen^{1*}, Konstantina Delli², Martha S van Ginkel¹, Alja J Stel¹, Frans GM Kroese¹, Fred KL Spijkervet², Arjan Vissink², Suzanne Arends¹, Hendrika Bootsma¹

*both authors contributed equally

¹ Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

² Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.

Edited version of Arthritis Care & Research 2020;72:583 – 590.

Abstract

Objective

To assess whether addition of salivary gland ultrasonography (SGUS) or replacement of current criteria items by SGUS influences the performance of the American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) criteria for primary Sjögren's syndrome.

Methods

Included were consecutive patients with complete data on all ACR-EULAR items (n=243) who underwent SGUS in our primary Sjögren's syndrome expertise center. Clinical diagnosis by the treating physician was used as gold standard. Separate analyses were performed for patients who underwent labial or parotid gland biopsies. The average score for hypoechogenic areas in 1 parotid and 1 submandibular gland was determined (range 0-3). Next, performance of the ACR-EULAR criteria was evaluated after addition of SGUS or replacement of current items by SGUS.

Results

Receiver operating characteristic analysis showed an optimal cut-off value of ≥ 1.5 for SGUS. The optimal weight for SGUS positivity was 1. Cut-off for ACR-EULAR fulfilment remained ≥ 4 . In patients who underwent a labial gland biopsy (n=124), the original criteria showed an area under the curve (AUC) of 0.965, sensitivity of 95.9%, and specificity of 92.2%. After addition of SGUS, AUC was 0.966, with a sensitivity of 97.3% and specificity of 90.2%. In patients who underwent a parotid gland biopsy (n=198), similar results were found. Sensitivity of the criteria decreased substantially when SGUS replaced salivary gland biopsy or anti-SSA/Ro antibodies, while performance remained equal when SGUS replaced ocular staining score, Schirmer's test or unstimulated whole saliva flow.

Conclusion

Validity of the ACR-EULAR criteria remains high after incorporation of SGUS. With SGUS, clinicians are offered a larger array of tests to evaluate fulfilment of the ACR-EULAR criteria.

Introduction

Primary Sjögren's syndrome (pSS) is a common systemic autoimmune disease affecting the exocrine glands, manifesting as keratoconjunctivitis sicca (dry eyes) and xerostomia (dry mouth).^{1,2} Patients also often experience fatigue and several extra-glandular manifestations.²

Several classification criteria sets for pSS have been developed during recent years. Of these, the 2002 American European Consensus Group (AECG) criteria have most often been used in daily clinical practice for many years.³⁻⁵ Cornerstones of this criteria set are a focus score ≥ 1 in a salivary gland biopsy and presence of anti-SSA/Ro/anti-SSB/La antibodies.³ These criteria take both subjective sicca symptoms and objective measures for the ocular and oral symptoms into account so that equal weights are assigned to the oral and ocular components.³ However, the AECG criteria have not been endorsed by the American College of Rheumatology (ACR) and European league against Rheumatism (EULAR).^{4,5} In an effort to reach international consensus regarding classification criteria for pSS, recently the 2016 ACR-EULAR criteria were developed, consisting of items from the 2002 AECG and the 2012 ACR criteria.³⁻⁶ Both EULAR and the ACR have endorsed the ACR-EULAR criteria, and the criteria have been validated in multiple external cohorts.⁷⁻⁹ The ACR-EULAR criteria show high sensitivity and specificity regardless of the type of biopsy (parotid or labial) taken to assess the salivary gland focus score.⁸

Upon a closer look at the ACR-EULAR criteria, a few key points become evident. First, salivary gland histopathology and the presence of anti-SSA/Ro antibodies deservedly remain cornerstones in the classification of pSS. Second, tear gland involvement is measured using a functional test (Schirmer's test) and by imaging of structural damage of the ocular surface (ocular staining score [OSS]), while salivary gland involvement is only evaluated using a functional test (unstimulated whole saliva flow [UWS]). Removal of sialography and scintigraphy from the criteria is an advantage of the ACR-EULAR criteria, considering the invasiveness and limited validity of these procedures.^{4,10-12} However, the ACR-EULAR criteria now lack a test that measures structural salivary gland damage.

Currently, B-mode salivary gland ultrasonography (SGUS) is increasingly applied to assess structural changes of the salivary glands in pSS. SGUS is non-invasive, non-irradiating, inexpensive, relatively easy to perform in an outpatient setting, and can be repeated for follow-up. Previous studies have demonstrated that SGUS has good accuracy to differentiate pSS from non-pSS.^{9,13-17} Many scoring systems are applied for SGUS, but recent analyses have shown that limiting scoring to hypoechoic areas in both the submandibular and parotid gland on 1 side suffices

for accurate differentiation between pSS and non-pSS.¹⁸ Scoring of hypoechogenic areas showed good intra- and interobserver reliability.^{19,20} This reduction of the scoring system further increases the feasibility of the technique for common application in a diagnostic setting.

In clinical cohort studies, the addition of SGUS to the AECG and ACR criteria has been shown to increase the sensitivity of these criteria, with a minor decrease in their specificity.^{10,21} Unfortunately, SGUS was not tested as a new diagnostic technique in the cohorts in which the ACR-EULAR criteria were developed and validated and not considered to be included in the criteria. Therefore, our primary objective was to assess whether the presence of hypoechogenic areas on SGUS as a criteria item influences the performance of the ACR-EULAR criteria. The second objective was to evaluate the performance of the ACR-EULAR criteria when replacing current items with SGUS. Both objectives were evaluated in a large cohort of patients clinically suspected of pSS.

Patients and Methods

Patients

The study population for this cohort study consisted of all eligible consecutive patients who underwent an SGUS examination in the University Medical Center Groningen between October 2014 and July 2017. SGUS was performed as a routine diagnostic imaging technique in new patients clinically suspected of pSS as well as during baseline visits of pSS patients included in the Efficacy and Safety of Abatacept in Patients with Primary Sjögren's Syndrome phase III trial (NCT02067910) or the REgistry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) observational cohort study.

Exclusion criteria were age <18 years, the presence of an associated systemic auto-immune disease (e.g., rheumatoid arthritis, systemic lupus erythematosus), or current use of biological disease-modifying antirheumatic drugs. Patients lacking a clinical diagnosis and patients with an incomplete diagnostic evaluation according to the ACR-EULAR criteria were also excluded. The clinical diagnosis by experienced treating physicians was used as the gold standard in all analyses. If a diagnosis was not clearcut, consensus was achieved by consulting at least 1 other experienced physician. This study was conducted in accordance with the Declaration of Helsinki. The research protocol was approved by the Medical Ethics Committee of the University Medical Center Groningen (METc 2018/309), which waived the requirement of written informed consent.

Salivary gland ultrasound

SGUS was performed using the MyLabSeven scanner (Esaote, Genova, Italy), equipped with a high-resolution linear probe (4-13 MHz). All SGUS images were scored by AJS, KD or JFvN, who previously showed good inter-observer agreement when scoring hypoechogenic areas.¹⁹ Median intraclass correlation coefficients were 0.74 for parotid glands and 0.71 for submandibular glands. The presence of hypoechogenic areas was scored as: 0 = no hypoechogenic areas, 1 = a few scattered areas, 2 = several areas, and 3 = numerous hypoechogenic areas.¹⁷ The average score for the presence of hypoechogenic areas (range 0-3) in the submandibular and parotid gland on the right side was determined, which was previously shown to accurately differentiate between pSS and non-pSS.¹⁸ If the right parotid or submandibular gland could not be scored (e.g., because of previous removal of that gland), scores of the left side were used. Receiver operating characteristic (ROC) analysis was used to determine the optimal cut-off value for SGUS to identify patients who were clinically diagnosed with pSS by the treating physicians by choosing the cut-off for which the sum of sensitivity and specificity was the highest.

Classification according to the original ACR-EULAR criteria

All included patients had been subjected to a complete multidisciplinary evaluation according to the ACR-EULAR criteria,^{4,5} including a labial gland biopsy, a parotid gland biopsy, or both. Separate analyses were performed, in which classification according to the ACR-EULAR criteria was determined using the outcomes of either labial or parotid gland biopsy results. Patients who underwent both a labial and parotid gland biopsy were included in both analyses, with the results of either their labial or parotid gland biopsy being used to determine ACR-EULAR classification.

Incorporation of SGUS into the ACR-EULAR criteria

SGUS positivity was added as an item to the ACR-EULAR criteria. To keep the original criteria applicable, the weight of the original criteria items was kept as they were, i.e., 3 points for the presence of anti-SSA/Ro antibodies and a focus score ≥ 1 ; and 1 point for an abnormal UWS, Schirmer's test and OSS score.^{4,5} To select the optimal weight of SGUS, separate analyses of the performance of the modified ACR-EULAR criteria were performed, assigning a weight of either 1, 2, or 3 points for a positive SGUS.

Replacement of current ACR-EULAR criteria items by SGUS

Next, 5 additional criteria sets were developed, in which SGUS replaced 1 of the current items. The weight of the original items was again kept equal to the original criteria, and the optimal weight of the SGUS item was determined by doing separate analyses using a weight of 1, 2, or 3 points for a positive SGUS.

Statistical analysis

Statistical analyses were executed using IBM SPSS Statistics 23 (SPSS, Chicago, IL, USA). ROC analysis was performed to determine the accuracy of the original ACR-EULAR score, the ACR-EULAR score with the addition of SGUS, and the ACR-EULAR score with SGUS as replacement of original items to predict the clinical diagnosis. The area under the curve (AUC) was interpreted as no discrimination (0-0.5), or with the accuracy judged as poor (0.5-0.7), fair (0.7-0.8), good (0.8-0.9) or excellent (0.9-1.0).²² Optimal cut-off values of the different ACR-EULAR scores were determined by choosing the cut-off for which the sum of sensitivity and specificity was the highest. Patients were then classified according to this cut-off for the original and modified criteria sets. Finally, absolute agreement, sensitivity and specificity of the original and modified ACR-EULAR criteria sets, with clinical diagnosis as the gold standard, were determined and compared.

Results

SGUS was performed in 363 patients. Of these, 243 patients were eligible for inclusion (figure 1). Of the 243 included patients, 45 patients underwent only a labial gland biopsy, 119 patients underwent only a parotid gland biopsy, and 79 patients underwent both a labial and parotid gland biopsy. Including the patients who underwent both biopsies, 124 patients underwent a labial biopsy, and 198 patients underwent a parotid gland biopsy.

Characteristics of pSS and non-pSS patients are shown in table 1. All included patients fulfilled the entry criteria of the ACR-EULAR criteria. The characteristics of the patients who underwent a labial gland biopsy were similar to those of the patients who underwent a parotid gland biopsy (data not shown). The median time between SGUS and salivary gland biopsies was 7 months for labial gland biopsies and 6 months for parotid gland biopsies.

Table 1.

Characteristics of study population.

Characteristics	pSS (n=147)	Non-pSS (n=96)
Age	53 ±14	52 ±14
Female	131 (89)	81 (84)
SGUS score	2.0 [1.0;2.5]	0.5 [0.5;1.0]
SGUS score ≥1.5	106 (72)	10 (10)
FS≥1 in labial gland biopsy*	64 (88)	5 (10)
FS≥1 in parotid gland biopsy**	89 (76)	2 (2)
Anti-SSA/Ro positive	125 (85)	9 (9)
Ocular Staining Score ≥5	70 (48)	11 (12)
Schirmer ≤5 mm/5 min	113 (77)	56 (58)
Unstimulated whole saliva ≤0.1 ml/min	105 (71)	42 (44)

Values are presented as mean ± SD, median [interquartile range] or number (%)

*n=124 of which 73 are pSS and 51 are non-pSS.

**n=198 of which 117 are pSS and 81 are non-pSS.

pSS = primary Sjögren's syndrome. SGUS = salivary gland ultrasound. FS = focus score (foci/4mm²).

Figure 1. Flow chart of patient inclusion.

bDMARDS = biological disease modifying anti-rheumatic drugs; pSS = primary Sjögren's syndrome;

SGUS = salivary gland ultrasound

* Missing items were: salivary gland biopsy (n=72); Schirmer's test (n=4); OSS (n=2) and UWS (n=2).

In the majority of these patients, either the patients could be classified as pSS without the need of a positive salivary gland biopsy, or a positive biopsy would not have resulted in a clinical diagnosis of pSS.

Performance of SGUS

The accuracy of SGUS to predict clinical diagnosis was good, with an AUC of 0.860 (95% confidence interval [95% CI] 0.821-0.900), and an optimal cut-off value of ≥ 1.5 . SGUS was therefore considered positive when the average score for the presence of hypoechoic areas in 1 parotid and 1 submandibular gland was ≥ 1.5 . Based on this cut-off point, SGUS was positive in 106 pSS and 6 non-pSS patients and negative in 41 pSS and 90 non-pSS patients. Absolute agreement with clinical diagnosis was 80.7%, sensitivity was 72.1% and specificity was 93.8%.

Performance of ACR-EULAR criteria with addition of SGUS

Supplementary table 1A-B shows the performance of the ACR-EULAR criteria when SGUS was added to the criteria, using a weight of 1, 2, or 3 for a positive SGUS. The performance of the ACR-EULAR criteria including SGUS was highest when a positive SGUS was assigned a weight of 1 point. The optimal cut-off point of the original ACR-EULAR score to discriminate between pSS and non-pSS was confirmed to be ≥ 4 . After the addition of SGUS to the ACR-EULAR criteria with a weight of 1 point, the optimal cut-off point of the modified ACR-EULAR score to discriminate between

pSS and non-pSS remained ≥ 4 (supplementary table 1A-B). Based on these results, in the following analyses a cut-off of ≥ 4 was used for the original and modified ACR-EULAR score. A positive SGUS results in an increase of 1 point in the modified ACR-EULAR score (table 2A-B).

Table 2.

Original and modified ACR-EULAR criteria incorporating salivary gland ultrasound.

Item	Weight
A. Original ACR-EULAR criteria	
Focal lymphocytic sialadenitis and FS ≥ 1	3 points
Anti-SSA/Ro positive	3 points
OSS ≥ 5 in at least 1 eye	1 point
Schirmer's test ≤ 5 mm/5min in at least 1 eye	1 point
UWS flow rate ≤ 0.1 ml/min	1 point
B. Modified ACR-EULAR criteria: addition of ultrasound	
Focal lymphocytic sialadenitis and FS ≥ 1	3 points
Anti-SSA/Ro positive	3 points
OSS ≥ 5 in at least 1 eye	1 point
Schirmer's test ≤ 5 mm/5min in at least 1 eye	1 point
UWS flow rate ≤ 0.1 ml/min	1 point
Average SGUS score for hypoechogenic areas ≥ 1.5	1 point

For both sets, patients with a score of ≥ 4 are classified as pSS.

FS = focus score (foci/4mm²); pSS = primary Sjögren's syndrome; OSS= ocular staining score; UWS = unstimulated whole saliva flow; SGUS = salivary gland ultrasound.

In patients who underwent a labial gland biopsy (n=124), the original ACR-EULAR criteria showed an AUC of 0.965 (95% CI 0.932 – 0.997) to predict clinical diagnosis (figure 2). Absolute agreement with clinical diagnosis was 94.4%, sensitivity was 95.9%, and specificity was 92.2%. After the addition of SGUS, the modified ACR-EULAR criteria showed an AUC of 0.966 (95% CI 0.934 – 0.998), absolute agreement remained the same, sensitivity slightly increased to 97.3%, and specificity slightly decreased to 88.2%.

The same analyses were performed in patients who underwent a parotid gland biopsy (n=198), and similar results were found (figure 2). In this group, the original criteria showed an AUC of 0.954 (95% CI 0.925 – 0.984) to predict clinical diagnosis. Absolute agreement with clinical diagnosis was 92.9%, sensitivity was 91.4%, and specificity was 95.1%. After addition of SGUS, the modified ACR-EULAR criteria showed an AUC of 0.964 (95% CI 0.939 – 0.989), absolute agreement remained the same, sensitivity slightly increased to 92.3%, and specificity slightly decreased to 93.8%.

To summarize, the addition of SGUS to the ACR-EULAR criteria resulted in negligible changes in the performance of the criteria and did not change its optimal cut-off point.

Figure 2. ROC curves of the original ACR-EULAR criteria and adjusted criteria with addition of SGUS. SGUS = salivary gland ultrasound

Performance of ACR-EULAR criteria with replacement of items by SGUS

For the following analysis, 5 modified sets of criteria were used, in which one of the original items was replaced with SGUS. When SGUS replaced current criteria items in patients who underwent a labial gland biopsy (n=124), the optimal weight for SGUS was again 1 point, regardless of which original criteria items was replaced by SGUS (supplementary table 2). The optimal cut-off point to discriminate between pSS and non-pSS remained ≥ 4 .

Table 3.

Performance of the original and modified ACR-EULAR criteria sets with SGUS replacing current items.

	AUC	95% CI	Agreement	Sensitivity	Specificity
A. Patients with labial gland biopsy (n=124)					
Original ACR-EULAR criteria	0.965	0.932-0.997	94.4%	95.9%	92.2%
SGUS replacing labial gland biopsy	0.903	0.849-0.957	87.9%	82.2%	94.1%
SGUS replacing anti-SSA/Ro antibodies	0.943	0.902-0.985	89.5%	86.3%	94.1%
SGUS replacing OSS	0.964	0.931-0.996	93.5%	95.9%	88.2%
SGUS replacing Schirmer's test	0.969	0.938-1.000	93.5%	94.5%	92.2%
SGUS replacing UWS	0.967	0.937-0.998	93.5%	97.3%	88.2%
B. Patients with parotid gland biopsy (n=198)					
Original ACR-EULAR criteria	0.954	0.925-0.984	92.9%	91.4%	95.1%
SGUS replacing parotid gland biopsy	0.925	0.887-0.962	88.4%	83.8%	95.1%
SGUS replacing anti-SSA/Ro antibodies	0.918	0.879-0.956	86.9%	79.5%	97.5%
SGUS replacing OSS	0.964	0.938-0.990	93.4%	92.3%	95.1%
SGUS replacing Schirmer's test	0.964	0.939-0.989	89.9%	84.6%	97.5%
SGUS replacing UWS	0.969	0.946-0.992	92.9%	90.6%	96.3%

In all criteria sets a weight of 1 point for SGUS and cut-off value of ≥ 4 for fulfilment of the criteria was used.
 OSS= Ocular Staining Score; UWS = unstimulated whole saliva flow; SGUS = salivary gland ultrasound

When SGUS replaced the labial gland biopsy or anti-SSA/Ro antibodies, there was a considerable decrease in accuracy and sensitivity, while there was only a slight decrease in specificity compared to the original criteria (table 3A, figure 3A). On the other hand, when SGUS replaced the OSS, Schirmer's test or UWS, no major changes in accuracy, sensitivity and specificity occurred.

The same analyses were performed in patients who underwent a parotid gland biopsy (n=198). When SGUS replaced the OSS, Schirmer's test or UWS, the optimal weight for SGUS was again 1 point (supplementary table 3), with only minor changes in sensitivity and specificity and even an increase in accuracy (table 3B, figure 3B). When SGUS replaced the parotid gland biopsy or anti-SSA/Ro antibodies, the optimal weight for SGUS would be 3 points. However, regardless of whether SGUS was assigned 1, 2, or 3 points in these analyses, accuracy of the ACR-EULAR criteria dropped substantially (supplementary table 3).

To summarize, SGUS can replace the OSS, Schirmer's test, or UWS in the classification of pSS without major changes in the performance of the criteria. The salivary gland biopsy or the measurement of anti-SSA/Ro antibodies, on the other hand, cannot be completely replaced by SGUS, since this led to a considerable decrease in the performance of the criteria.

Figure 3. ROC curves of the original ACR-EULAR criteria and adjusted criteria with replacement of original items by SGUS.

A. ACR-EULAR criteria including labial gland biopsy outcome. B. ACR-EULAR criteria including parotid gland biopsy outcome. SGUS = salivary gland ultrasound

Discussion

In this large clinical cohort study, we aimed to investigate the performance of the ACR-EULAR criteria when a positive SGUS was added to the criteria. The performance of the ACR-EULAR criteria was best when SGUS was assigned a weight of 1 point.

In the first part of this study, we showed that addition of SGUS to the ACR-EULAR criteria only marginally increased sensitivity and marginally decreased specificity, while overall accuracy remained the same. Although the addition of SGUS did not improve the accuracy of the ACR-EULAR criteria in our cohort, it improves their feasibility in clinical practice, by allowing rheumatologists to choose from a larger array of tests.

Previously, two other studies incorporated SGUS in the ACR-EULAR classification criteria.^{23,24} In the study by Le Goff et al.,²³ in which the AECG and ACR-EULAR classification criteria were compared, the addition of SGUS to the ACR-EULAR criteria was also investigated. However, the authors arbitrarily assigned a weight of 1 point to a positive SGUS and used the same cut-off value as the original ACR-EULAR criteria (i.e., ≥ 4).

In our study, we confirmed statistically that the optimal weight to assign to SGUS was indeed 1 point and that the optimal cut-off value to classify a patient as having pSS remained ≥ 4 . In the study by Le Goff et al.,²³ similar results were found regarding the performance of the ACR-EULAR criteria after addition of SGUS, i.e., sensitivity was slightly increased and specificity slightly decreased.

In the study by Takagi et al.,²⁴ the weight of the original criteria items was also kept. In contrast to our study, 3 points were assigned to SGUS positivity, and the optimal cut-off point to discriminate between SS and non-SS increased to ≥ 5 . The combined ACR-EULAR and SGUS scoring system showed an improved accuracy compared to the original criteria. Unfortunately, a fair comparison between the study of Takagi et al.²⁴ and ours cannot be made since the methodology of their study differed greatly from ours. Importantly, complete data regarding the ACR-EULAR items was only available in a small subset of the included patients (62 out of 213 patients); Saxon's test, which measures stimulated whole saliva, was used instead of UWS, and patients with secondary SS were not excluded. Furthermore, a different, more complicated SGUS score was used.

In the second part of this study, the performance of the ACR-EULAR criteria was evaluated when SGUS replaced current classification items. We found that SGUS could replace the OSS, Schirmer's test, or UWS in the classification of pSS without decreasing the accuracy of the ACR-EULAR criteria. However, when SGUS replaced the salivary gland biopsy in the classification of pSS or the measurement of anti-SSA/Ro antibodies, the performance of the criteria significantly decreased.

In a previous study, we showed that the combination of a positive SGUS and presence of anti-SSA/Ro antibodies had a positive predictive value of 97% for classification as pSS, according to the ACR-EULAR criteria.¹⁴ Based on these results, Mossel et al. suggested that for classification purposes, the first step of a classification evaluation could be SGUS and determination of anti-SSA/Ro positivity. When both are positive, patients can already be classified as pSS. The current study confirms these results, because the combination of anti-SSA/Ro positivity and SGUS positivity is indeed enough for fulfillment of the adjusted ACR-EULAR criteria. As the next step in the evaluation for classification, we recommend a salivary gland biopsy since the sensitivity of the ACR-EULAR criteria decreased substantially when the salivary gland biopsy was completely replaced by SGUS. When it comes to clinically diagnosing a patient with pSS, on the other hand, we prefer a full evaluation, including SGUS and as many items of the ACR-EULAR criteria as possible, to allow a clinician to decide on the best possible treatment for that particular patient.

When SGUS is added to the ACR-EULAR criteria, the cut-off of 4 points can be fulfilled solely based on the Schirmer's test, OSS, UWS, and SGUS. In our database, this combination only occurred in 1 patient, who was clinically diagnosed as non-SS. Therefore, we cannot draw a definite conclusion about the validity of the ACR-EULAR criteria in this specific subgroup. Based on our expert opinion, we would recommend only classifying such a patient as pSS if a positive biopsy result or anti-SSA/Ro antibodies are also present until there is more data available regarding this subgroup.

In this study, we used a simplified SGUS scoring system, similar to the ones used by other groups.^{15,21,23} However, the lack of a consensus scoring system complicates the incorporation of SGUS into the ACR-EULAR criteria. Jousse-Joulin et al.²⁰ recently published an atlas with consensual definitions of SGUS abnormalities. The next step will be to agree on a consensus scoring system with a validated cut-off. As soon as a validated cut-off is set, SGUS hopefully will be incorporated into the ACR-EULAR criteria. The addition of SGUS, as a measure for structural damage of the salivary glands, would balance the ACR-EULAR criteria by including 2 items to measure tear as well as salivary gland involvement.

A strong point of our study is the use of a large cohort of patients from daily clinical practice, including pSS as well as non-pSS sicca patients, with complete data for all ACR-EULAR items. Furthermore, analyses were performed separately for patients who underwent a labial and/or a parotid gland biopsy, which makes our data relevant to all diagnostic centers, regardless of the type of biopsy performed. A potential limitation of the study is the use of clinical diagnosis performed by expert clinicians working in a tertiary referral center for pSS, instead of expert consensus,

as the gold standard. However, using expert consensus as the gold standard would also have introduced bias depending on the familiarity of the experts with SGUS in pSS.

In conclusion, the validity of the ACR-EULAR criteria remains high after incorporation of SGUS. SGUS is non-invasive, non-irradiating, inexpensive, and relatively easy to perform in an outpatient setting and could replace OSS, Schirmer's test, or UWS in centers with less access to these tests. Incorporation of SGUS into the ACR-EULAR criteria improves their feasibility in clinical practice by allowing rheumatologists to choose from a larger array of tests. The modified criteria enable a step-wise approach for classification, starting with determination of anti-SSA/Ro-antibodies and SGUS, which decreases the number of invasive salivary gland biopsies needed for classification.

Funding:

This work was supported by: the Dutch Arthritis Foundation (research grant for the Registry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) cohort); Bristol-Myers Squibb (unrestricted research grant for the RESULT cohort and ASAPIII trial, (NCT02067910); Horizon 2020, a research project supported by European Commission [H2020-SC1-2016-RTD, proposal 731944]; National Institutes of Health NIDCR 11-074-E Salivary Biomarkers for Sjögren's Syndrome Detection – A Multi-Center Study.

Acknowledgements

Part of this study has been presented at the ISSS 2018* in Washington and at the EULAR 2018** in Amsterdam, both as a poster, and the corresponding abstracts have been published as a conference abstract.

* Van Nimwegen JF, Mossel E, Van Ginkel MS, et al. P-30: Addition of salivary gland ultrasound increases the feasibility of the ACR-EULAR classification criteria in primary Sjögren's syndrome. Clin Exp Rheumatol, 2018 May-Jun; 36 Suppl 112(3): 256-332.

** Mossel E, van Nimwegen JF, van Ginkel MS, et al. The ACR-EULAR classification criteria in primary Sjögren's syndrome: the contributing role of ultrasound. Ann Rheum Dis, 77, supplement Suppl, 2018, page A1075.

References

1. Brito-Zeron P, Baldini C, Bootsma H, *et al.* Sjögren's syndrome. *Nat Rev Dis Prim* 2016;2:16047
2. Mariëtte X, Criswell LA. Primary Sjögren's Syndrome. *N Engl J Med* 2018;378:931-9.
3. 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.
4. Shiboski CH, Shiboski SC, Seror R, *et al.* 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:9-16.
5. 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.
6. 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 (Hoboken)* 2012;64:475-87.
7. Tsuboi H, Hagiwara S, Asashima H, *et al.* Comparison of performance of the 2016 ACR-EULAR classification criteria for primary Sjögren's syndrome with other sets of criteria in Japanese patients. *Ann Rheum Dis* 2017;76:1980-5.
8. Van Nimwegen JF, van Ginkel MS, Arends S, *et al.* Validation of the ACR-EULAR criteria for primary Sjögren's syndrome in a Dutch prospective diagnostic cohort. *Rheumatology (Oxford)* 2018;57:818-25.
9. Delli K, Dijkstra PU, Stel AJ, *et al.* Diagnostic properties of ultrasound of major salivary glands in primary Sjögren's syndrome: a meta-analysis. *Oral Dis* 2015;21:792-800.
10. Cornec D, Jousse-Joulin S, Pers, J-O, *et al.* Contribution of salivary gland ultrasonography to the diagnosis of Sjögren's syndrome: Toward new diagnostic criteria? *Arthritis Rheum* 2013;65:216-25.
11. Takagi Y, Kimura Y, Nakamura H, *et al.* Salivary gland ultrasonography: can it be an alternative to sialography as an imaging modality for Sjögren's syndrome? *Ann Rheum Dis* 2010;69:1321-4.

12. 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 (Oxford)* 2012;51:1081-5.
13. Jousse-Joulin S, Milic V, Jonsson MV, *et al.* Is salivary gland ultrasonography a useful tool in Sjögren's syndrome? A systematic review. *Rheumatology (Oxford)* 2016;55:789-800.
14. Mossel E, Delli K, van Nimwegen JF, *et al.* Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1883-9.
15. Theander E, Mandl T. Primary Sjögren's Syndrome: Diagnostic and Prognostic Value of Salivary Gland Ultrasonography Using a Simplified Scoring System. *Arthritis Care Res (Hoboken)* 2014;66:1102-7.
16. Baldini C, Luciano N, Tarantini G, *et al.* Salivary gland ultrasonography: a highly specific tool for the early diagnosis of primary Sjögren's syndrome. *Arthritis Res Ther* 2015;17:146.
17. Hocevar A, Ambrozic A, Rozman B, *et al.* Ultrasonographic changes of major salivary glands in primary Sjögren's syndrome. Diagnostic value of a novel scoring system. *Rheumatology (Oxford)* 2005;44:768-72.
18. 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.
19. Delli K, Arends S, van Nimwegen JF, *et al.* Ultrasound of the Major Salivary Glands is a Reliable Imaging Technique in Patients with Clinically Suspected Primary Sjögren's syndrome. *Ultraschall Med* 2018;39:328-33.
20. 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.
21. 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.
22. Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. Third edition 2013.

23. 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.
24. 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.

Supplementary table 1.

Optimal weight of SGUS in the ACR-EULAR classification criteria and optimal cut-off of the modified ACR-EULAR score.

Weight SGUS	AUC	95% CI	Cut-off point	Sensitivity	Specificity
A. Modified ACR-EULAR criteria incl. labial gland biopsy					
1 point	0.966	0.934-0.998	3	97.3%	78.4%
			4	97.3%	88.2%
			5	84.9%	94.1%
2 points	0.965	0.932-0.997	3	97.3%	78.4%
			4	97.3%	84.3%
			5	86.3%	92.2%
3 points	0.962	0.929-0.995	3	97.3%	76.5%
			4	97.3%	84.3%
			5	86.3%	88.2%
B. Modified ACR-EULAR criteria incl. parotid gland biopsy					
1 point	0.964	0.939-0.989	3	94.0%	86.4%
			4	92.3%	93.8%
			5	82.1%	98.8%
2 points	0.967	0.943-0.991	3	96.6%	85.2%
			4	92.3%	91.4%
			5	82.9%	97.5%
3 points	0.965	0.941-0.990	3	96.6%	85.2%
			4	94.9%	90.1%
			5	82.9%	95.1%

Analysis of the optimal weight of SGUS and optimal cut-off of the modified ACR-EULAR score when SGUS is added to the ACR-criteria, based on the best combination of sensitivity and specificity. The optimal weight and cut-off point are highlighted in bold.
SGUS = salivary gland ultrasound.

Supplementary Table 2.

Optimal weight and cut-off after replacement of items with SGUS in patients with labial gland biopsies.

Weight SGUS	AUC	95% CI	Cut-off point	Sensitivity	Specificity
A. SGUS replacing anti-SSA/Ro antibodies					
1 point	0.943	0.902 – 0.985	3	91.8%	86.3%
			4	86.3%	94.1%
			5	69.9%	100%
2 points	0.941	0.898 – 0.983	3	93.2%	86.3%
			4	87.7%	90.2%
			5	71.2%	98.0%
3 points	0.936	0.892 – 0.980	3	93.2%	84.3%
			4	89.0%	90.2%
			5	72.6%	94.1%
B. SGUS replacing labial gland biopsy					
1 point	0.903	0.849 – 0.957	3	84.9%	88.2%
			4	82.2%	94.1%
			5	72.6%	94.1%
2 points	0.911	0.860 – 0.963	3	87.7%	88.3%
			4	82.2%	90.2%
			5	72.6%	94.1%
3 points	0.910	0.858 – 0.962	3	89.0%	84.3%
			4	84.9%	90.2%
			5	72.6%	90.2%
C. SGUS replacing Ocular Staining Score					
1 point	0.964	0.931 – 0.996	3	97.3%	80.4%
			4	95.9%	88.2%
			5	83.6%	94.1%

Weight SGUS	AUC	95% CI	Cut-off point	Sensitivity	Specificity
2 points	0.963	0.931 – 0.996	3	97.3%	78.4%
			4	95.9%	86.3%
			5	86.3%	92.2%
3 points	0.961	0.927 – 0.994	3	97.3%	76.5%
			4	95.9%	84.3%
			5	86.3%	90.2%

D. SGUS replacing Schirmer's I Test

1 point	0.969	0.938 – 1.000	3	97.3%	82.4%
			4	94.5%	92.2%
			5	80.8%	100%
2 points	0.969	0.937 – 1.000	3	97.3%	78.4%
			4	94.5%	92.2%
			5	84.9%	98.0%
3 points	0.965	0.932 – 0.998	3	97.3%	76.5%
			4	94.5%	88.2%
			5	84.9%	98.0%

E. SGUS replacing Unstimulated Whole Saliva Flow

1 point	0.967	0.937 – 0.997	3	97.3%	78.4%
			4	97.3%	88.2%
			5	84.9%	94.1%
2 points	0.964	0.933 – 0.995	3	97.3%	78.4%
			4	90.4%	86.3%
			5	84.9%	98.0%
3 points	0.961	0.929 – 0.993	3	97.3%	76.5%
			4	90.4%	84.3%
			5	84.9%	96.1%

Analysis of the optimal weight of SGUS and optimal cut-off of the modified ACR-EULAR score when current items are replaced by SGUS in patients who underwent labial biopsies, based on the best combination of sensitivity and specificity. The optimal weight and cut-off point are highlighted in bold. SGUS = salivary gland ultrasound

Supplementary Table 3.

Optimal weight and cut-off after replacement of items with SGUS in patients with parotid gland biopsies.

Weight SGUS	AUC	95% CI	Cut-off point	Sensitivity	Specificity
A. SGUS replacing anti-SSA/Ro antibodies					
1 point	0.918	0.879 – 0.956	3	83.8%	92.6%
			4	79.5%	97.5%
			5	65.0%	100%
2 points	0.922	0.884 – 0.960	3	87.2%	91.4%
			4	81.2%	95.1%
			5	71.8%	98.8%
3 points	0.921	0.883 – 0.959	3	87.2%	91.4%
			4	84.6%	93.8%
			5	73.5%	96.3%
B. SGUS replacing parotid gland biopsy					
1 point	0.925	0.887 – 0.962	3	86.3%	88.9%
			4	83.8%	95.1%
			5	70.1%	98.8%
2 points	0.935	0.900 – 0.960	3	91.5%	87.7%
			4	84.6%	92.6%
			5	71.8%	97.5%
3 points	0.935	0.900 – 0.970	3	92.3%	87.7%
			4	89.7%	91.4%
			5	72.6%	95.1%
C. SGUS replacing Ocular Staining Score					
1 point	0.964	0.938 – 0.990	3	93.2%	88.9%
			4	92.3%	95.1%
			5	80.3%	98.8%

Weight SGUS	AUC	95% CI	Cut-off point	Sensitivity	Specificity
2 points	0.966	0.941 – 0.991	3	95.7%	86.4%
			4	92.3%	92.6%
			5	82.9%	98.8%
3 points	0.964	0.939 – 0.990	3	95.7%	86.4%
			4	94.9%	90.1%
			5	82.9%	96.3%

D. SGUS replacing Schirmer's I Test

1 point	0.964	0.939 – 0.989	3	93.2%	90.1%
			4	84.6%	97.5%
			5	76.9%	100%
2 points	0.966	0.943 – 0.990	3	94.0%	86.4%
			4	84.6%	96.3%
			5	79.5%	100%
3 points	0.965	0.940 – 0.989	3	95.7%	86.4%
			4	85.5%	92.6%
			5	79.5%	98.8%

E. SGUS replacing Unstimulated Whole Saliva Flow

1 point	0.969	0.946 – 0.992	3	93.2%	88.9%
			4	90.6%	96.3%
			5	78.6%	100%
2 points	0.967	0.943 – 0.991	3	96.6%	85.2%
			4	92.3%	91.4%
			5	82.9%	97.5%
3 points	0.969	0.947 – 0.992	3	95.7%	86.4%
			4	92.3%	92.6%
			5	79.5%	97.5%

Analysis of the optimal weight of SGUS and optimal cut-off of the modified ACR-EULAR score when current items are replaced by SGUS in patients who underwent parotid gland biopsies, based on the best combination of sensitivity and specificity. The optimal weight and cut-off point are highlighted in bold. SGUS = salivary gland ultrasound.

CHAPTER 6

Clinical phenotyping of primary Sjögren's patients using salivary gland ultrasonography – data from the RESULT cohort

Esther Mossel¹, Jolien F van Nimwegen¹, Alja J Stel¹, Robin F Wijnsma¹, Konstantina Delli², Greetje S van Zuiden¹, Lisette Olie³, Jelle Vehof³, Leonoor I Los³, Arjan Vissink², Frans GM Kroese¹, Suzanne Arends¹, Hendrika Bootsma¹

¹ Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

² Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

³ Department of Ophthalmology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Edited version of *The Journal of Rheumatology* 2020; published online first: 1 October 2020

Abstract

Objective

To investigate salivary gland ultrasound (SGUS) abnormalities in relation to clinical phenotype and patient characteristics, disease activity and disease damage in patients with primary Sjögren's syndrome (pSS).

Methods

Consecutive outpatients included in our REgistry of Sjögren Syndrome LongiTudinal (RESULT) cohort were selected. Included pSS patients were classified according to the ACR-EULAR criteria and underwent full ultrasonographic examination (Hocevar score 0-48) at baseline. Total SGUS scores of ≥ 15 were considered positive. Patient characteristics, disease activity and disease damage were compared between the different SGUS groups.

Results

In total, 172/186 pSS patients were eligible, of whom 136 (79%) were SGUS positive. SGUS positive patients had significantly longer disease duration, higher ESSDAI, higher SSDDI, more often a positive parotid gland biopsy, anti-SSA/Ro/SSB/La antibodies, abnormal unstimulated whole saliva (UWS) and ocular staining score (OSS), and higher levels of IgG and rheumatoid factor compared with SGUS negative patients. Regarding patient-reported outcome measurements (PROMs), SGUS positive patients scored significantly lower on ESSPRI fatigue and pain, and more often found their disease state acceptable compared with SGUS negative patients. SGUS total score showed significant associations with various clinical and serological parameters, and with PROMs. Highest associations were found for UWS ($p=0.551$) and OSS ($p=0.532$).

Conclusion

SGUS positive patients show a distinct clinical phenotype compared with SGUS negative patients in all aspects of the disease: clinical, functional, serological and PROMs. SGUS could be a helpful tool in selecting patients for clinical trials and estimating treatment need.

Introduction

Primary Sjögren's syndrome (pSS) is a common systemic auto-immune disease.¹ Women are affected nine times more often than men.² pSS is a highly heterogeneous disease, which is reflected by the many different manifestations patients can have. Common symptoms, such as extreme fatigue and sicca symptoms have a major impact on the quality of life.³ This heterogeneity already emerges during the diagnostic work-up of pSS, i.e., not every pSS patient has auto-antibodies or a focus score positive salivary gland biopsy, which suggests that there are different subgroups of patients. It would be of great value to be able to identify individual patients at high risk for a severe disease outcome. Prospective cohort studies are gaining more and more importance in this quest.⁴ Since treatment options for pSS patients are eagerly awaited, but unfortunately still very limited, the search for patient stratification and proper selection methods for clinical trials is currently ongoing.

Regarding the care for (suspected) pSS patients, there is a unique collaboration between different departments at the University Medical Center Groningen (UMCG). The REgistry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) cohort has been set up to identify biomarkers and clinical parameters that determine and predict the longitudinal course of pSS. Observational studies, like the RESULT cohort, are important as they provide information on long-term outcome of pSS, which reflects daily clinical practice.

Salivary gland ultrasonography (SGUS) is increasingly gaining acceptance as an imaging tool of the salivary glands in pSS. Nowadays, ultrasound is widely accessible in outpatient rheumatology clinics. SGUS is non-invasive and non-irradiating, which makes it patient-friendly and an ideal imaging modality for repeated use.⁵⁻⁷

Previously, we have studied the validity of SGUS and found that a positive ultrasound, based on the total Hocevar score,⁶ predicts classification according to the American College of Rheumatology – European League Against Rheumatism (ACR-EULAR) criteria.⁸ Subsequently we provided evidence that measuring only hypoechogenic areas in one parotid and one submandibular gland is sufficient to predict ACR-EULAR classification, which increases the feasibility of SGUS.⁹ Although a simpler scoring system suffices for classification purposes, it is not yet known whether SGUS abnormalities can also be used for patient stratification, long-term follow-up or even as selection method for clinical trials. Therefore, a full SGUS evaluation according to the Hocevar score is performed in each patient included in the RESULT cohort.

The aim of this study was to investigate SGUS abnormalities in relation to clinical phenotype and patient characteristics, disease activity and disease damage in patients with pSS.

Patients and Methods

Registry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) cohort

The observational RESULT cohort combines up-to-date quality of care with gathering long-term prospective follow-up data in a large cohort of patients. For participation in the RESULT cohort, we consider all consecutive patients with probable or confirmed pSS who visit the outpatient clinic of the Department of Rheumatology and Clinical Immunology in the UMCG, a tertiary referral expertise center. Inclusion in the RESULT cohort is ongoing and duration of follow-up will be 10 years.

The present cross-sectional analysis included the baseline visit of all patients who were included in the RESULT cohort between January 2016 and December 2018. Patients with missing ultrasonographic examination as well as patients who did not fulfill the ACR-EULAR criteria for pSS (i.e. probable pSS patients)^{10,11} were excluded.

This study was conducted in accordance with the Declaration of Helsinki. The research protocol was approved by the Medical Ethics Committee of the UMCG; METC 2014/491. All subjects provided informed consent.

Assessments

Imaging, clinical, functional, histopathological, serological parameters and patients-reported outcome measurements (PROMs) were obtained according to a fixed protocol.

Salivary gland ultrasound

B-mode SGUS was performed using the MyLabSeven scanner (Esaote, Genova, Italy), equipped with a high-resolution linear probe (4-13 MHz). All ultrasonographic images were scored real-time by trained readers (AJS, KD, JFN, EM, RW). Test-retest reliability in our center was demonstrated previously.¹² The scoring system by Hocevar et al.⁶ was applied (range 0-48), including the components parenchymal echogenicity, homogeneity, presence of hypoechogenic areas, hyperechogenic reflections and clearness of the salivary gland border. Total SGUS score of ≥ 15 was considered positive.⁸

Other assessments

Demographic characteristics, EULAR Sjögren's Syndrome Disease Activity Index (ESSDAI),¹³ DAS28-ESR/CRP,^{14,15} number of tenderpoints, physician global disease activity (physician GDA), Sjögren's Syndrome Disease Damage Index (SSDDI),¹⁶ unstimulated whole saliva flow (UWS),¹⁷ Schirmer's test and ocular staining score (OSS)¹⁸

were determined. Two methods were applied for Schirmer's test and OSS, i.e. when dividing in normal/abnormal the worst eye was selected and when applied as a continuous variable the mean of both eyes was used. A salivary gland biopsy was not mandatory for participation in the RESULT cohort and therefore, parotid and labial salivary gland focus score were recorded if available.¹⁹⁻²¹

Serological parameters were determined, including presence of anti-SSA/Ro/SSB/La antibodies, immunoglobulin G (IgG) level, rheumatoid factor (RF) level, complement C3 and C4 levels and leukocyte count.

Patients completed a questionnaire, which included EULAR Sjögren's Syndrome Patient Reported Index (ESSPRI) dryness, fatigue and pain,²² patient acceptable symptom state (PASS), patient GDA and the 5-level EuroQol-5 dimensions health status questionnaire (EQ-5D-5L).²³

Statistics

Statistical analyses were performed using IBM SPSS Statistics 23 (SPSS, Chicago, IL, USA). Descriptive parameters were expressed as number (%) of patients for categorical data and mean (SD) or median (IQR) for continuous data.

Patient characteristics, disease activity and damage were compared between SGUS negative (score <15) and positive (score ≥15) patients. Subsequently, based on the median score of the SGUS positive group, SGUS positive patients were arbitrarily divided into two equal groups; patients with scores ≥15, but <27 were defined as medium-positive and patients with scores ≥27 were defined as high-positive.

Fisher's exact test or Chi square were used as appropriate to evaluate differences in categorical variables between the ultrasound groups. Independent Samples T-test or Mann-Whitney U test were used as appropriate to evaluate differences in continuous variables between the ultrasound groups. ESSDAI subdomains were summarized descriptively.

The association between SGUS total score and continuous variables was analyzed using Spearman correlation coefficient (ρ), and interpreted as poor association (0.0-0.2), fair (0.2-0.4), moderate (0.4-0.6), good (0.6-0.8) or excellent (0.8-1.0).²⁴ All parameters were also evaluated using univariate logistic regression analysis with SGUS outcome (positive vs. negative) as dependent variable. In case of residuals with non-Gaussian distribution, variables were transformed (log or square root), before being entered into the model. The explained variance was evaluated using Nagelkerke R^2 . P-values ≤0.05 were considered statistically significant.

All analyses were repeated when only taking the average score for 'hypoechoic areas' in the right parotid and submandibular gland into account,⁹ instead of the total SGUS score as described by Hocevar et al.⁶ For this score, a cut-off value of ≥1.5 was considered positive.²⁵

Results

Between January 2016 and December 2018, 186 patients were included in the RESULT cohort. Fourteen patients were excluded for the present analysis due to a missing (n=3) or incomplete (n=5) ultrasonographic examination, or because they did not fulfill the ACR-EULAR criteria (n=6). Of the eligible patients (n=172), mean age was 53 years (SD 13.9), 156 patients (91%) were female, 136 patients (79%) were SGUS positive (i.e. SGUS score ≥ 15) and median time since diagnosis was 8 years (table 1).

Comparison of SGUS negative and positive patients

Table 1 shows the characteristics of the total group of pSS patients, as well as of the patients with a positive or negative SGUS. There were no significant differences in general patient characteristics between the two groups, except for disease duration, which was longer in de SGUS positive patients.

SGUS positive patients had significantly higher ESSDAI scores, higher DAS28-ESR and higher physician GDA compared with SGUS negative patients, indicating higher disease activity (table 1; figure 1A,B; suppl. figure 1). Moreover, a parotid gland focus score ≥ 1 , UWS ≤ 0.1 ml/min and OSS ≥ 5 were more often seen in SGUS positive patients (table 1). SSDDI, UWS, Schirmer's test and OSS also differed significantly between both groups, with more damage and worse salivary and lacrimal gland function in SGUS positive patients (table 1; figure 1C-E).

Regarding the serological parameters, anti-SSA/Ro and anti-SSB/La antibodies were more often present in SGUS positive patients. Furthermore, SGUS positive patients showed higher levels of IgG and RF, lower complement C3 and C4 levels and lower leucocyte counts compared with SGUS negative patients (table 1; figure 1F).

Regarding PROMs, SGUS positive patients scored significantly lower on ESSPRI fatigue and pain, and more often found their disease state acceptable, which indicates that SGUS positive patients experienced less symptoms (table 1).

Results were confirmed with univariate logistic regression analyses (table 2). The explained variance of individual parameters varied from 0.1% for body mass index (BMI) to 22.4% for parotid gland biopsy (focus score ≥ 1).

As an overview of the available data, a heatmap of the characteristics of the individual pSS patients is shown in supplementary figure 2. The patients order has been determined based upon the total SGUS score. Overall, our data show that SGUS positive patients have a distinct clinical phenotype compared with SGUS negative patients. These findings illustrate the results described above in another way.

Comparison of patients with medium-positive or high-positive SGUS scores

When subdividing the group of SGUS positive patients into medium- and high-positive patients, we observed that compared with patients with a medium-positive SGUS score, patients with a high-positive SGUS score significantly more often had an UWS ≤ 0.1 ml/min, Schirmer's test ≤ 5 mm/5 min and OSS ≥ 5 (table 3).

Furthermore, SSDDI, UWS, Schirmer's test and OSS differed significantly between medium- and high-positive SGUS patients, showing more damage and a worse salivary and lacrimal gland function in the high-positive patients (table 3).

Patients with high-positive SGUS scores experienced significantly more dryness, but less fatigue and pain compared with patients with a medium-positive SGUS score (table 3).

Figure 1. Ultrasound total score (negative/positive) compared with

A. total ESSDAI;

B. Physician global disease activity;

C. total SSDDI;

D. Unstimulated whole saliva flow;

E. Ocular staining score and

F. total IgG level

ESSDAI = EULAR Sjögren's Syndrome Disease Activity Index;

SSDDI = Sjögren's Syndrome Disease Damage Index;

Correlations of SGUS total score

Significant associations were found between SGUS total score and disease duration ($p=0.279$), symptom duration ($p=0.234$), ESSDAI ($p=0.196$), DAS28-ESR ($p=0.159$), physician GDA ($p=0.217$), SSDDI ($p=0.398$), UWS ($p=-0.551$), Schirmer's test ($p=-0.349$) and OSS ($p=0.532$) (suppl. table 1; figure 2A-F and figure 3A).

Furthermore, significant associations were found between SGUS total score and IgG level ($p=0.264$), RF level ($p=0.343$), complement C4 level ($p=-0.200$) and leucocyte count ($p=-0.244$) (suppl. table 1; figure 3B,C).

Moreover, SGUS total scores showed significant association with PROMs; ESSPRI total score ($p=-0.157$), dryness ($p=0.223$), fatigue ($p=-0.209$) and pain ($p=-0.314$) (suppl. table 1; figure 3D-F).

To summarize, an increase in SGUS abnormalities is associated with longer disease duration, more damage and worse gland function, and with more dryness symptoms.

Figure 2. Scatterplots of ultrasound total score compared with

- A. Disease duration;
- B. total ESSDAI;
- C. SSDDI;
- D. Unstimulated whole saliva flow;
- E. Schirmer's test and
- F. Ocular staining score

For Schirmer's test and ocular staining score, the mean score of both eyes was calculated.

ESSDAI = EULAR Sjögren's Syndrome Disease Activity Index; SSDDI = Sjögren's Syndrome Disease Damage Index

SGUS – hypoechogenic areas only

When using only hypoechogenic areas to define SGUS positivity,⁹ multiple parameters showed similar results as when total Hoyer score was applied, except that no significant differences were found for: ESSDAI, DAS28-ESR, physician GDA, complement C3 and C4 levels, leucocyte counts and PASS (suppl. tables 2&3).

Figure 3. Scatterplots of ultrasound total score compared with

- A. DAS28-ESR;
- B. IgG level;
- C. Rheumatoid Factor level;
- D. ESSPRI dryness;
- E. ESSPRI fatigue and
- F. ESSPRI pain.

DAS28 = 28 joint Disease Activity Score; ESR = erythrocyte sedimentation rate; ESSPRI = EULAR Sjögren's Syndrome Patient Reported Index

Table 1.

Patients characteristics and comparison of SGUS negative and positive patients.

Characteristic	Total group (n=172)	SGUS ≤14 (n=36)	SGUS ≥15 (n=136)	P value
General characteristics				
Age, years	52.9 (13.9)	56.0 (14.0)	52.0 (13.8)	0.13
Females	156 (90.7%)	31 (86.1%)	125 (91.9%)	0.29
Disease duration, years	8.0 (4.0-13.0)	5.0 (3.0-8.8)	8.5 (5.0-13.8)	0.003
Symptom duration, years***	15.0 (9.0-21.0)	11.0 (6.0-19.0)	15.0 (10.0-22.0)	0.06
BMI (kg/m ²)*	24.9 (4.2)	24.6 (3.6)	24.8 (4.3)	0.79
Clinical parameters				
ESSDAI total score*	4.0 (2.0-8.0)	2.0 (0.0-6.5)	4.0 (2.0-8.0)	0.028
ESSDAI categories*				0.024
ESSDAI = 0	25 (14.6%)	10 (27.8%)	15 (11.1%)	
ESSDAI = 1-4	75 (43.9%)	16 (44.4%)	59 (43.7%)	
ESSDAI ≥5	71 (41.5%)	10 (27.8%)	61 (45.2%)	
DAS28-ESR**	3.2 (1.0)	2.9 (0.8)	3.3 (1.0)	0.027
DAS28-CRP**	2.3 (1.9-2.6)	2.3 (1.9-2.5)	2.3 (1.8-2.7)	0.74
Tenderpoints**	1.5 (0.0-8.0)	2.0 (0.0-12.0)	1.0 (0.0-8.0)	0.34
Physician GDA***	2.0 (1.0-3.0)	2.0 (1.0-3.0)	3.0 (1.0-4.0)	0.026
SSDDI total score***	2.0 (1.0-3.0)	1.5 (1.0-2.0)	2.0 (1.0-3.0)	0.018
UWS ≤0.1 mL/min**	111 (68.5%)	16 (45.7%)	95 (74.8%)	0.001
UWS flow, mL/min**	0.05 (0.01-0.13)	0.12 (0.03-0.27)	0.03 (0.00-0.11)	<0.001
Parotid gland biopsy, FS ≥1 ¹	85 (81.0%)	12 (50.0%)	73 (90.1%)	<0.001
Labial gland biopsy, FS ≥1 ²	47 (81.0%)	11 (68.8%)	36 (85.7%)	0.14
Schirmer's test ≤5mm/5min**	121 (74.7%)	25 (69.4%)	96 (76.2%)	0.41
Schirmer's test ODS, mm/5min**	4.0 (0.9-10.0)	5.5 (2.6-11.1)	3.5 (0.0-9.6)	0.020
OSS ≥5*	58 (34.1%)	3 (8.3%)	55 (41.0%)	<0.001
OSS ODS total score*	2.5 (0.9-5.0)	0.5 (0.0-2.0)	3.5 (1.0-5.0)	<0.001

Characteristic	Total group (n=172)	SGUS ≤14 (n=36)	SGUS ≥15 (n=136)	P value
Serological parameters				
Anti-SSA/Ro antibodies*	154 (90.1%)	27 (75.0%)	127 (94.1%)	0.001
Anti-SSB/La antibodies*	92 (53.8%)	9 (25.0%)	83 (61.5%)	<0.001
IgG level >16.0 g/mL*	81 (47.4%)	5 (13.9%)	76 (56.3%)	<0.001
IgG level, g/mL*	15.5 (11.2-20.3)	11.2 (9.3-13.0)	16.9 (12.1-21.8)	<0.001
RF level >5.0 IU/mL*	115 (67.3%)	12 (33.3%)	103 (76.3%)	<0.001
RF level, IU/mL*	15.0 (2.6-42.0)	2.1 (0.6-10.6)	21.0 (5.2-51.0)	<0.001
Complement C3 level (g/L)*	1.12 (0.23)	1.20 (0.24)	1.10 (0.22)	0.012
Complement C4 level (g/L)*	0.19 (0.15-0.24)	0.20 (0.18-0.24)	0.18 (0.14-0.24)	0.015
Leucocyte count 10 ⁹ /L*	5.4 (1.9)	6.3 (2.0)	5.2 (1.8)	0.002
Patient-reported outcome measurements				
ESSPRI total score*	6.0 (4.3-7.0)	6.7 (5.0-7.7)	5.7 (4.3-7.0)	0.016
ESSPRI dryness*	6.0 (5.0-8.0)	6.0 (4.0-8.0)	7.0 (5.0-8.0)	0.26
ESSPRI fatigue*	7.0 (5.0-8.0)	8.0 (5.0-8.0)	7.0 (4.3-8.0)	0.024
ESSPRI pain*	5.0 (2.0-7.0)	7.0 (5.0-8.0)	4.5 (2.0-7.0)	<0.001
Patient GDA**	6.0 (4.0-8.0)	7.0 (4.3-8.0)	6.0 (4.0-8.0)	0.15
EQ-5D-5L****	0.77 (0.14)	0.73 (0.17)	0.80 (0.12)	0.23
PASS, acceptable**	117 (71.8%)	21 (58.3%)	96 (75.6%)	0.042

Data are expressed as number of patients (%), mean (SD) or median (IQR). *<5% missing data; **5-10% missing data; ***10-15% missing data; ****22% missing data. Data available for 161% and 134% of patients.

Schirmer's test ≤5mm/min and OSS ≥5 were considered positive if criteria were met in at least one eye. For Schirmer's test ODS and OSS ODS, the mean score of both eyes was calculated.

SGUS = salivary gland ultrasonography; BMI = body mass index; DAS28 = 28-joint Disease Activity Score; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; ESSDAI = EULAR Sjögren's Syndrome Disease Activity Index; GDA = global disease activity; SDDI = Sjögren's Syndrome Disease Damage Index; UWS = unstimulated whole saliva; FS = focus score; OSS = ocular staining score; RF = rheumatoid factor; ESSPRI = EULAR Sjögren's Syndrome Patient Reported Index; EQ-5D-5L = 5-level EuroQol-5 dimensions health status questionnaire; PASS = patient acceptable symptom state.

Tabel 2.

Logistic regression analyses of demographic, clinical, serological and patient reported outcome parameters to predict ultrasound outcome.

Characteristic	Univariate analysis OR (95% CI)	P value	R ²
General characteristics			
Age, years	0.979 (0.952-1.007)	0.13	0.021
Females	1.833 (0.593-5.662)	0.29	0.009
Disease duration, years	1.108 (1.028-1.195)	0.007	0.082
Symptom duration, years***	1.036 (0.991-1.083)	0.12	0.028
BMI (kg/m ²)*	0.988 (0.906-1.078)	0.79	0.001
Clinical parameters			
ESSDAI total score†*	1.438 (1.040-1.988)	0.028	0.046
DAS28-ESR**	1.607 (1.047-2.466)	0.030	0.048
DAS28-CRP**	1.276 (0.738-2.205)	0.38	0.008
Tenderpoints**	0.976 (0.924-1.032)	0.40	0.007
Physician GDA***	1.473 (1.062-2.043)	0.020	0.064
SSDDI total score***	1.357 (1.053-1.748)	0.018	0.079
UWS ≤0.1 mL/min**	3.525 (1.622-7.663)	0.001	0.094
UWS flow, mL/min**	0.010 (0.001-0.138)	0.001	0.120
Parotid gland biopsy, FS ≥1 ¹	9.125 (3.089-26.953)	<0.001	0.224
Labial gland biopsy, FS ≥1 ²	2.727 (0.696-10.684)	0.15	0.049
Schirmer's test ≤5mm/5min**	1.408 (0.621-3.194)	0.41	0.006
Schirmer's test ODS, mm/5min‡**	0.658 (0.459-0.942)	0.022	0.051
OSS ≥5*	7.658 (2.236-26.227)	0.001	0.141
OSS ODS total score*	1.598 (1.274-2.005)	<0.001	0.212

Characteristic	Univariate analysis OR (95% CI)	P value	R ²
Serological parameters			
Anti-SSA/Ro antibodies*	5.292 (1.872-14.956)	0.002	0.084
Anti-SSB/La antibodies*	4.788 (2.088-10.984)	<0.001	0.136
IgG level >16.0 g/mL*	7.986 (2.927-21.795)	<0.001	0.192
IgG level g/mL*	1.129 (1.049-1.215)	0.001	0.121
RF level >5.0 IU/mL*	6.438 (2.897-14.305)	<0.001	0.192
RF level IU/mL*	1.020 (1.004-1.036)	0.012	0.094
Complement C3 level (g/L)*	0.132 (0.026-0.672)	0.015	0.055
Complement C4 level (g/L)*	0.026 (0.000-1.991)	0.10	0.024
Leucocyte count 10 ⁹ /L*	0.756 (0.622-0.919)	0.005	0.075
Patient-reported outcome measurements			
ESSPRI total score*	0.814 (0.662-1.001)	0.051	0.038
ESSPRI dryness‡*	1.680 (0.795-3.550)	0.17	0.016
ESSPRI fatigue*	0.837 (0.701-0.998)	0.047	0.040
ESSPRI pain‡*	0.380 (0.179-0.803)	0.011	0.075
Patient GDA†*	0.808 (0.427-1.529)	0.51	0.004
EQ-5D-5L***	10.489 (0.483-227.980)	0.14	0.026
PASS, acceptable**	2.212 (1.018-4.809)	0.045	0.036

† SQRT transformation; ‡LN transformation ; *<5% missing data; **5-10% missing data; ***10-15% missing data; ****22% missing data. Data available for 161% and 734% of patients.

BMI = body mass index; DAS28 = 28-joint Disease Activity Score; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; ESSDAI = EULAR Sjögren's Syndrome Disease Activity Index; GDA = global disease activity; SSDDI = Sjögren's Syndrome Disease Damage Index; UWS = unstimulated whole saliva; FS = focus score; OSS = ocular staining score; RF = rheumatoid factor; ESSPRI = EULAR Sjögren's Syndrome Patient Reported Index; EQ-5D-5L = 5-level EuroQoL-5 dimensions health status questionnaire; PASS = patient acceptable symptom state.

Tabel 3.

Comparison of SGUS positive patients with medium or high SGUS scores.

Characteristic	SGUS 15-26 (n=67)	SGUS 27-41 (n=69)	P value
General characteristics			
Age, years	53.1 (13.6)	51.0 (13.9)	0.39
Females	63 (94.0%)	62 (89.9%)	0.53
Disease duration, years	8.0 (4.0-14.0)	9.0 (6.0-13.5)	0.35
Symptom duration, years***	14.5 (8.0-21.8)	16.0 (11.0-22.0)	0.22
BMI (kg/m ²)*	24.8 (4.7)	24.8 (4.0)	0.99
Clinical parameters			
ESSDAI total score*	4.0 (2.0-8.0)	4.0 (2.0-8.0)	0.76
ESSDAI categories*			0.92
ESSDAI = 0	7 (10.6%)	8 (11.6%)	
ESSDAI = 1-4	30 (45.5%)	29 (42.0%)	
ESSDAI ≥5	29 (43.9%)	32 (46.4%)	
DAS28-ESR**	3.3 (1.0)	3.3 (1.0)	0.88
DAS28-CRP**	2.3 (1.7-2.7)	2.3 (2.0-2.7)	0.59
Tenderpoints**	2.0 (0.0-9.0)	0.0 (0.0-5.8)	0.19
Physician GDA***	2.0 (1.0-3.0)	3.0 (1.0-4.0)	0.28
SDDI total score***	2.0 (1.0-3.0)	2.0 (2.0-5.8)	0.001
UWS ≤0.1 mL/min**	40 (60.6%)	55 (90.1%)	<0.001
UWS flow, mL/min**	0.08 (0.01-0.15)	0.01 (0.00-0.04)	<0.001
Parotid gland biopsy, FS ≥1 ¹	36 (85.7%)	37 (94.9%)	0.27
Labial gland biopsy, FS ≥1 ²	18 (81.8%)	18 (90.0%)	0.67
Schirmer's test ≤5mm/5min**	41 (67.2%)	55 (84.6%)	0.022
Schirmer's test ODS, mm/5min**	5.0 (1.0-12.0)	2.0 (0.0-5.3)	0.017
OSS ≥5*	17 (25.8%)	38 (55.9%)	<0.001
OSS ODS total score*	2.0 (1.0-4.0)	4.0 (2.5-6.4)	<0.001

Characteristic	SGUS 15-26 (n=67)	SGUS 27-41 (n=69)	P value
Serological parameters			
Anti-SSA/Ro antibodies*	60 (90.9%)	67 (97.1%)	0.16
Anti-SSB/La antibodies*	38 (57.6%)	45 (65.2%)	0.36
IgG level >16.0 g/mL*	37 (56.1%)	39 (56.5%)	0.96
IgG level, g/mL*	16.8 (12.0-19.9)	17.4 (12.1-22.6)	0.57
RF level >5.0 IU/mL*	47 (71.2%)	56 (81.2%)	0.17
RF level, IU/mL*	15.5 (3.0-36.3)	32.0 (8.5-57.5)	0.037
Complement C3 level (g/L)*	1.10 (0.23)	1.10 (0.22)	0.88
Complement C4 level (g/L)*	0.19 (0.15-0.24)	0.18 (0.13-0.22)	0.16
Leucocyte count 10 ⁹ /L*	5.3 (1.6)	5.1 (2.0)	0.64
Patient-reported outcome measurements			
ESSPRI total score*	6.0 (4.3-7.2)	5.7 (4.0-6.7)	0.30
ESSPRI dryness*	6.0 (4.0-8.0)	7.0 (5.0-8.0)	0.050
ESSPRI fatigue*	7.0 (5.0-8.0)	6.0 (4.0-7.0)	0.042
ESSPRI pain*	6.0 (3.0-7.0)	4.0 (2.0-6.0)	0.019
Patient GDA**	6.0 (4.0-7.5)	6.0 (4.0-8.0)	0.80
EQ-5D-5L***	0.78 (0.14)	0.78 (0.11)	0.94
PASS, acceptable**	45 (73.8%)	51 (77.3%)	0.65

Data are expressed as number of patients (%), mean (SD) or median (IQR). *<5% missing data; **5-10% missing data; ***10-15% missing data; ****22% missing data. Data available for 161% and 234% of patients.

Schirmer's test ≤5mm/min and OSS ≥5 were considered positive if criteria were met in at least one eye. For Schirmer's test ODS and OSS ODS, the mean score of both eyes was calculated. SGUS = salivary gland ultrasonography; BMI = body mass index; DAS28 = 28-joint Disease Activity Score; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; ESSDAI = EULAR Sjögren's Syndrome Disease Activity Index; GDA = global disease activity; SSDDI = Sjögren's Syndrome Disease Damage Index; UWS = unstimulated whole saliva; FS = focus score; OSS = ocular staining score; RF = rheumatoid factor; ESSPRI = EULAR Sjögren's Syndrome Patient Reported Index; EQ-5D-5L = 5-level EuroQol-5 dimensions health status questionnaire; PASS = patient acceptable symptom state.

Discussion

In our prospective observational RESULT cohort, we showed that SGUS positive patients have a distinct clinical phenotype compared with SGUS negative patients. This difference was found in all aspects of the disease; clinical, functional, serological and PROMs. SGUS could give an overall indication about the observable and experienced severity of pSS.

SGUS positive patients have higher systemic disease activity, measured by ESDAI, DAS28-ESR and physician GDA, compared with SGUS negative patients. Of interest, SGUS positive patients score significantly worse on all individual items of the ACR-EULAR criteria, i.e. parotid gland biopsy, anti-SSA/Ro antibodies, Schirmer's test, OSS and UWS, compared with SGUS negative patients. Overall, total SGUS score showed the strongest association with OSS and UWS. In addition to these differences, SGUS positive patients score worse on SSDDI and serological parameters. These results show that SGUS enables us to identify patients with higher clinical and serological disease activity and more damage due to pSS.

Interestingly, SGUS positive patients experienced less fatigue and pain, both measured by ESSPRI, and more often found their disease state acceptable, which implies that these patients have a lower symptom burden. Perhaps patients who already have pSS (or symptoms) for several years are more used to it and developed their own coping strategy or they adjusted their expectations. Another possibility for the differences between SGUS negative and positive patients, is that there are indeed different phenotypic clusters of pSS patients. Very recently, Tarn et al.²⁶ defined four subgroups of pSS patients based upon the patient-reported outcomes dryness, fatigue, pain, anxiety and depression. Our data suggest that patients with high SGUS scores belong to a subgroup of patients with low symptom burden. Unfortunately, the Hospital Anxiety and Depression Scale (HADS) is not part of the questionnaires within our RESULT cohort. Therefore we were unable to verify whether SGUS scores also differ within these four subgroups of patients.

In the current study, we not only compared SGUS negative and positive patients based on a previously defined diagnostic cut-off point,⁸ but also zoomed in on the broad range of SGUS positive patients. As expected, patients with a high-positive SGUS score showed more pSS-related damage (SSDDI) and lower salivary and lacrimal gland function and more glandular damage, compared with patients with a medium-positive SGUS score. Interestingly, there were no differences in the percentage of patients with a positive biopsy or presence of anti-SSA/Ro antibodies between both groups. This could be because most patients within our cohort score positive on these items, which makes it more difficult to see differences within sub-

groups of patients. Moreover, both, focus score and anti-SSA/Ro antibodies, were collected as absent/present rather than on a continuous scale. Furthermore, the differences in ESSPRI fatigue and pain remain, with less patient symptoms in the high-positive group. In contrast however, high-positive SGUS patients do indeed experience more dryness compared with the medium-positive patients, which is logical considering the relationship between SGUS and glandular function.

The association between SGUS and disease duration suggests that there is an increase in ultrasonographic abnormalities over time. In contrast, when looking solely at the SGUS positive patients, there is no difference in disease duration between patients with medium-positive or high-positive scores. This raises the question how long it takes for these SGUS abnormalities to develop and how long these abnormalities continue to worsen. Gazeau et al.²⁷ showed that a nearly two-year interval between consecutive SGUS examinations was not enough to see significant progression over time in a group of 49 suspected pSS patients. A possible explanation for the lack of difference in disease duration in medium-positive and high-positive SGUS patients might be inter-observer differences, as it was previously shown that SGUS scores between different observers show more variability when total score exceeds 20.¹² Alternatively, it could be postulated that after a certain disease duration SGUS lesions stabilize, as is the case with the production of saliva.²⁸

In our previous studies, we have shown that for diagnostic purposes it suffices to only measure hypoechoic areas in one parotid and one submandibular gland and that optimal cut-off for a positive SGUS is ≥ 1.5 .²⁵ Since the use of SGUS to stratify pSS patients is essentially different from the use of SGUS for diagnostic purposes, we assessed whether results would be similar when using total SGUS score compared with only measuring hypoechoic areas. Regarding UWS, Schirmer's test, OSS and disease damage measured by SDDI, results were the same when only the component hypoechoic areas was taken into account. This suggests that evaluation of hypoechoic areas can be used to identify patients with glandular dysfunction and overall pSS-related damage. However, no differences in ESSDAI, physician GDA and DAS28-ESR were found when SGUS positivity was solely based on hypoechoic areas, although there were significant differences in serological activity. Therefore, the ultrasonographic component hypoechoic areas should not be used to identify patients with high disease activity. For this purpose, a more comprehensive scoring system, like the Hocevar scoring system,⁶ may be preferred above a scoring system including only 1 component.

Previously, several groups studied associations between SGUS and clinical, serological, and patient-reported parameters.²⁹⁻³⁷ However, there are considerable differences between some of these studies and our current study. The most important

difference is that most studies focus on the possible diagnostic purposes of SGUS rather than its possible use for stratification of already classified pSS patients.^{30,33-35} In our study, differences between the SGUS negative and positive patients cannot be attributed to the fact that there are non-SS sicca controls included, as we only included pSS patients in this study. In comparison with the previous studies, we included a considerable higher number of pSS patients. Nevertheless, previous studies found significant differences between SGUS negative and positive patients regarding ESSDAI,³¹ tear- and saliva production,²⁹⁻³² presence of anti-SSA/Ro antibodies and/or anti-SSB/La antibodies,²⁹⁻³² RF positivity,^{30,31} VAS dry mouth³² and ESSPRI dryness,²⁹ and, with the exception of the patient-reported dryness symptoms, we were able to confirm these results. In contrast, other studies did not find differences in ESSDAI^{29,30} and SSDDI³⁰ between SGUS negative and positive patients. In a study including pSS as well as non-SS sicca controls, SGUS positive patients had higher labial gland focus score and more often had an OSS ≥ 3 , UWS ≤ 0.1 mL/min, were anti-SSA/Ro/SSB/La and RF positive and had hypergammaglobulinemia, compared with SGUS negative patients.³³ In a large mixed-population of pSS patients and healthy controls, Milic et al.³⁶ found significant correlations between SGUS score and age, minor salivary gland biopsy, SSDDI, SSDAI (activity index) and ESSDAI. However, in contrast to our findings, the authors did not find a significant correlation between SGUS and disease duration and ESSPRI. Other studies also found associations between SGUS and ESSDAI³⁴ and several serological parameters,^{34,35,37} but again in a mixed population of pSS and non-SS sicca controls.

Other differences between previously performed studies and our current study relate to the applied SGUS scoring system and criteria set used for classification. Some studies, including this current study, applied the Hocevar scoring system,⁶ but different cut-off points were applied.^{29,30} Furthermore, we applied the ACR-EULAR classification criteria, as did Kim et al.³³ and La Paglia et al.,³⁷ whereas in all other studies, including the more recent ones, the AECG criteria were applied.^{29-32,34-36}

To confirm our results in different populations, a consensus scoring system with a validated cut-off is needed. Very recently, the first steps in reaching international expert consensus have indeed been taken by the OMERACT task force on Sjögren's syndrome.³⁸ Furthermore, the development of an SGUS endpoint for use in future clinical trials is part of the innovative medicines initiative (IMI) project 'NECESSITY'.³⁹ Two recent studies showed that the addition of SGUS improves the performance of the ACR-EULAR classification criteria.^{25,40} In addition to the potential value of SGUS for diagnostic purposes, our results indicate that SGUS could also be used for patient stratification, e.g., for the selection of subgroups of patients for clinical trials. Although our results seem promising, the value of SGUS for patient

stratification needs to be confirmed by other research groups. Currently, within the European Union, initiatives, i.e., the HarmonicSS research project, are already taken to improve stratification of pSS patients, also including the use of SGUS.⁴¹

Our prospective observational cohort revealed that the majority of patients is SGUS positive. These patients have a longer disease duration, a higher disease activity and more pSS-related damage compared with SGUS negative patients, whereas SGUS negative patients experience more fatigue and pain. In the future, SGUS hopefully can be used as a valid selection method for clinical trials, as it gives an overall indication of the disease.

Acknowledgement

The authors would like to thank all participating patients.

Part of this study was presented as an oral presentation at the Dutch Rheumatology congress on September 26, 2019.

References

1. Mariëtte X, Criswell LA. Primary Sjögren's Syndrome. *N Engl J Med* 2018;378:931-9.
2. Qin B, Wang J, Yang Z, *et al.* Epidemiology of primary Sjögren's syndrome: a systematic review and meta-analysis. *Ann Rheum Dis* 2015;74:1983-9.
3. Pijpe J, Meijer JM, Bootsma, *et al.* Clinical and histologic evidence of salivary gland restoration supports the efficacy of rituximab treatment in Sjögren's syndrome. *Arthritis Rheumatism* 2009;60:3251-6.
4. Rischmueller M, Tieu J, Lester S. Primary Sjögren's syndrome. *Best Pract Res Clin Rheumatol* 2016;30:189-220.
5. 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.
6. Hocevar A, Ambrozic A, Rozman B, *et al.* Ultrasonographic changes of major salivary glands in primary Sjögren's syndrome. Diagnostic value of a novel scoring system. *Rheumatology (Oxford)* 2005;44:768-72.
7. Delli K, Dijkstra PU, Stel AJ, *et al.* Diagnostic properties of ultrasound of major salivary glands in Sjögren's syndrome: a meta-analysis. *Oral Dis* 2015;21:792-800.
8. Mossel E, Delli K, van Nimwegen JF, *et al.* Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1883-9.
9. 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.
10. Shiboski CH, Shiboski SC, Seror R, *et al.* 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:9-16.
11. 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.
12. Delli K, Arends S, van Nimwegen JF, *et al.* Ultrasound of the Major Salivary Glands is a Reliable Imaging Technique in Patients with Clinically Suspected Primary Sjögren's syndrome. *Ultraschall Med* 2018;39:328-33.

13. Seror R, Ravaud P, Bowman SJ, *et al.* EULAR Sjögren's syndrome disease activity index: development of a consensus systemic disease activity index for primary Sjögren's syndrome. *Ann Rheum Dis* 2010;69:1103-9.
14. Prevoo M, van 't Hof M, Kuper H, *et al.* Modified disease activity score that include twenty-eight-joint-counts development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44-8.
15. Moerman RV, Arends S, Meiners PM, *et al.* Detailed analysis of the articular domain in patients with primary Sjögren's syndrome. *J Rheumatol* 2017;44:292-6.
16. Vitali C, Palombi G, Baldini C, *et al.* Sjögren's syndrome disease damage index and disease activity index scoring system for the assessment of disease damage in disease activity in Sjögren's syndrome, derived from an analysis of a cohort of Italian patients. *Arthritis Rheum* 2007;56:2223-31.
17. Jensen SB, Vissink A. Salivary Gland Dysfunction and Xerostomia in Sjögren's Syndrome. *Oral Maxillofac Surg Clin NA* 2014;26:35-53.
18. Whitcher JP, Shiboski CH, Shiboski SC, *et al.* A simplified quantitative method for assessing keratoconjunctivitis sicca from the Sjögren's syndrome international registry. *Am J Ophthalmol* 2010;149:405-15.
19. Chisholm DM, Mason DK. Labial salivary gland biopsy in Sjögren's disease. *J Clin Pathol* 1968;21:656-60.
20. Pijpe J, Kalk WWI, van der Wal JE, *et al.* Parotid gland biopsy compared with labial biopsy in the diagnosis of patients with primary Sjögren's syndrome. *Rheumatology (Oxford)* 2007;46:335-41.
21. Greenspan JS, Daniels TE, Talal N, *et al.* The histopathology of Sjögren's syndrome in labial salivary gland biopsies. *Oral Surgery, Oral Med Oral Pathol* 1974;37:217-29.
22. Seror R, Ravaud P, Mariette X, *et al.* EULAR Sjögren's syndrome patient reported index (ESSPRI) : development of a consensus patient index for primary Sjögren's syndrome. *Ann Rheum Dis* 2011;70:968-72.
23. Richardson J, McKie J, Bariola E. Multiattribute utility instruments and their use. In Cuylen AJ, editors. *Encyclopedia of Health Economics*. San Diego: Elsevier 2014;341-57.
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
25. Van Nimwegen J, Mossel E, Delli K, *et al.* Incorporation of salivary gland ultrasonography into the ACR-EULAR criteria for primary Sjögren's syndrome. *Arthritis Care Res* 2020;72:583-90.

26. Tarn JR, Howard-Tripp N, Lendrem DW, *et al.* Symptom-based stratification of patients with primary Sjögren's syndrome: multi-dimensional characterization of international observational cohorts and reanalyses of randomised clinical trials. *Lancet Rheumatol* 2019;1:85-94.
27. Gazeau P, Cornec D, Jousse-Joulin S, *et al.* Time-course of ultrasound abnormalities of major salivary glands in suspected Sjögren's syndrome. *Jt Bone Spine* 2018;85:227-32.
28. Pijpe J, Kalk W, Bootsma H, *et al.* Progression of salivary gland dysfunction in patients with Sjögren's syndrome. *Ann Rheum Dis* 2007;66:107-13.
29. Inanc N, Şahinkaya Y, Mumcu G, *et al.* Evaluation of salivary gland ultrasonography in primary Sjögren's syndrome: does it reflect clinical activity and outcome of the disease? *Clin Exp Rheumatol* 2019;37Suppl118:140-5.
30. Lee K, Lee S, Kim H. Diagnosis and predictive evaluation using salivary gland ultrasonography in primary Sjögren's syndrome. *Clin Exp Rheumatol* 2018;36Suppl112:165-72.
31. Fidelix T, Czapkowski A, Azjen S, *et al.* Salivary gland ultrasonography as a predictor of clinical activity in Sjögren's syndrome. *PLoS One* 2017;12:e0182287.
32. Hammenfors DS, Brun, JG, Jonsson R, *et al.* Diagnostic utility of major salivary gland ultrasonography in primary Sjögren's syndrome. *Clin Exp Rheumatol* 2015;33:56-62.
33. Kim J, Lee H, Park S, *et al.* Salivary gland ultrasonography findings are associated with clinical, histological, and serologic features of Sjögren's syndrome. *Scand J Rheumatol* 2018;47:303-10.
34. Theander E, Mandl T. Primary Sjögren's syndrome: diagnostic and prognostic value of salivary gland ultrasonography using a simplified scoring system. *Arthritis Care Res* 2014;66:1102-7.
35. Wernicke D, Hess H, Gromnica-Ihle E, *et al.* Ultrasonography of salivary glands – a highly specific imaging procedure for diagnosis of Sjögren's syndrome. *J Rheumatol* 2008;35:285-93.
36. Milic V, Colic J, Cirkovic A, *et al.* Disease activity and damage in patients with primary Sjögren's syndrome: Prognostic value of salivary gland ultrasonography. *PLoS One* 2019;14:1-12.

37. La Paglia GMC, Sanchez-Pernaute O, Alunno A, *et al.* Ultrasound salivary gland involvement in Sjögren's syndrome vs. other connective tissue diseases: is it autoantibody and gland dependent? *Clin Rheumatol* 2020;39:1207-15.
38. 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.
39. New clinical endpoints in primary Sjögren's syndrome: an interventional trial based on stratifying patients (NECESSITY). *Innovative Medicines Initiative*. Grant agreement number 806975.
40. Jousse-Joulin S, Gatlneau F, Baldini C, *et al.* Weight of salivary gland ultrasonography compared to other items of the 2016 ACR/EULAR classification criteria for primary Sjögren's syndrome. *J Intern Med* 2020;287:180-8.
41. Harmonization and integrative analysis of regional, national and international cohorts on primary Sjögren's syndrome (pSS) towards improved stratification, treatment and health policy making (HarmonicSS). *Horizon2020 project*. Grant agreement number 731944.

Supplementary Table 1.

Associations of SGUS with demographic, clinical, serological and patient reported outcome measurements.

*<5% missing data; **5-10% missing data; ***10-15% missing data; ****22% missing data.

For Schirmer's test ODS and OSS ODS, the mean score of both eyes was calculated.

SGUS = salivary gland ultrasonography; BMI = body mass index; DAS28 = 28-joint Disease Activity Score; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; ESSDAI = EULAR Sjögren's Syndrome Disease Activity Index; GDA = global disease activity; SSDDI = Sjögren's Syndrome Disease Damage Index; UWS = unstimulated whole saliva; OSS = ocular staining score; RF = rheumatoid factor; ESSPRI = EULAR Sjögren's Syndrome Patient Reported Index; EQ-5D-5L = 5-level EuroQol-5 dimensions health status questionnaire.

Characteristic	Spearman's ρ	P value
General characteristics		
Age, years	-0.085	0.27
Disease duration, years	0.279	<0.001
Symptom duration, years***	0.234	0.003
BMI (kg/m ²)*	0.005	0.95
Clinical parameters		
ESSDAI total score*	0.196	0.010
DAS28-ESR**	0.159	0.045
DAS28-CRP**	0.091	0.25
Tenderpoints**	-0.080	0.31
Physician GDA***	0.217	0.007
SSDDI total score***	0.398	<0.001
UWS flow, mL/min**	-0.551	<0.001
Schirmer's test ODS, mm/5min**	-0.349	<0.001
OSS ODS total score*	0.532	<0.001
Serological parameters		
IgG level, g/mL*	0.264	<0.001
RF level, IU/mL*	0.343	<0.001
Complement C3 level (g/L)*	-0.089	0.25
Complement C4 level (g/L)*	-0.200	0.009
Leucocyte count 10 ⁹ /L*	-0.244	0.001
Patient-reported outcome measurements		
ESSPRI total score*	-0.157	0.043
ESSPRI dryness*	0.223	0.004
ESSPRI fatigue*	-0.209	0.007
ESSPRI pain*	-0.314	<0.001
Patient GDA**	-0.033	0.68
EQ-5D-5L****	0.037	0.67

Supplementary Table 2.

Comparison of SGUS negative and positive patients based on presence of hypoechogenic areas.

Characteristic	SGUS HYPO <1.5 (n=38)	SGUS HYPO ≥1.5 (n=134)	P value
General characteristics			
Age, years	55.2 (14.7)	52.2 (13.6)	0.25
Females	32 (84.2%)	124 (92.5%)	0.12
Disease duration, years	5.5 (3.0-9.5)	8.5 (5.0-13.3)	0.011
Symptom duration, years***	11.0 (6.3-19.0)	15.5 (10.0-22.0)	0.050
BMI (kg/m ²)*	25.4 (4.3)	24.7 (4.1)	0.38
Clinical parameters			
ESSDAI total score*	2.0 (0.8-7.3)	4.0 (2.0-8.0)	0.12
ESSDAI categories*			0.18
ESSDAI = 0	9 (23.7%)	16 (12.0%)	
ESSDAI = 1-4	16 (42.1%)	59 (44.4%)	
ESSDAI ≥5	13 (34.2%)	58 (43.6%)	
DAS28-ESR**	2.9 (1.0)	3.2 (1.0)	0.10
DAS28-CRP**	2.3 (2.0-2.5)	2.3 (1.8-2.7)	0.97
Tenderpoints**	2.0 (0.0-10.0)	1.0 (0.0-8.0)	0.55
Physician GDA***	2.0 (1.0-3.0)	3.0 (1.0-3.3)	0.08
SSDDI total score***	1.0 (1.0-2.0)	2.0 (1.0-3.0)	0.015
UWS ≤0.1 mL/min**	16 (43.2%)	95 (76.0%)	<0.001
UWS flow, mL/min**	0.13 (0.04-0.26)	0.03 (0.00-0.11)	<0.001
Parotid gland biopsy, FS ≥1 ¹	15 (60.0%)	70 (87.5%)	0.002
Labial gland biopsy, FS ≥1 ²	10 (66.7%)	37 (86.0%)	0.10
Schirmer's test ≤5mm/5min**	24 (64.9%)	97 (77.6%)	0.12
Schirmer's test ODS, mm/5min**	5.5 (2.8-14.0)	3.0 (0.0-9.0)	0.006
OSS ≥5*	3 (7.9%)	55 (41.7%)	<0.001
OSS ODS total score*	0.5 (0.0-2.0)	3.5 (1.1-5.4)	<0.001

Characteristic	SGUS HYPO <1.5 (n=38)	SGUS HYPO ≥1.5 (n=134)	P value
Serological parameters			
Anti-SSA/Ro antibodies*	28 (73.7%)	126 (94.7%)	<0.001
Anti-SSB/La antibodies*	10 (26.3%)	82 (61.7%)	<0.001
IgG level >16.0 g/mL*	8 (21.1%)	73 (54.9%)	<0.001
IgG level, g/mL*	11.5 (9.5-14.2)	16.7 (12.1-21.7)	<0.001
RF level >5.0 IU/mL*	13 (34.2%)	102 (76.7%)	<0.001
RF level, IU/mL*	2.2 (0.8-7.5)	22.0 (5.4-48.5)	<0.001
Complement C3 level (g/L)*	1.19 (0.24)	1.10 (0.23)	0.09
Complement C4 level (g/L)*	0.21 (0.18-0.24)	0.18 (0.14-0.23)	0.30
Leucocyte count 10 ⁹ /L*	6.1 (1.8)	5.2 (1.9)	0.06
Patient-reported outcome measurements			
ESSPRI total score*	6.7 (5.0-7.9)	5.7 (4.3-7.0)	0.007
ESSPRI dryness*	7.0 (5.0-8.0)	6.0 (5.0-8.0)	0.91
ESSPRI fatigue*	8.0 (5.0-8.0)	7.0 (4.0-8.0)	0.025
ESSPRI pain*	7.0 (4.5-8.0)	5.0 (2.0-7.0)	<0.001
Patient GDA**	7.0 (4.5-8.0)	6.0 (4.0-8.0)	0.21
EQ-5D-5L***	0.75 (0.13)	0.78 (0.14)	0.41
PASS, acceptable**	24 (64.9%)	93 (73.8%)	0.29

Data are expressed as number of patients (%), mean (SD) or median (IQR). *<5% missing data; **5-10% missing data; ***10-15% missing data; ****22% missing data. Data available for ¹61% and ²34% of patients.

Schirmer's test ≤5mm/min and OSS ≥5 were considered positive if criteria were met in at least one eye. For Schirmer's test ODS and OSS ODS, the mean score of both eyes was calculated.

SGUS = salivary gland ultrasonography; BMI = body mass index; DAS28 = 28-joint Disease Activity Score; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; ESSDAI = EULAR Sjögren's Syndrome Disease Activity Index; GDA = global disease activity; SSDDI = Sjögren's Syndrome Disease Damage Index; UWS = unstimulated whole saliva; FS = focus score; OSS = ocular staining score; RF = rheumatoid factor; ESSPRI = EULAR Sjögren's Syndrome Patient Reported Index; EQ-5D-5L = 5-level EuroQol-5 dimensions health status questionnaire; PASS = patient acceptable symptom state.

Supplementary Table 3.

Comparison of SGUS patients with medium or high SGUS scores based on presence of hypoechogenic areas.

Characteristic	SGUS HYPO 1.5-2.0 (n=72)	SGUS HYPO 2.5-3.0 (n=62)	P value
General characteristics			
Age, years	52.7 (13.7)	51.7 (13.7)	0.69
Females	68 (94.4%)	56 (90.3%)	0.51
Disease duration, years	8.5 (4.0-13.8)	8.5 (5.0-13.3)	0.55
Symptom duration, years***	16.0 (8.0-22.0)	15.0 (11.0-22.0)	0.45
BMI (kg/m ²)*	24.6 (4.3)	24.8 (3.9)	0.80
Clinical parameters			
ESSDAI total score*	4.0 (2.0-6.0)	5.0 (3.0-10.0)	0.006
ESSDAI categories*			0.044
ESSDAI = 0	11 (15.5%)	5 (8.1%)	
ESSDAI = 1-4	36 (50.7%)	23 (37.1%)	
ESSDAI ≥5	24 (33.8%)	34 (46.6%)	
DAS28-ESR**	3.2 (1.0)	3.3 (0.9)	0.74
DAS28-CRP**	2.2 (1.7-2.7)	2.4 (2.0-2.6)	0.44
Tenderpoints**	2.0 (0.0-10.0)	0.0 (0.0-5.3)	0.29
Physician GDA***	2.0 (1.0-3.0)	3.0 (1.0-4.0)	0.07
SSDDI total score***	2.0 (1.0-2.3)	2.0 (2.0-6.0)	0.001
UWS ≤0.1 mL/min**	47 (67.1%)	48 (87.3%)	0.009
UWS flow, mL/min**	0.07 (0.01-0.14)	0.01 (0.00-0.05)	<0.001
Parotid gland biopsy, FS ≥1 ¹	39 (83.0%)	31 (93.4%)	0.18
Labial gland biopsy, FS ≥1 ²	21 (84.0%)	16 (88.9%)	1.00
Schirmer's test ≤5mm/5min**	47 (70.1%)	50 (86.2%)	0.032
Schirmer's test ODS, mm/5min**	3.5 (1.0-11.5)	1.8 (0.0-5.1)	0.019
OSS ≥5*	21 (29.6%)	34 (55.7%)	0.002
OSS ODS total score*	2.5 (1.0-4.0)	4.0 (2.5-6.0)	<0.001

Characteristic	SGUS HYPO 1.5-2.0 (n=72)	SGUS HYPO 2.5-3.0 (n=62)	P value
Serological parameters			
Anti-SSA/Ro antibodies*	66 (93.0%)	60 (96.8%)	0.45
Anti-SSB/La antibodies*	42 (59.2%)	40 (64.5%)	0.53
IgG level >16.0 g/mL*	34 (47.2%)	39 (62.9%)	0.08
IgG level, g/mL*	15.9 (11.7-18.9)	17.7 (13.1-22.9)	0.06
RF level >5.0 IU/mL*	53 (74.6%)	49 (79.0%)	0.55
RF level, IU/mL*	20.0 (4.9-37.0)	24.5 (6.3-65.0)	0.12
Complement C3 level (g/L)*	1.10 (0.23)	1.10 (0.22)	0.93
Complement C4 level (g/L)*	0.19 (0.15-0.22)	0.18 (0.14-0.23)	0.31
Leucocyte count 10 ⁹ /L*	5.3 (1.8)	5.1 (2.0)	0.49
Patient-reported outcome measurements			
ESSPRI total score*	6.0 (4.3-7.0)	5.3 (4.0-6.9)	0.58
ESSPRI dryness*	6.0 (4.0-7.0)	7.0 (5.0-8.0)	0.002
ESSPRI fatigue*	7.0 (5.0-8.0)	6.0 (3.5-7.0)	0.06
ESSPRI pain*	5.0 (3.0-7.0)	4.0 (2.0-6.0)	0.038
Patient GDA**	6.0 (4.0-7.3)	6.0 (4.0-8.0)	0.69
EQ-5D-5L***	0.77 (0.16)	0.79 (0.11)	0.44
PASS, acceptable**	47 (71.2%)	46 (76.7%)	0.49

Data are expressed as number of patients (%), mean (SD) or median (IQR). *<5% missing data; **5-10% missing data;

10-15% missing data; *22% missing data. Data available for 161% and 134% of patients.

Schirmer's test ≤5mm/min and OSS ≥5 were considered positive if criteria were met in at least one eye. For Schirmer's test ODS and OSS ODS, the mean score of both eyes was calculated.

SGUS = salivary gland ultrasonography; BMI = body mass index; DAS28 = 28-joint Disease Activity Score; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; ESSDAI = EULAR Sjögren's Syndrome Disease Activity Index; GDA = global disease activity; SSSDI = Sjögren's Syndrome Disease Damage Index; UWS = unstimulated whole saliva; FS = focus score; OSS = ocular staining score; RF = rheumatoid factor; ESSPRI = EULAR Sjögren's Syndrome Patient Reported Index; EQ-5D-5L = 5-level EuroQol-5 dimensions health status questionnaire; PASS = patient acceptable symptom state.

Supplementary Figure 1.

Comparison of total and subdomain ESSDAI activity between SGUS negative (total score 0-14), medium-positive (total score 15-26) and high-positive (total score 27-48) patients.

Supplementary Figure 2.

Patients characteristics. A heatmap of the parameters for the 172 individual patients. The values have been scaled between zero (turquoise) and one (red), with 1 being the worst score. Missing values are shown as white bars.

SGUS = salivary gland ultrasonography; BMI = body mass index; DAS28 = 28-joint Disease Activity Score; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; ESSDAI = EULAR Sjögren's Syndrome Disease Activity Index; GDA = global disease activity; SSDDI = Sjögren's Syndrome Disease Damage Index; UWS = unstimulated whole saliva; FS = focus score; OSS = ocular staining score; RF = rheumatoid factor; ESSPRI = EULAR Sjögren's Syndrome Patient Reported Index; EQ-5D-5L = 5-level EuroQoL-5 dimensions health status questionnaire; PASS = patient acceptable symptom state.

CHAPTER 7

Can ultrasound of the major salivary glands assess histopathological changes induced by treatment with rituximab in primary Sjögren's syndrome?

Esther Mossel¹, Konstantina Delli², Suzanne Arends¹, Erlin A Haacke^{1,3}, Bert van der Vegt³, Jolien F van Nimwegen¹, Alja J Stel¹, Fred KL Spijkervet², Arjan Vissink², Frans GM Kroese¹ and Hendrika Bootsma¹.

¹ Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

² Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

³ Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Edited version of Ann Rheum Dis 2019;78:e27.

With great interest we have read the recent publication by Fisher et al.,¹ entitled 'Effect of rituximab on a salivary gland ultrasound score in primary Sjögren's syndrome: results of the TRACTISS randomized double-blind multicenter substudy' in which the authors demonstrated a significant improvement in total ultrasound score (TUS) after rituximab treatment compared with placebo at weeks 16 and 48 in 52 patients with primary Sjögren's syndrome (pSS).

We and others have shown that treatment with rituximab, a chimeric anti-CD20 monoclonal antibody, affects the histopathology of the salivary glands and results in a decrease in area of lymphocytic infiltrate in the labial and parotid glands.²⁻⁵ This reduction in infiltrated area, that is, up to 50%, is mainly caused by B cell depletion.² Rituximab treatment also causes a significant loss of germinal centres.² Interestingly, rituximab treatment also leads to a significant restoration of the ductal epithelium glandular tissue itself. This is illustrated by a significant decrease in number and severity of lymphoepithelial lesions (LELs) in parotid gland tissue 12 weeks after rituximab treatment.^{2,6} The normalisation of the epithelium appears to be a direct consequence of depletion of intraepithelial B cells.⁷

Significant improvement of the parotid gland ultrasound score of patients with pSS has been observed before in another randomized controlled trial (RCT) with rituximab, the Tolerance and Efficacy of Rituximab in Primary Sjögren's Syndrome (TEARS) study.⁸ This is an RCT with rituximab versus placebo, performed by a French group. Together with the improvement of the histopathology of the gland, it might not be a surprise that Fisher et al.¹ also observed a significant improvement in TUS. This newly developed scoring system is composed of five domains: echogenicity, consistency, definition, glands involved and hypoechoic foci size. In the study by Fisher et al.,¹ the difference in TUS between rituximab and placebo-treated patients was only attributed to the 'definition' domain. This domain addresses the visibility of the salivary gland posterior border. Other studies have shown that this domain is hard to assess,^{1,9} which is also supported by the accompanying low interobserver reliability.^{10,11} None of the other components showed a significant improvement at weeks 16 and 48 after rituximab treatment in patients with pSS.¹ In a previous study, we have shown that the presence of hypoechoic areas is the most important component in the evaluation of pSS.¹² Hypoechoic areas contribute to several domains scored when using the ultrasound scoring system applied by Fisher et al.,¹ which include echogenicity, consistency and hypoechoic foci size. Although no significant improvement was found in these ultrasonographic components, there seems to be a tendency that the consistency domain does somewhat improve after rituximab treatment.¹ On the basis of the observed histopathological changes seen in parotid glands after rituximab treatment, however, it might have been expected that more domains of the TUS would have been improved.

The question becomes now why Fisher et al.,¹ in our opinion, observe only minimal changes in ultrasound score after rituximab treatment. First, as long as we do not exactly know which histopathological parameters in salivary glands are being measured by ultrasound, it is hard to understand the effects of treatment on the ultrasound scores. While in the TEARS study, a good association between focus score in the labial gland biopsy and their ultrasound score was found, our group found only fair to moderate associations between the Hocevar ultrasound score¹³ and focus score in parotid and labial salivary gland biopsies, respectively.^{14,15} The focus score is not appropriate to measure the local effect of rituximab on the periductal lymphocytic infiltration.² Moreover, a focus score does not reflect the size of the infiltrate and the inflamed area can be considerably reduced as a consequence of treatment, while the number of foci does not or only slightly decrease. Indeed, after rituximab treatment, the focus score was not altered, although there was a reduction in the size of the periductal infiltrate.² Taking this limitation of the focus score into account, other measurements, such as the relative area of infiltrating CD45⁺ cells, the number of B cells/mm² of glandular parenchyma, the presence and stage of LELs as well as the presence of germinal centres, more accurately define the magnitude of the inflammatory status of the gland.² Since these histopathological features are too small to be represented by the hypoechogenic areas, it is not surprising that only minor changes in TUS were observed by Fisher et al.¹ Apparently, the change in the area of the inflammatory infiltrates might not be measured directly by the change in percentage of hypoechogenic areas seen on ultrasound.

Second, another possible explanation why no more domains of the TUS change after rituximab treatment is that TUS is not sensitive enough to detect the histopathological changes. The consistency component is scored on a categorical scale, based on the ultrasonographers' estimate of the amount of hypoechogenic areas present. It is questionable whether the decrease in the glandular area occupied by infiltrating cells is reflected by a similar change in the ultrasound image. In other words, it is uncertain whether the applied scoring system (TUS) is sensitive enough to detect this change. With only a few categories, there has to be a relative large change in the amount of hypoechogenic areas present in the ultrasonographic image of a particular salivary gland to result in a different score. Furthermore, with regard to the ultrasonographic scoring, there are some other points to keep in mind. For example, ultrasound scoring was done live, but for an unbiased evaluation, blinded scoring is warranted. In addition, as already pointed out by Fisher et al.,¹ the ultrasound devices between the different centres were not standardised and intraobserver and interobserver reliability were not assessed. This last point is particularly important since previously there was more inconsistency in scoring

the severity of the ultrasonographic characteristics¹⁰ as used in the study of Fisher et al.¹

On the other hand, the ultrasonographers were able to reliably identify in which patients the ultrasonographic characteristics support the diagnosis of pSS.¹⁰

Finally, the most simple explanation for the minor TUS changes is that the TRACTISS study, a randomized double blind placebo controlled clinical TRial of Anti-B-Cell Therapy In patients with primary Sjögren's Syndrome, failed to meet its primary endpoint, that is, a change of 30% in oral dryness or fatigue on the visual analogue scale.⁸ Taking this into account, we should have been surprised if major differences in TUS were found. Another important point to mention is that in the TRACTISS study,⁸ the effect of rituximab treatment on the histopathology of the glandular tissue itself has not been studied, nor is any information given about how many patients had a positive salivary gland biopsy at the start of the study. Because of this, it is impossible to make a statement about whether or not the right subgroup of patients was included in the first place, to detect changes in TUS, as it appears that rituximab shows efficacy in a subgroup of pSS patients.¹⁶

In conclusion, rituximab treatment leads to significant improvements in histopathological characteristics of salivary gland biopsies. This improvement was not reflected by a similar improvement in the amount of hypoechogenic areas identified with ultrasound by Fisher et al.,¹ but this could be due to several reasons, that is, ultrasound is not sensitive enough to detect the histopathological differences or does not reflect the changes seen in the biopsies and the TRACTISS study was a negative study, perhaps due to the chosen inclusion criteria. In order to make a statement about whether or not ultrasound is the right tool to measure treatment effect, the sensitivity to change should be determined in positive clinical trials. At this point, the baseline histopathological characteristics of the parotid gland should be used to predict the effect of rituximab therapy and to help us to determine a personalised treatment approach.

References

1. Fisher BA, Everett CC, Rout J, *et al.* Effect of rituximab on a salivary gland ultrasound score in primary Sjögren's syndrome: results of the TRACTISS randomized double-blind multicentre substudy. *Ann Rheum Dis* 2018;77:412-6.
2. Delli K, Haacke EA, Kroese FGM, *et al.* Towards personalised treatment in primary Sjögren's syndrome: baseline parotid histopathology predicts responsiveness to rituximab treatment. *Ann Rheum Dis* 2016;75:1933-8.
3. Delli K, Haacke EA, Kroese FGM, *et al.* In primary Sjögren's syndrome high absolute numbers and proportions of B cells in parotid glands predict responsiveness to rituximab as defined by ESSDAI, but not by SSRI. *Ann Rheum Dis* 2016;75:e34.
4. Cornec D, Costa S, Devauchelle-Pensec V, *et al.* Blood and salivary-gland BAFF-driven B-cell hyperactivity is associated to rituximab inefficacy in primary Sjögren's syndrome. *J Autoimmun* 2016;67:102-10.
5. Cornec D, Costa S, Devauchelle-Pensec V, *et al.* Do high numbers of salivary gland-infiltrating B cells predict better or worse outcomes after rituximab in patients with primary Sjögren's syndrome? *Ann Rheum Dis* 2016;75:2016-8.
6. Pijpe J, Meijer JM, Bootsma H, *et al.* Clinical and histologic evidence of salivary gland restoration supports the efficacy of rituximab treatment in Sjögren's syndrome. *Arthritis Rheum* 2009;60:3251-6.
7. Haacke EA, van der Vegt B, Meiners PM, *et al.* Abatacept treatment of patients with primary Sjögren's syndrome results in a decrease of germinal centres in salivary gland tissue. *Clin Exp Rheumatol* 2017;35:317-20.
8. Bowman SJ, Everett CC, O'Dwyer JL, *et al.* Randomized Controlled Trial of Rituximab and Cost-Effectiveness Analysis in Treating Fatigue and Oral Dryness in Primary Sjögren's Syndrome. *Arthritis Rheumatol* 2017;69:1440-50.
9. Jousse-Joulin S, Devauchelle-Pensec V, Cornec D, *et al.* Brief Report: Ultrasonographic assessment of salivary gland response to rituximab in primary Sjögren's syndrome. *Arthritis Rheumatol* 2015;67:1623-8.
10. Delli K, Arends S, van Nimwegen JF, *et al.* Ultrasound of the major salivary glands is a reliable imaging technique in patients with clinically suspected primary Sjögren's syndrome. *Ultraschall Med* 2018;39:328-33.
11. 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.

12. 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;77:556-562.
13. Hocevar A, Ambrozic A, Rozman B, *et al.* 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. Cornec D, Jousse-Joulin S, Costa S, *et al.* High-grade salivary-gland involvement, assessed by histology or ultrasonography, is associated with a poor response to a single rituximab course in primary Sjögren's syndrome: Data from the TEARS randomized trial. *PLoS One* 2016;11:1-11.
15. Mossel E, Delli K, van Nimwegen JF, *et al.* The parotid gland connection: ultrasound and biopsies in primary Sjögren's syndrome. *Ann Rheum Dis* 2018;77:e38.
16. Bootsma H, Kroese FGM, Vissink A. Editorial: Rituximab in the treatment of Sjögren's syndrome: is it the right or wrong drug? *Arthritis Rheumatol* 2017;69:1346-9.

CHAPTER 8

Abatacept treatment for patients with early active primary Sjögren's syndrome: a single-centre, randomised, double-blind, placebo-controlled, phase 3 trial (ASAP-III study)

Jolien F van Nimwegen¹, Esther Mossel¹, Greetje S van Zuiden¹, Robin F Wijnsma¹, Konstantina Delli², Alja J Stel¹, Bert van der Vegt³, Erlin A Haacke^{1,3}, Lisette Olie⁴, Leonoor I Los⁴, Gwenny M Verstappen¹, Sarah A Pringle¹, Fred KL Spijkervet², Frans GM Kroese¹, Arjan Vissink², Suzanne Arends¹, Hendrika Bootsma¹.

¹ Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

² Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

³ Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

⁴ Department of Ophthalmology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

Edited version of *Lancet Rheumatology* 2020;3:e153-e163

Abstract

Background.

Several small open-label studies have suggested efficacy of abatacept—a co-stimulation inhibitor—in patients with primary Sjögren’s syndrome. These promising results warranted further evaluation. We therefore aimed to further assess the safety and efficacy of abatacept compared with placebo in patients with primary Sjögren’s syndrome.

Methods.

We did a single-centre, randomised, double-blind, placebo-controlled, phase 3 trial at the University Medical Center Groningen (Groningen, Netherlands). We included patients with primary Sjögren’s syndrome fulfilling the American European Consensus Group criteria, aged 18 years or older, with positive salivary gland biopsies, time from diagnosis of 7 years or less, and a European League Against Rheumatism Sjögren’s Syndrome Disease Activity Index (ESSDAI) score of 5 or more. Independent pharmacists randomly allocated patients (1:1) to either the abatacept group or placebo group using a computer-generated sequence stratified by previous use of disease-modifying anti-rheumatic drugs. Patients received at-home subcutaneous injections of abatacept (125 mg) or placebo once a week for 24 weeks. The primary outcome was the between-group difference in ESSDAI score at week 24. Efficacy was analysed in patients who received at least one drug dose and for whom post-baseline data were collected. Safety was analysed in all patients who received at least one drug dose.

Findings.

Between Aug 14, 2014, and Aug 23, 2018, 580 patients were reviewed for eligibility, of which 80 patients were randomly assigned to receive study treatment. Efficacy was analysed in 40 patients receiving abatacept and 39 patients receiving placebo (one patient in this group was lost to follow-up). The primary outcome did not significantly differ between the treatment groups. The adjusted mean difference in ESSDAI score at week 24 between the abatacept group and placebo group was -1.3 (95% CI -4.1 to 1.6). No deaths or treatment-related serious adverse events occurred. In 38 (95%) of 40 patients in the abatacept group, 103 adverse events occurred, including one serious adverse event and 46 infections. In 38 (95%) of 40 patients in the placebo group, 87 adverse events occurred, including four serious adverse events and 49 infections.

Interpretation.

On the basis of this trial, we cannot recommend abatacept treatment as standard of care to reduce systemic disease activity in patients with primary Sjögren's syndrome. Further studies should evaluate whether patients with specific clinical manifestations and biological characteristics might benefit from abatacept treatment.

Research in context

Evidence before this study

No systemic treatment options have been approved for patients with primary Sjögren's syndrome. Several studies confirmed the efficacy and safety of abatacept in rheumatoid arthritis. We searched PubMed with the terms "Sjögren's syndrome", "abatacept", and "therapy" to identify trials of abatacept therapy for primary Sjögren's syndrome up to Dec 8, 2019. No randomised, double-blind trials were found. Three small open-label studies were identified. The first study, in which intravenous abatacept was given to 15 patients with primary Sjögren's syndrome, short disease duration, and active disease, reported improvements in the European League Against Rheumatism Sjögren's Syndrome Disease Activity Index (ESSDAI) and Patient-Reported Index (ESSPRI), laboratory parameters, fatigue, and health-related quality of life. The second study reported amelioration of symptoms in some patients, although no validated questionnaires were used. The third study reported improvement of ESSDAI and salivary flow.

Added value of this study

The Abatacept Sjögren Active Patients phase III (ASAP-III) study is, to our knowledge, the first randomised, double-blind trial to compare the efficacy and safety of subcutaneous administration of abatacept with placebo over 24 weeks in patients with primary Sjögren's syndrome that have a short disease duration and active systemic disease. Abatacept therapy did not result in a significant difference in ESSDAI score compared with placebo treatment at week 24. No significant differences in ESSPRI score were found between abatacept and placebo-treated patients, although abatacept-treated patients showed slightly larger improvements in ESSPRI score, resulting in a higher proportion of ESSPRI responders. Abatacept attenuated B-cell hyperactivity, as shown by a decrease in IgG and rheumatoid factor concentrations. Sexual function in female patients was improved in the abatacept group, but no difference was seen for dryness symptoms, fatigue, and quality of life. Salivary and tear gland function were not significantly improved. In this trial, abatacept treatment showed a good safety profile (no deaths or suspected unexpected serious adverse reaction was reported) in patients with primary Sjögren's syndrome.

Implications of all the available evidence

On the basis of our results, and considering the high cost of abatacept, we cannot recommend the use of abatacept as standard of care to alleviate systemic disease activity in patients with primary Sjögren's syndrome. Further studies could assess whether patients with specific clinical manifestations and biological characteristics might benefit from abatacept treatment and whether longer treatment duration will improve efficacy of abatacept

Introduction

Primary Sjögren's syndrome is a systemic autoimmune disease, characterised by lymphocytic infiltration of the salivary and tear glands, with an estimated prevalence of 0.06%.^{1,2} Patients with primary Sjögren's syndrome experience a wide range of symptoms, including glandular enlargement, sicca symptoms, disabling fatigue, arthritis, skin involvement, renal and lung involvement, and peripheral neuropathy. Despite the major impact of this syndrome on health-related quality of life and socioeconomic status,³ treatment for primary Sjögren's syndrome mostly focuses on symptom relief, and an unmet need exists for systemic treatment options. Efficacy of traditional and biological disease-modifying anti-rheumatic drugs (DMARDs) in primary Sjögren's syndrome has not been confirmed in large randomised controlled trials.⁴

Abatacept is a fully human soluble fusion protein of cytotoxic T-lymphocyte antigen 4 coupled to the Fc tail of IgG. This biological DMARD prevents T-cell activation by inhibiting co-stimulation. Because T-cell-dependent B-cell hyperactivity has a central role in the pathogenesis of primary Sjögren's syndrome, inhibition of co-stimulation is a promising treatment approach.⁵ In an open-label study of intravenous abatacept in 15 patients with primary Sjögren's syndrome (time since diagnosis ≤ 5 years and active disease), patients showed improvements in the European League Against Rheumatism Sjögren's Syndrome Disease Activity Index (ESSDAI) and the European League Against Rheumatism Patient Reported Index (ESSPRI), laboratory parameters, fatigue, and health-related quality of life after 24 weeks.⁶ Another open-label study found that some patients reported amelioration of their symptoms during 28 weeks of abatacept treatment, although no validated questionnaires were used.⁷ In 2019, a third group reported improvement of ESSDAI and salivary flow after 24 months of intravenous abatacept treatment.⁸ Abatacept has also shown efficacy in an open-label study of patients with rheumatoid arthritis and associated Sjögren's syndrome.⁹

These promising results warranted further evaluation of abatacept for the treatment of primary Sjögren's syndrome. We therefore aimed to assess the safety and efficacy of subcutaneous abatacept compared with placebo in patients with primary Sjögren's syndrome.

Patients and Methods

Study design and participants

We did a single-centre, randomised, double-blind, placebo-controlled, phase 3 trial (the Abatacept Sjögren Active Patients phase III [ASAP-III] study). This study was done in the multidisciplinary tertiary referral expertise centre for primary Sjögren's syndrome at the University Medical Center Groningen (UMCG; Groningen, Netherlands). We obtained ethical approval from the UMCG institutional review board (METc2014.118). Patients were selected from the UMCG population or referred by other Dutch clinics for trial participation.

Our main inclusion criteria for patients for this study were fulfilment of the American-European Consensus Group criteria for primary Sjögren's syndrome,¹⁰ age 18 years or older, positive gland biopsy, time since diagnosis of 7 years or less, and an ESSDAI score of 5 or more.¹¹ Anti-Sjögren's-syndrome-related antigen A (anti-SSA/Ro) positivity was not required. Retrospectively, all patients also fulfilled the American College of Rheumatology (ACR)-European League Against Rheumatism (EULAR) Classification criteria.¹² Patients who had used prednisone (>10 mg/day), non-biological DMARDs including hydroxychloroquine, or pilocarpine for 1 month or less before enrolment were excluded. Previous use of biological DMARDs was not allowed. 6 months after trial commencement, the protocol was amended to allow patients who previously used abatacept (≥6 months before inclusion) or rituximab (≥12 months before inclusion) to participate, as some of these patients showed high disease activity and were eager to join our trial. The protocol (online supplement) shows a full list of the inclusion and exclusion criteria. We obtained written informed consent from all participants.

Randomisation and masking

Patients were screened for eligibility (JFvN, EM, GSvZ, RFW, HB). Following enrolment, participants received a study number and were randomly allocated (1:1) to either the abatacept group or placebo group by independent pharmacists according to a computer-generated sequence, which was only known to the pharmacists. Allocation was stratified by previous use of DMARDs (including hydroxychloroquine), using block randomisation with a block size of four. Block size was pre-specified in the protocol and therefore could be known to the investigators and outcome assessors. Abatacept and placebo injections were identical in appearance. Participants, investigators, outcome assessors, and care providers were masked to the treatment groups, until the last patient completed week 24 and the database was locked.

Procedures

Patients received instructions to administer subcutaneous injections at home, once a week, for 24 weeks. Injections contained 125 mg of abatacept or placebo. Concomitant use of pilocarpine or DMARDs was not permitted. A stable low dose of prednisone (≤ 10 mg) was permitted. Non-steroidal anti-inflammatory drugs and topical treatments were allowed, but patients were asked to discontinue non-steroidal anti-inflammatory drugs 3 days before each visit, and eye drops 1 h before ocular examinations. Rescue therapy with prednisone or cyclophosphamide was permitted after week 12. Patients visited the UMCG at baseline and weeks 4, 8, 12, and 24. Clinical assessments including laboratory parameters and safety outcomes were done by JFvN, EM, GSvZ, RFW, or HB at each visit. At each visit, patients also completed online questionnaires. Glandular function tests were done by oral medicine specialists and ophthalmologists during the screening visit, week 12, and week 24. Patients from both treatment groups received open-label abatacept for another 24 weeks after the double-blind phase. Safety and efficacy results of the extension phase will be analysed separately.

Outcomes

The primary outcome was the difference in ESSDAI score between the abatacept and placebo groups at week 24.¹¹ Secondary clinical outcomes were ESSDAI score at other timepoints, physician global assessment of disease activity using a numeric rating scale (range 0-10), and Disease Activity Score 28 joint count (DAS-28). Secondary laboratory outcomes were rheumatoid factor, IgG, IgA, IgM, erythrocyte sedimentation rate, and complement C3 and C4. Secondary patient-reported outcomes were the ESSPRI score,¹³ patient global assessment of disease activity, and Patient Acceptable Symptom State at each timepoint; ocular and oral dryness (numeric rating scale [range 0-10]), Multidimensional Fatigue Inventory, Short-Form 36 (SF-36), EuroQoL five dimensions health status questionnaire (EQ-5D-5L), and Work Participation and Activity Impairment Questionnaire (WPAI) at weeks 12 and 24. Physical and mental component scores of the SF-36 were calculated using the QualityMetric Health Outcomes Scoring Software, version 5.1. EQ-5D-5L index values were calculated using syntax provided by EuroQoL, based on a set of weights for the Netherlands. Additionally, EuroQoL five dimensions (EQ-5D) visual analogue scale for general health was used as an outcome. For female participants, patient-reported outcomes included vaginal dryness (numeric rating scale [range 0-10]) at weeks 12 and 24, and Female Sexual Function Index at week 24. Secondary glandular function outcomes were unstimulated whole salivary flow and citric acid-stimulated whole salivary flow, ocular staining score,¹⁴ tear break-up time, and

Schirmer's test without anaesthesia. A full description of the study outcomes is included in the protocol (online supplement).

Exploratory outcomes included salivary gland ultrasound, histology, gut microbiome, and additional laboratory outcomes, which will be analysed and reported separately to provide an in-depth translational evaluation of the efficacy of abatacept.

ESSDAI and ESSPRI response rates were not included in the original protocol, as the minimally clinically important change of ESSDAI and ESSPRI had not yet been published. These were added as secondary outcomes in the statistical analysis plan before unmasking. ESSDAI response was defined as a decrease of three points or more from baseline, and ESSPRI response as a decrease of one point or more or a decrease of 15%.¹⁵ Post-hoc analysis included the evaluation of Sjögren Syndrome Responder Index, which defines response as a 30% improvement or more from baseline in two or more of the five domains (fatigue, oral dryness, ocular dryness, unstimulated whole salivary flow, and erythrocyte sedimentation rate).¹⁶

Data for adverse events were collected by clinicians from baseline to week 24, by open-ended questioning and clinical examination, and clinicians specifically inquired about infections. Investigators assessed all adverse events for severity and potential causality.¹⁷ Adverse events were coded according to the Medical Dictionary for Regulatory Activities (version 19.0). Other safety endpoints included laboratory tests and treatment discontinuation.

Statistical analysis

The sample size was based on a difference of two points in the ESSDAI score between the abatacept group and placebo group, as the minimally clinically important change of three points or more had not yet been determined at the time the protocol was written.¹⁵ A sample size of 44 patients per treatment group was needed to detect a two point difference in ESSDAI (estimated SD of three points on the basis of results from previous trials^{6,18,19}) between the abatacept group and placebo group at 24 weeks, with a two-sided 5% significance level and power of 80%, allowing for a 15% dropout rate. However, after enrolment of 61 patients, the dropout rate appeared to be much lower (<5%), after which the predefined sample size was adjusted to 40 patients per group, in consultation with our data and safety monitoring board. An independent data and safety monitoring board reviewed unblinded study data once a year to monitor safety and overall conduct of the clinical trial. There were no predefined stopping rules. No formal interim analysis for efficacy was done.

Before unmasking, a detailed analysis plan was prepared (online supplement). According to the modified intention-to-treat principle, all participants who recei-

ved at least one dose of study medication were included in the efficacy analysis, irrespective of protocol violations, with exception of one patient for whom no post-baseline efficacy data were available (lost to follow-up before week 4). We planned to do a per-protocol analysis of the primary endpoint only if there was a difference of more than 10% between the modified intention-to-treat and per-protocol population. The safety analysis included all participants who received at least one dose of study medication.

Data collected during visits were considered non-valid and coded as missing when three or more injections were skipped within 4 weeks before a visit, when cyclophosphamide was used, or when a dose of 5 mg or more of prednisone or equivalent was used within 2 weeks before the visit, unless corticosteroid dose was stable since baseline.

We considered p values less than 0.05 statistically significant. For the primary endpoint and all secondary efficacy outcomes, except the Patient Acceptable Symptom State and response according to ESSPRI, ESSDAI, and Sjögren's Syndrome Responder Index, the difference between abatacept and placebo groups was evaluated using linear generalised estimating equations. Missing data were not imputed. The generalised estimating equations model included previous DMARD use, baseline values of the efficacy outcome, treatment (abatacept or placebo), visits, and interactions of treatment by visits. In case of residuals with non-Gaussian distribution, the variable was transformed (second power for ESSPRI total, ESSPRI fatigue, ESSPRI pain, and oral and vaginal dryness; square root for IgG; and natural logarithm for erythrocyte sedimentation rate). Different correlation structures (exchangeable, M-dependent, unstructured) were tested and the model with the lowest information criterion was used, which was the exchangeable correlation structure for all variables. The comparison of main interest was the difference between groups at 24 weeks (treatment-by-visit interaction). Differences at 4, 8, and 12 weeks were assessed to investigate early treatment efficacy. Differences in Patient Acceptable Symptom State between treatment groups were similarly evaluated using logistic generalised estimating equations.

Response according to ESSDAI, ESSPRI, and Sjögren's Syndrome Responder Index were evaluated using binary logistic regression with treatment and previous use of disease-modifying anti-rheumatic drugs as factors. For response analyses, missing values were imputed as non-responses. Baseline characteristics, ESSDAI subdomains, and safety endpoints were summarised descriptively.

We did all the statistical analyses using IBM SPSS Statistics (version 23). This study is registered with ClinicalTrials.gov, number NCT02067910.

Role of the funding source

The funder of the study (Bristol-Myers Squibb) had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all data in the study and had final responsibility for the decision to submit for publication.

Results

Between Aug 14, 2014, and Aug 23, 2018, 580 patients were reviewed for eligibility of which 81 patients were enrolled (figure 1). One enrolled patient was excluded before being randomly allocated to a treatment group because their disease activity became too severe to be treated with study medication (abatacept or placebo) between the screening and baseline visits. After randomisation, 40 patients received abatacept and 40 received placebo. The last visit of the double-blind phase was completed on Feb 21, 2019. Table 1 shows the baseline characteristics.

Table 1

Baseline characteristics of participants

Characteristics	Abatacept (n=40)	Placebo (n=40)	Total (n=80)
Age (years)	48 (15)	49 (16)	49 (16)
Men	3 (8%)	3 (8%)	6 (8%)
Women	37 (93%)	37 (93%)	74 (93%)
Time since diagnosis (years)	2 (0-4)	2 (1-4)	2 (1-4)
Time since symptom started (years)	11 (4-14)	8 (4-14)	8 (4-14)
Anti-SSA/Ro	34 (85%)	37 (93%)	71 (89%)
Anti-SSB/La	20 (50%)	23 (58%)	43 (54%)
Baseline SWS ≥ 0.05 ml/min	31 (78%)	28 (70%)	59 (74%)
Previous use of DMARDs	18 (45%)	16 (40%)	34 (43%)
<i>Oral corticosteroids</i>	10 (25%)	10 (25%)	20 (25%)
<i>Hydroxychloroquine</i>	15 (38%)	14 (35%)	29 (36%)
<i>Methotrexate</i>	3 (8%)	2 (5%)	5 (6%)
<i>Abatacept^a</i>	2 (5%)	3 (8%)	5 (6%)
<i>Rituximab^a</i>	1 (3%)	0	1 (1%)

Characteristics	Abatacept (n=40)	Placebo (n=40)	Total (n=80)
Current medication use			
<i>Corticosteroids</i>	0	1 (3%)	1 (1%)
<i>NSAIDs</i>	16 (40%)	11 (28%)	27 (34%)
<i>Artificial tears</i>	32 (80%)	33 (83%)	65 (81%)
<i>Corticosteroid eye-drops</i>	6 (15%)	5 (13%)	11 (14%)
<i>Artificial saliva</i>	7 (18%)	7 (18%)	14 (18%)
ESSDAI subdomain activity^b			
<i>Constitutional</i>	20 (50%)	17 (43%)	37 (46%)
<i>Lymphadenopathy</i>	10 (25%)	13 (33%)	23 (29%)
<i>Glandular</i>	36 (90%)	37 (93%)	73 (91%)
<i>Articular</i>	23 (58%)	24 (60%)	47 (59%)
<i>Cutaneous</i>	11 (28%)	7 (18%)	18 (23%)
<i>Pulmonary</i>	3 (8%)	2 (5%)	5 (6%)
<i>Renal</i>	0	0	0
<i>Muscular</i>	1 (3%)	0	1 (1%)
<i>Peripheral nervous system</i>	2 (5%)	5 (13%)	7 (9%)
<i>Central nervous system</i>	0	0	0
<i>Haematological</i>	14 (35%)	20 (50%)	34 (43%)
<i>Biological</i>	32 (80%)	31 (78%)	63 (79%)

Data are mean (SD), median (IQR), or n (%). Some percentages do not add up to 100 because of rounding.

^a Patients with previous abatacept use were participants in the open-label ASAP study.⁶ Time between previous treatment with rituximab or abatacept and inclusion was at least 2 years.

^b Total number of patients with low, moderate or high activity in ESSDAI subdomains. DMARDs: disease-modifying anti-rheumatic drugs; ESSDAI: EULAR Sjögren syndrome disease activity index; NSAIDs: non-steroidal anti-inflammatory drugs; SWS: stimulated whole salivary flow.

Figure 1. Trial profile. ^a Numbers of patients fulfilling different exclusion criteria are not available because patients often fulfilled more than one criterion and only the first one that was detected was noted. ^b One patient in the placebo group was excluded from the efficacy analysis because she was lost to follow-up in week 3 and no post-baseline was collected.
ESSDAI: European League Against Rheumatism Sjögren's Syndrome Disease Activity Index; ITT: intention-to-treat; MALT: mucosa-associated lymphoid tissue; SCLE: subacute cutaneous lupus erythematosus.

None of the abatacept-treated patients withdrew from treatment or were lost to follow-up (figure 1). Two placebo-treated patients withdrew from treatment because of adverse events, of whom one was excluded from the efficacy analysis, as no post-baseline data were available. The modified intention-to-treat population consisted of 40 patients in the abatacept group and 39 in the placebo group. The per-protocol population consisted of 40 patients in the abatacept group and 37 in the placebo group. As the difference between modified intention-to-treat and per-protocol populations was less than 10%, no per-protocol analysis was done.

Week 12 data for one patient in the placebo group were coded as missing for the efficacy analysis, as this patient skipped three injections between week 8 and week 12. One patient used a maintenance dose of prednisone (5 mg/day) throughout the trial because of autoimmune hepatitis, which was in remission at inclusion. Although rescue therapy was allowed after week 12 according to protocol, two patients received corticosteroids before week 12, intended as bridging therapy while awaiting the effect of study treatment. The first patient received abatacept and used corticosteroids between week 10 and week 17 because of active cutaneous involvement. The second patient received placebo and was treated with corticosteroids between weeks 2-10 and 12-24 for severe arthritis and myositis. Data from week 12 of the first patient and data from weeks 4, 8, and 24 of the second patient were therefore coded as missing for the efficacy analysis. No additional patients received rescue therapy after week 12. After coding non-valid data as missing, all variables had less than 5% missing values per visit, with exception of IgG at week 4, ocular staining score, tear break-up time, Schirmer's test, unstimulated whole salivary flow, and stimulated whole salivary flow at week 12, and stimulated whole salivary flow at week 24, which had 6-10% missing values.

Table 2.

Baseline values and differences between groups in week 12 and 24

Variable	Baseline		Week 12			Week 24		
	Abatacept	Placebo	AD	95% CI	P value	AD	95% CI	P value
ESSDAI score	14.0 (9.0-16.8)	13.0 (8.0-18.0)	-2.4	-4.7 to -0.1	0.039	-1.3	-4.1 to 1.6	0.385
ESSPRI score ^a	7.0 (5.4-7.7)	7.3 (5.3-8.0)	-4.7	-12.5 to 3.0	0.232	-5.0	-12.7 to 2.8	0.208
<i>Dryness</i>	7.0 (5.3-8.0)	7.0 (7.0-8.0)	-0.3	-1.0 to 0.3	0.311	-0.3	-0.9 to 0.4	0.455
<i>Fatigue</i> ^a	7.5 (7.0-8.0)	8.0 (6.0-9.0)	-2.0	-12.5 to 8.6	0.716	-2.6	-14.1 to 9.0	0.664
<i>Pain</i> ^a	7.0 (5.0-8.0)	7.0 (3.0-8.0)	-4.5	-14.7 to 5.7	0.389	-5.2	-14.5 to 4.2	0.279
Ocular dryness (NRS)	6.0 (3.3-7.0)	7.0 (6.0-8.0)	0.4	-0.5 to 1.2	0.388	0.0	-0.8 to 0.9	0.972
Oral dryness (NRS) ^a	7.0 (5.0-8.0)	7.0 (6.0-8.0)	-3.8	-11.6 to 4.1	0.345	-0.6	-10.1 to 8.9	0.902
Vaginal dryness (NRS) ^a	5.0 (3.0-8.0)	6.5 (5.0-8.8)	-1.4	-10.2 to 7.3	0.751	-0.1	-10.0 to 9.7	0.981
Patient GDA	8.0 (7.0-8.0)	7.0 (7.0-9.0)	-0.8	-1.7 to 0.0	0.063	-0.4	-1.3 to 0.6	0.442
MFI								
<i>General fatigue</i>	16.5 (13.3-19.0)	16.0 (13.0-20.0)	-0.4	-2.0 to 1.2	0.583	-0.8	-2.8 to 1.2	0.448
<i>Physical fatigue</i>	15.5 (14.0-17.0)	15.0 (13.0-18.0)	-0.1	-1.4 to 1.2	0.907	-1.3	-3.0 to 0.4	0.144
<i>Reduced activity</i>	14.0 (11.0-15.0)	13.0 (11.0-16.0)	-0.5	-1.9 to 1.0	0.509	-0.4	-2.1 to 1.3	0.619
<i>Reduced motivation</i>	12.0 (9.0-13.0)	11.0 (9.0-14.0)	-0.6	-2.2 to 1.0	0.487	-1.3	-2.9 to 0.2	0.099
<i>Mental fatigue</i>	12.0 (9.3-15.0)	13.0 (8.0-16.0)	-0.9	-2.2 to 0.5	0.216	-0.4	-1.6 to 0.9	0.554

Variable	Baseline			Week 12		Week 24		
	Abatacept	Placebo	AD	95% CI	P value	AD	95% CI	P value
Short Form 36								
PCS	37.1 (7.9)	38.2 (9.8)	2.5	-0.2 to 5.2	0.065	2.0	-0.7 to 4.7	0.141
MCS	44.0 (38.0-52.1)	44.4 (36.5-51.6)	-1.9	-5.5 to 1.8	0.312	-1.0	-4.6 to 2.6	0.594
EQ-5D-5L	0.71 (0.60-0.80)	0.71 (0.50-0.79)	-0.04	-0.11 to 0.03	0.273	-0.03	-0.10 to 0.04	0.343
EQ-5D VAS	60.5 (42.3-70.0)	60.0 (45.0-71.0)	6.2	-1.5 to 13.9	0.117	3.1	-5.0 to 11.3	0.454
FSFI	11.5 (4.4-23.2)	14.7 (5.1-23.4)	NA	NA	NA	3.8	0.1 to 7.4	0.042
DAS-28 (ESR)	4.82 (1.19)	5.00 (1.48)	-0.37	-0.80 to 0.06	0.089	-0.46	-1.02 to 0.09	0.101
DAS-28 (CRP)	4.35 (2.63-5.08)	4.20 (2.90-5.10)	-0.44	-0.87 to -0.02	0.041	-0.46	-1.00 to 0.08	0.097
Physician GDA	5.7 (1.5)	5.7 (1.6)	-0.7	-1.3 to -0.1	0.022	-0.6	-1.4 to 0.2	0.144
IgG (g/L) ^a	17.4 (13.4-26.7)	18.7 (14.8-24.7)	-0.1	-0.2 to 0.0	0.215	-0.1	-0.2 to -0.01	0.028
RF (IU/mL)	32.5 (2.1-71.0)	24.0 (6.8-83.0)	-11.7	-16.6 to -6.8	<0.0001	-13.8	-20.7 to -6.0	<0.0001
Complement C3 (g/L)	1.2 (1-1.3)	1.1 (1.0-1.3)	0.0	-0.1 to 0.1	0.564	0.0	-0.1 to 0.1	0.451
Complement C4 (g/L)	0.19 (0.09)	0.20 (0.07)	0.01	-0.01 to 0.03	0.393	0.01	-0.01 to 0.03	0.381
ESR (mm/h) ^a	28.0 (13.3-47.0)	33.0 (17.0-54.0)	-0.01	-0.20 to 0.18	0.881	-0.20	-0.41 to 0.02	0.068
Ocular staining score ^b	4.0 (0.5-6.5)	4.5 (2.0-7.0)	-0.9	-1.9 to 0.1	0.071	0.0	-1.0 to 1.0	0.993

Variable	Baseline		Week 12			Week 24		
	Abatacept	Placebo	AD	95% CI	P value	AD	95% CI	P value
Schirmer's test ^b	3.5 (0.6-14.0)	2.5 (0.0-8.5)	0.1	-2.9 to 3.2	0.927	1.1	-1.2 to 3.3	0.367
TBUT (s)	5.3 (2.5-7.5)	4.0 (2.0-7.0)	-0.4	-1.7 to 0.9	0.548	-0.6	-2.1 to 0.9	0.435
UWS (mL/min)	0.05 (0.01-0.12)	0.05 (0.01-0.13)	-0.01	-0.04 to 0.03	0.744	0.02	-0.02 to 0.05	0.375
SWS (mL/min)	0.16 (0.06-0.33)	0.10 (0.02-0.43)	0.03	-0.05 to 0.10	0.481	0.01	-0.08 to 0.09	0.861

Baseline data are mean (SD) or median (IQR). Raw outcome data for week 12 and week 24 are in the appendix (p 3). The adjusted difference is the unstandardised regression coefficient from the linear generalised estimating equations models, which represents the difference between the treatment groups (abatacept–placebo), adjusted for baseline values and previous disease-modifying anti-rheumatic drug use. Vaginal dryness and the Female Sexual Function Index include only female patients. ^aRaw values were transformed before generalised estimating equations analysis. Estimates of the adjusted differences in weeks 12 and 24 therefore represent differences on the transformed scale (square root for IgG, natural logarithm for ERS and second power for ESSPRI total, ESSPRI fatigue, ESSPRI pain, oral dryness, and vaginal dryness). ^bAverage of right and left eye.

AD: Adjusted difference; CRP: C-reactive protein; DAS-28: disease activity score 28 joint count; ESR: erythrocyte sedimentation rate; ESSDAI: European League Against Rheumatism Sjögren Syndrome Disease Activity Index; ESSPRI: European League Against Rheumatism Sjögren Syndrome Patient Reported Index; EQ-5D: EuroQoL five dimensions; EQ-5D-5L: EuroQoL five dimensions health status questionnaire; FSFI: Female Sexual Function Index; GDA: global disease activity; MCS: mental component summary; MFI: Multidimensional Fatigue Inventory; NA: not available; NRS: numerical rating scale; PCS: physical component summary; RF: rheumatoid factor; SWS: stimulated whole salivary flow; TBUT: tear break-up time; UWS: unstimulated whole salivary flow; VAS: visual analogue scale.

Figure 2. ESSDAI score and secondary efficacy outcomes. The difference between the treatment groups are adjusted for baseline values and previous disease-modifying anti-rheumatic drug use. Error bars are IQR for medians, SD for means, and 95% CI for the ADs. AD: adjusted difference; CRP: C-reactive protein; DAS-28: Disease Activity Score 28 joint count; ESSDAI: European League Against Rheumatism Sjögren's Syndrome Disease Activity Index; ESSPRI: European League Against Rheumatism Sjögren's Syndrome Patient-Reported Index.

The primary endpoint (the difference in ESSDAI score at week 24) did not significantly differ between treatment groups (table 2, figure 2, supplementary table 1). Differences in ESSDAI score between treatment groups, adjusted for baseline values and previous DMARD use, were -2.4 (95% CI -4.7 to -0.1) at week 12 and -1.3 (-4.1 to 1.6) at week 24. ESSDAI score at week 12 was significantly lower in patients receiving abatacept than in those receiving placebo. No significant differences were found in proportions of patients reaching the minimally clinically important change for ESSDAI (figure 3). At week 12, 19 (48%) of 40 patients in the abatacept group and 13 (33%) of 39 in the placebo group were ESSDAI responders (odds ratio [OR] 1.8, 95% CI 0.7-4.5; $p=0.207$). At week 24, 23 (58%) of 40 patients in the abatacept group and 20 (51%) of 39 in the placebo group were ESSDAI responders (OR 1.3, 95% CI 0.5-3.1; $p=0.600$). The largest decrease in activity was seen in the articular, glandular, and constitutional ESSDAI subdomains (figure 3, supplementary table 2). Improvements in these domains occurred in both treatment groups, although improvement in the articular domain was larger in the abatacept group. ESSDAI scores of individual patients over time are shown in supplementary figure 1.

No significant differences between groups were found for ESSPRI score (table 2, figure 2, supplementary table 1), although abatacept did increase the number of patients reaching the minimally clinically important change for ESSPRI (figure 3). At week 12, 21 (53%) of 40 patients in the abatacept group and 11 (28%) of 39 in the placebo group were ESSPRI responders (OR 2.8, 95% CI 1.1-7.1; $p=0.031$). At week 24, 23 (58%) patients in the abatacept group and eight (21%) in the placebo group were ESSPRI responders (OR 5.7, 95% CI 2.0-15.7; $p=0.001$).

Figure 3. Categorical efficacy outcomes. (A) Response according to the ESSDAI, ESSPRI, and Sjögren's Syndrome Responder Index. (B) ESSDAI subdomain and total activity. ESSDAI response is a decrease of three points or more from baseline.¹⁵ ESSPRI response is a decrease of one point or more, or a decrease of 15% from baseline.¹⁵ Response according to the Sjögren Syndrome Responder Index is an improvement of 30% or more from baseline in two or more of the five domains (fatigue, oral dryness, ocular dryness, unstimulated whole salivary flow, and erythrocyte sedimentation rate).¹⁶ ESSDAI: European League Against Rheumatism Sjögren's Syndrome Disease Activity Index; ESSPRI: European League Against Rheumatism Sjögren's Syndrome Patient-Reported Index.

The Female Sexual Function Index at week 24 was significantly better in patients receiving abatacept than in those receiving placebo (adjusted difference 3.8, 95% CI 0.1-7.4; $p=0.042$; table 2, supplementary table 1, supplementary figure 2). No differences were found for oral, ocular, and vaginal dryness, patient global disease activity, fatigue (Multidimensional Fatigue Index domains), or health-related quality of life (SF-36 physical and mental component scores, EQ-5D-5L, EQ-5D visual analogue scale) at any timepoint. No differences were found in proportions of patients with acceptable symptoms at week 12 (OR 0.4, 95% CI 0.1-1.7; $p=0.221$) or week 24 (1.1, 0.2-5.3; $p=0.893$; supplementary table 1). As only 20 (50%) of 40 patients in the abatacept group and 21 (54%) of 39 patients in the placebo group had paid employment at baseline, the number of patients with complete WPAs was small and differences in WPAI outcomes were not statistically tested. At week 24, 21 (53%) of 40 patients in the abatacept group and 19 (53%) of 36 patients in the placebo group had paid employment.

Abatacept significantly decreased physician global disease activity compared with placebo at week 12 (adjusted difference -0.7, 95% CI -1.3 to -0.1; $p=0.022$), but not at week 24 (table 2, figure 2, supplementary table 1). Although both DAS-28 (C-reactive protein) and DAS-28 (erythrocyte sedimentation rate) were decreased by abatacept treatment, a significant difference was only seen in DAS-28 (C-reactive protein) at week 12 (adjusted difference -0.44, 95% CI -0.87 to -0.02; $p=0.041$). IgG was significantly lower in patients receiving abatacept than in those receiving placebo at week 24. IgA and IgM did not significantly differ between groups (data not shown). Rheumatoid factor was significantly lower in patients in the abatacept group than in those in the placebo group at week 4 and all following visits. With the exception of rheumatoid factor, other laboratory efficacy outcomes did not show significant differences at weeks 4 and 8. No differences were seen in erythrocyte sedimentation rate or complement. No significant differences between treatment groups were found in tear or salivary gland function (table 2, figure 2, supplementary table 1).

Table 3.

Summary of adverse events by intervention group

	Number of events		Number of patients with events	
	Abatacept	Placebo	Abatacept (n=40)	Placebo (n=40)
Death	0	0	0	0
SUSAR	0	0	0	0
SAE (total) ^a	1	4	1 (3%)	4 (10%)
SAE with possible relation to intervention ^b	0	1	0	1 (3%)
AE, severe ^c	1	3	1 (3%)	3 (8%)
AE, moderate ^d	3	7	3 (8%)	7 (18%)
AE, total	103	87	38 (95%)	38 (95%)
AE with possible relation to intervention ^b	80	66	36 (90%)	32 (80%)
Infection	46	49	29 (73%)	28 (70%)
Treatment withdrawal due to AE	0	2	0	2 (5%)
Temporary treatment discontinuation due to AE	6	12	5 (13%)	8 (20%)

Data are n (%), unless otherwise specified.

^a A definition of serious adverse events is shown in the protocol (online supplement).^b Possible, probable, or definite relation to treatment, as assessed by the investigators during the trial.^c Severe adverse events were defined as those that make activities of daily living impossible and usually require treatment or other interventions.^d Moderate adverse events were defined as those that limit activities of daily living and for which treatment or another intervention is sometimes necessary.

AE: adverse event; SAE: serious adverse event; SUSAR: suspected unexpected serious adverse reaction.

No deaths or unexpected serious adverse reaction occurred (table 3). One serious adverse event occurred in the abatacept group, which was deemed not treatment related, whereas four serious adverse events occurred in the placebo group (supplementary table 3). None of the patients receiving abatacept withdrew from treatment because of adverse events. One patient in the placebo group discontinued treatment in week 3 because of the development of autoimmune hepatitis. Another patient in the placebo group discontinued treatment in week 8 because of severe exacerbation of disease activity, with glomerulonephritis, pulmonary involvement, cutaneous vasculitis, and development of high anti-double-stranded DNA concentrations, after which the patient's diagnosis was changed to Sjögren's syndrome combined with systemic lupus erythematosus (SLE). A total of 103 adverse events (including one serious adverse event and 46 infections) occurred in 38 (95%) of 40 patients in the abatacept group compared with 87 adverse events (including four serious adverse events and 49 infections) in 38 (95%) of 40 in the placebo group. Most adverse events were mild. The most common adverse events were infections, occurring in 29 (73%) patients in the abatacept group and 28 (70%) in the placebo group (supplementary table 4). General and administration site conditions (including fatigue, malaise, and pyrexia) occurred in 12 (30%) patients in the abatacept group and four (10%) in the placebo group. Gastrointestinal disorders (including abdominal pain, dyspepsia, nausea, and diarrhoea) occurred in 11 (28%) patients in the abatacept group and five (13%) in the placebo group. Laboratory safety outcomes are shown in supplementary figure 3.

Post-hoc analysis showed significantly more responders according to the Sjögren's Syndrome Responder Index in the abatacept group than in the placebo group at week 12, but not week 24 (figure 3). In week 12, 16 (40%) of 40 patients in the abatacept group and five (13%) of 39 in the placebo group were responders (OR 4.5, 95% CI 1.5-14.1; $p=0.009$). At week 24, 13 (33%) patients in the abatacept group and nine (23%) in the placebo group were responders (OR 1.7, 95% CI 0.6-4.6; $p=0.311$).

Discussion

In the ASAP-III study, we found no significant difference between the abatacept group and placebo group in the primary endpoint of ESSDAI score after 24 weeks of treatment. Abatacept significantly decreased ESSDAI score compared with placebo at week 12, although no difference between treatment groups was found in the proportion of patients reaching the minimally clinically important change of three points or more in ESSDAI. Physician global disease activity was significantly lower in abatacept-treated patients than in placebo-treated patients only at week 12. These results indicate that while disease activity was lowered in both treatment groups, improvements in ESSDAI and physician global disease activity were seen at earlier timepoints in patients receiving abatacept.

In previous open-label trials, ESSDAI score was significantly decreased during abatacept treatment.^{6,8} In the current trial, improvement of ESSDAI score was found in both treatment groups. The improvements seen over 24 weeks in the placebo and abatacept groups might partly be explained by regression to the mean. A study in which patients were treated with standard of care therapy showed that high baseline ESSDAIs were associated with improvement of ESSDAI after 12 months.²⁰ As only patients with moderate or high ESSDAI scores were included in the ASAP-III study, a natural decrease in ESSDAI scores is expected to occur for some patients, regardless of the treatment group.

Sjögren's syndrome is a highly heterogeneous disease. Therefore, not all patients might respond equally to a particular DMARD, as was shown for rituximab.²¹ Our trial was not powered to show effects on specific systemic manifestations. Further studies should assess whether abatacept is effective in patients with specific characteristics—e.g., high articular activity.

Besides ESSDAI, the patient symptom index ESSPRI, which includes questions regarding sicca symptoms, fatigue, and pain, is an important outcome, because both indices are complementary. Abatacept significantly increased the number of ESSPRI responders, based on the minimally clinically important change, compared with placebo at weeks 12 and 24, despite the absence of significant differences in ESSPRI score (which is a continuous variable). This finding shows that although abatacept-treated patients reached minimally clinically important change more often, the difference between groups was small. Female abatacept-treated patients showed improved sexual function at week 24, which might be important, considering the negative effect of primary Sjögren's syndrome on sexual function.²² No differences were found in other patient-reported outcomes including global disease activity, dryness, fatigue, and health-related quality of life.

No significant differences were found in glandular function. During our previous open-label trial,⁶ glandular function also did not improve, but stimulated whole salivary flow remained stable during treatment and deteriorated after treatment discontinuation. Abatacept reduced the number of germinal centres in parotid gland biopsies, but no changes were found in focus score or size of infiltration.²³ As a recent open-label study found improvement of salivary flow after 24 months of abatacept treatment,⁸ longer treatment might be needed to improve salivary gland function. Evaluation of long-term efficacy of abatacept on glandular function in the ASAP-III open-label extension phase, and the effect on ultrasound and histological findings, will be of interest.

In line with previous results,⁶ abatacept attenuated B-cell hyperactivity, as reflected by the decrease in rheumatoid factor and IgG. This decrease in B-cell hyperactivity might be caused by effects of abatacept on T-follicular helper (Tfh) cells.²⁴ Abatacept decreased the number of circulating Tfh-cells, and expression of inducible T-cell co-stimulator, a marker of T-cell activation.²⁴ Decreased expression of inducible T-cell co-stimulator was associated with ESSDAI improvement,²⁴ Abatacept also decreased Bruton's tyrosine kinase concentrations in naive and memory B cells.²⁵ Unfortunately, this biological effect did not translate into improvement of systemic disease activity, raising the question whether future trials should focus on different treatment targets or on different clinical endpoints, or both. Ongoing in-depth analysis of the biological effects of abatacept might contribute to a better understanding of the molecular pathways involved in primary Sjögren's syndrome.

The choice of endpoints can greatly influence the conclusion of a trial, as previously described in SLE.²⁶ Although development of the ESSDAI, 10 years ago, has been a major step forward in the assessment of primary Sjögren's syndrome, it has certain limitations as a primary endpoint. The ESSDAI does not reflect prominent symptoms experienced by patients (sicca, fatigue, and pain). Furthermore, it is not always possible to separate signs of active disease from irreversible damage, and consequently some of the ESSDAI domains are likely not to improve, such as the pulmonary and peripheral nervous system domains. Finally, the ESSDAI can be insensitive to partial improvements within subdomains, which has also been described for similar outcome measures in SLE.²⁷ For example, a patient with an IgG of more than 20 g/L might show a reduction of IgG, but if IgG does not drop below 20 g/L, the ESSDAI biological subdomain activity does not improve. For these reasons, it is of great importance to develop validated composite endpoints for primary Sjögren's syndrome, which have high sensitivity to change, and adequately reflect disease activity as well as symptoms of primary Sjögren's syndrome. The composite endpoint Sjögren Syndrome Responder Index was created and validated

using data from trials with rituximab and infliximab, and includes fatigue, oral and ocular dryness, unstimulated whole salivary flow, and erythrocyte sedimentation rate.¹⁶ At week 12, the number of responders for the Sjögren Syndrome Responder Index was significantly higher in the abatacept group than in the placebo group. The development of a validated response index is the focus of the recently initiated NECESSITY project.²⁸

Abatacept was well tolerated by patients with primary Sjögren's syndrome during the ASAP-III study. No treatment-related serious adverse events or treatment withdrawals occurred in the abatacept group. Prevalence of infections was not increased during abatacept treatment, and most adverse events were mild.

The results of a sponsor-initiated, multicentre trial of abatacept treatment for primary Sjögren's syndrome have recently been presented.²⁹ Similar to our results, despite favourable effects on biological activity, abatacept was no better than placebo for improving ESSDAI, ESSPRI, or stimulated whole salivary flow after 24 weeks of treatment. There are some important differences between the two studies. The sponsor-initiated trial allowed hydroxychloroquine as concomitant treatment and included only patients positive for SSA/Ro with no limit to disease duration; by contrast, in our investigator-initiated trial, the use of hydroxychloroquine was not allowed and only patients with positive biopsies and disease duration of 7 years or less were included. The number of patients using concomitant corticosteroid treatment in the ASAP-III study was very low compared with previous trials of primary Sjögren's syndrome.

Our trial has some limitations. Because our sample size was based on the primary endpoint, we cannot rule out that clinically relevant differences in secondary endpoints with large variance, such as salivary flow rate,³⁰ were not detected with the current sample size. Second, because of the large number of secondary outcomes, some statistically significant findings might result from chance. Third, as none of the included patients had renal disease or CNS involvement, the effect of abatacept on these domains could not be assessed. Finally, only a small proportion of our patient population was eligible for participation. Efficacy of abatacept treatment in other patients with primary Sjögren's syndrome—e.g., with longer disease duration or low ESSDAI scores—remains unknown.

To conclude, abatacept was well tolerated by patients with primary Sjögren's syndrome. Despite the beneficial biological effects of abatacept, the ASAP-III study does not support the use of abatacept as standard of care to reduce systemic disease activity. As primary Sjögren's syndrome is a highly heterogeneous disease, further studies could evaluate whether patients with specific clinical manifestations and biological characteristics might benefit from abatacept treatment.

Declaration of interests

JFvN is a speaker and consultant for Bristol-Myers Squibb. GSvZ is a speaker for Roche. FGMK received an unrestricted grant from Bristol-Myers Squibb, is a consultant for Bristol-Myers Squibb, and a speaker for Bristol-Myers Squibb, Roche, and Janssen-Cilag. HB received unrestricted grants from Bristol-Myers Squibb and Roche; is a consultant for Bristol-Myers Squibb, Roche, Novartis, Medimmune, and Union Chimique Belge; a speaker for Bristol-Myers Squibb and Novartis; and is a member of the advisory board of Bristol-Myers Squibb, Novartis, and Sanofi. All other authors declare no competing interests.

Acknowledgments

This study was funded by Bristol-Myers Squibb. Abatacept and placebo injections were provided by Bristol-Myers Squibb free of charge. We thank all patients for participation in our trial. We thank N J G M Veeger, C A Stegeman, and E N van Roon for their participation in our data safety monitoring board, and H Strooisma for monitoring our study. We also thank J Bulthuis-Kuiper, E P Herder, and R Tjemmes for their logistical support during the study. Finally, we thank our rheumatology colleagues in the Netherlands for referral of patients to participate in our trial.

References

1. Qin B, Wang J, Yang Z, *et al.* Epidemiology of primary Sjögren's syndrome: a systematic review and meta-analysis. *Ann Rheum Dis* 2014;74:1983–9.
2. Brito-Zerón P, Baldini C, Bootsma H, *et al.* Sjögren syndrome. *Nat Rev Dis Prim* 2016;2:16047.
3. Meijer JM, Meiners PM, Huddleston Slater JJ, *et al.* Health-related quality of life, employment and disability in patients with Sjögren's syndrome. *Rheumatology (Oxford)* 2009;48:1077–82.
4. Price EJ, Baer AN. How to treat Sjögren's syndrome. *Rheumatology (Oxford)*; published online first: 15 February 2019.
5. Verstappen GM, Kroese FGM, Bootsma H. T cells in primary Sjögren's syndrome: targets for early intervention. *Rheumatology (Oxford)*; published online first: 15 February 2019.
6. Meiners PM, Vissink A, Kroese FGM, *et al.* Abatacept treatment reduces disease activity in early primary Sjögren's syndrome (open-label proof of concept ASAP study). *Ann Rheum Dis* 2014;73:1–4.
7. Adler S, Körner M, Förger F, *et al.* Evaluation of histological, serological and clinical changes in response to abatacept treatment of primary Sjögren's syndrome: A pilot study. *Arthritis Care Res* 2013;65:1862–8.
8. Machado AC, dos Santos LC, Fidelix T, *et al.* Effectiveness and safety of abatacept for the treatment of patients with primary Sjögren's syndrome. *Clin Rheumatol* 2020;39:243–8.
9. Tsuboi H, Matsumoto I, Hagiwara S, *et al.* Efficacy and safety of abatacept for patients with Sjögren's syndrome associated with rheumatoid arthritis: rheumatoid arthritis with orencia trial toward Sjögren's syndrome Endocrinopathy (ROSE) trial-an open-label, one-year, prospective study-Interim analysis of 32 patients for 24 weeks. *Mod Rheumatol* 2015;25:187–93.
10. 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.
11. Seror R, Ravaud P, Bowman SJ, *et al.* EULAR Sjögren's syndrome disease activity index: development of a consensus systemic disease activity index for primary Sjögren's syndrome. *Ann Rheum Dis* 2010;69:1103–9.

12. 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.
13. Seror R, Ravaud P, Mariette X, *et al.* EULAR Sjögren's Syndrome Patient Reported Index (ESSPRI): development of a consensus patient index for primary Sjögren's syndrome. *Ann Rheum Dis* 2011;70:968–72.
14. Whitcher JP, Shiboski CH, Shiboski SC, *et al.* A simplified quantitative method for assessing keratoconjunctivitis sicca from the Sjögren's Syndrome International Registry. *Am J Ophthalmol* 2010;149:405–15.
15. Seror R, Bootsma H, Saraux A, *et al.* Defining disease activity states and clinically meaningful improvement in primary Sjögren's syndrome with EULAR primary Sjögren's syndrome disease activity (ESSDAI) and patient-reported indexes (ESSPRI). *Ann Rheum Dis* 2016;75:382–9.
16. Cornec D, Devauchelle-Pensec V, Mariette X, *et al.* Development of the Sjögren's Syndrome Responder Index, a data-driven composite endpoint for assessing treatment efficacy. *Rheumatology (Oxford)* 2015;54:1699–708.
17. Naranjo CA, Busto U, Sellers EM, *et al.* A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther* 1981;30:239–45.
18. Meiners PM, Arends S, Brouwer E, *et al.* Responsiveness of disease activity indices ESSPRI and ESSDAI in patients with primary Sjögren's syndrome treated with rituximab. *Ann Rheum Dis* 2012;71:1297–302.
19. Meijer JM, Meiners PM, Vissink A, *et al.* Effectiveness of rituximab treatment in primary Sjögren's syndrome: a randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2010;62:960–8.
20. Baldini C, Ferro F, Luciano N, *et al.* Clinically Meaningful Improvement of Essdai and Esspri in Patients with Primary Sjögren's Syndrome in Real Life: A 12-Month Longitudinal Study. *Arthritis Rheumatol* 2016;68(suppl 10) (abstr).
21. Delli K, Haacke EA, Kroese FGM, *et al.* Towards personalised treatment in primary Sjögren's syndrome: baseline parotid histopathology predicts responsiveness to rituximab treatment. *Ann Rheum Dis* 2016;75:1933–8.
22. van Nimwegen JF, Arends S, van Zuiden GS, *et al.* The impact of primary Sjögren's syndrome on female sexual function. *Rheumatology (Oxford)* 2015;54:1286–93.

23. Haacke EA, van der Vegt B, Meiners PM, *et al.* Abatacept treatment of patients with primary Sjögren's syndrome results in a decrease of germinal centres in salivary gland tissue. *Clin Exp Rheumatol* 2017;35:317–20.
24. Verstappen GM, Meiners PM, Corneth OBJ, *et al.* Attenuation of Follicular Helper T Cell-Dependent B Cell Hyperactivity by Abatacept Treatment in Primary Sjögren's Syndrome. *Arthritis Rheumatol* 2017;69:1850–61.
25. Corneth OBJ, Verstappen GMP, Paulissen SMJ, *et al.* Enhanced Bruton's Tyrosine Kinase Activity in Peripheral Blood B Lymphocytes From Patients With Autoimmune Disease. *Arthritis Rheumatol* 2017;69:1313–24.
26. Wofsy D, Hillson JL, Diamond B. Abatacept for lupus nephritis: Alternative definitions of complete response support conflicting conclusions. *Arthritis Rheum* 2012;64:3660–5.
27. Miles A, Pope JE. A comparison of rheumatoid arthritis and systemic lupus erythematosus trial design: a commentary on ways to improve the number of positive trials in SLE. *Clin Exp Rheumatol* 2015;33:671–80.
28. New clinical endpoints in primary Sjögren's Syndrome: an interventional trial based on stratifying patients (NECESSITY). *Innovative Medicines Initiative*. Grant agreement number 806975.
29. Baer A, Gottenberg J-E, St. Clair WE, *et al.* OP0039 Efficacy and safety of abatacept in active primary Sjögren's syndrome: results of a randomised placebo-controlled phase III trial. *Ann Rheum Dis* 2019;78(suppl 2) (abstr).
30. Burlage FR, Pijpe J, Coppes RP, *et al.* Variability of flow rate when collecting stimulated human parotid saliva. *Eur J Oral Sci* 2005;113:386–90.

Supplementary table 1.

Efficacy outcomes in abatacept and placebo treated patients.

	Baseline		Week 12		Week 24	
	Abatacept	Placebo	Abatacept	Placebo	Abatacept	Placebo
ESSDAI score	14.0 (9.0-16.8)	13.0 (8.0-18.0)	9.0 (5.0-12.0)	10.5 (7.0-16.0)	8.0 (4.0-14.0)	8.0 (5.0-14.5)
ESSPRI score	7.0 (5.4-7.7)	7.3 (5.3-8.0)	6.0 (4.3-7.0)	6.3 (5.2-7.8)	6.0 (4.4-7.3)	6.7 (4.6-7.9)
Dryness	7.0 (5.3-8.0)	7.0 (7.0-8.0)	6.0 (4.0-7.0)	7.0 (6.0-8.0)	6.0 (4.0-7.0)	7.0 (5.0-8.0)
Fatigue	7.5 (7.0-8.0)	8.0 (6.0-9.0)	7.0 (4.0-8.0)	7.0 (5.0-8.0)	6.0 (5.0-8.0)	7.0 (4.3-8.0)
Pain	7.0 (5.0-8.0)	7.0 (3.0-8.0)	6.0 (3.0-7.0)	6.0 (3.5-8.0)	6.0 (4.3-8.0)	6.0 (3.3-7.8)
Ocular dryness NRS	6.0 (3.3-7.0)	7.0 (6.0-8.0)	6.0 (4.0-7.0)	7.0 (5.5-8.0)	5.0 (3.3-7.0)	7.0 (5.0-8.8)
Oral dryness NRS	7.0 (5.0-8.0)	7.0 (6.0-8.0)	6.0 (4.0-8.0)	7.0 (5.5-8.0)	6.0 (4.0-7.0)	6.0 (5.0-8.0)
Vaginal dryness NRS	5.0 (3.0-8.0)	6.5 (5.0-8.8)	5.0 (3.0-7.0)	5.5 (3.0-8.0)	5.0 (3.0-7.5)	6.0 (3.0-8.0)
Patient GDA	8.0 (7.0-8.0)	7.0 (7.0-9.0)	6.0 (5.0-7.0)	7.0 (6.0-8.0)	6.0 (5.0-7.8)	7.0 (5.0-8.0)
MFI						
General fatigue	16.5 (13.3-19.0)	16.0 (13.0-20.0)	15.0 (12.0-18.0)	14.0 (12.0-19.0)	15.5 (13.0-17.8)	16.0 (11.3-18.8)
Physical fatigue	15.5 (14.0-17.0)	15.0 (13.0-18.0)	14.0 (12.0-16.0)	14.0 (11.0-17.0)	14.0 (11.3-16.0)	14.5 (11.0-18.0)
Reduced activity	14.0 (11.0-15.0)	13.0 (11.0-16.0)	13.0 (10.0-14.0)	12.0 (8.0-16.0)	12.5 (9.3-15.0)	13.0 (10.3-16.0)
Reduced motivation	12.0 (9.0-13.0)	11.0 (9.0-14.0)	11.0 (7.0-14.0)	10.0 (8.0-14.0)	10.0 (8.0-12.0)	12.0 (7.3-14.0)

	Baseline		Week 12		Week 24	
	Abatacept	Placebo	Abatacept	Placebo	Abatacept	Placebo
Mental fatigue	12.0 (9.3-15.0)	13.0 (8.0-16.0)	12.0 (8.0-15.0)	14.0 (7.5-17.0)	12.0 (9.5-14.0)	13.0 (8.0-16.0)
SF36						
PCS	37.1 (7.9)	38.2 (9.8)	40.1 (8.6)	39.1 (9.6)	39.3 (8.6)	38.7 (11.1)
MCS	44.0 (38.0-52.1)	44.4 (36.5-51.6)	49.2 (39.3-52.2)	47.3 (40.8-55.6)	47.0 (39.7-53.7)	50.0 (42.4-54.0)
EQ-5D-5L	0.71 (0.60-0.80)	0.71 (0.50-0.79)	0.71 (0.65-0.81)	0.77 (0.67-0.84)	0.74 (0.57-0.82)	0.75 (0.59-0.81)
EQ-5D VAS	60.5 (42.3-70.0)	60.0 (45.0-71.0)	63.0 (51.0-72.0)	59.0 (39.5-73.0)	61.0 (52.0-75.0)	56.5 (40.3-67.8)
FSFI	11.5 (4.4-23.2)	14.7 (5.1-23.4)	NA	NA	18.8 (5.1-25.7)	14.9 (3.7-22.2)
DAS-28 (ESR)	4.82 (1.19)	5.00 (1.48)	4.04 (1.13)	4.58 (1.55)	3.74 (0.88)	4.30 (1.29)
DAS-28 (CRP)	4.35 (2.63-5.08)	4.20 (2.90-5.10)	3.00 (2.20-4.00)	3.20 (2.73-4.38)	2.50 (2.10-3.40)	3.10 (2.30-3.90)
Physician GDA	5.7 (1.5)	5.7 (1.6)	4.1 (1.7)	4.7 (2.0)	3.7 (1.9)	4.3 (2.0)
IgG, g/L	17.4 (13.4-26.7)	18.7 (14.8-24.7)	16.4 (12.7-26.4)	18.4 (14.8-24.2)	17.0 (12.9-26.0)	19.0 (13.7-25.5)
RF, IU/ml	32.5 (2.1-71.0)	24.0 (6.8-83.0)	17.0 (1.9-49.0)	27.0 (9.7-89.5)	17.5 (1.7-42.0)	29.0 (8.0-90.0)
Complement C3, g/L	1.2 (1-1.3)	1.1 (1.0-1.3)	1.2 (1.0-1.4)	1.2 (1.0-1.3)	1.2 (1.0-1.4)	1.1 (1.0-1.2)
Complement C4, g/L	0.19 (0.09)	0.20 (0.07)	0.20 (0.08)	0.21 (0.07)	0.21 (0.07)	0.20 (0.08)
ESR, mm/hour	28.0 (13.3-47.0)	33.0 (17.0-54.0)	35.0 (11.0-60.0)	30.5 (21.3-59.0)	30.0 (15.0-57.0)	44.0 (21.0-66.0)
OSS ^a	4.0 (0.5-6.5)	4.5 (2.0-7.0)	2.0 (0.5-5.3)	3.5 (2.0-6.4)	3.0 (1.0-6.4)	3.5 (1.3-7.3)

	Baseline		Week 12		Week 24	
	Abatacept	Placebo	Abatacept	Placebo	Abatacept	Placebo
Schirmer's test ^a	3.5 (0.6-14.0)	2.5 (0.0-8.5)	4.3 (0.4-9.1)	2.5 (0.0-6.3)	5.3 (2.1-10.3)	1.0 (0.0-4.0)
TBUT, seconds ^a	5.3 (2.5-7.5)	4.0 (2.0-7.0)	4.0 (2.9-7.6)	4.0 (2.6-7.9)	4.0 (2.5-6.9)	4.0 (2.0-7.5)
UWS, ml/min	0.05 (0.01-0.12)	0.05 (0.01-0.13)	0.07 (0.01-0.15)	0.03 (0.00-0.17)	0.06 (0.01-0.15)	0.04 (0.01-0.10)
SWS, ml/min	0.16 (0.06-0.33)	0.10 (0.02-0.43)	0.21 (0.05-0.41)	0.15 (0.03-0.35)	0.20 (0.08-0.43)	0.10 (0.03-0.29)
PASS ^b , n (%)	16 (40%)	17 (44%)	18 (46%)	23.0 (62%)	18 (45%)	17 (47%)
ESSDAI responder, n (%)	NA	NA	19 (48%)	13 (33%)	23 (58%)	20 (51%)
ESSPRI responder, n (%)	NA	NA	21 (53%)	11 (28%)	23 (58%)	8 (21%)
SSRI responder, n (%)	NA	NA	16 (40%)	5 (13%)	13 (33%)	9 (23%)

Values are mean (SD), median (25th-75th percentile) unless otherwise indicated. Vaginal dryness and FSFI include only female patients (n=74). ^a Average of right and left eye. ^b Number (%) of patients with acceptable symptoms. CRP: C-reactive protein; DAS-28: disease activity score 28 joint count; ESR: erythrocyte sedimentation rate; ESSDAI: EULAR Sjögren syndrome disease activity index; ESSPRI: EULAR Sjögren syndrome patient-reported index; EQ-5D-5L: summary index value of the five level version of the EuroQoL five dimensions health status questionnaire; FSFI: Female Sexual Function Index; GDA: global disease activity; IgG: immunoglobulin G; NA: not applicable; NRS: numeric rating scale (range 0-10); MCS: Mental component summary; MFI: multidimensional fatigue inventory; OSS: ocular staining score; PASS: patient acceptable symptom state; PCS: physical component summary; RF: rheumatoid factor; SF-36: Short-Form 36; SSRI: Sjögren syndrome responder index; SWS: stimulated whole salivary flow; UWS: unstimulated whole salivary flow; VAS: visual analogue scale (range 0-100).

Supplementary table 2.

ESSDAI subdomain and total activity.

Domains	Baseline	Week 12		Week 24		
	Abatacept	Placebo	Abatacept	Placebo	Abatacept	Placebo
Activity	n = 40	n = 39	n = 39	n = 37 ^a	n = 40	n = 37
Constitutional						
None	20 (50)	23 (59)	26 (67)	20 (54)	29 (73)	25 (68)
Low	16 (40)	9 (23)	10 (26)	12 (32)	8 (20)	10 (27)
Moderate	4 (10)	7 (18)	3 (8)	5 (14)	3 (8)	2 (5)
Lymphadenopathy						
None	30 (75)	26 (67)	35 (90)	30 (81)	35 (88)	31 (84)
Low	8 (20)	12 (31)	3 (8)	6 (16)	4 (10)	6 (16)
Moderate	2 (5)	1 (3)	1 (3)	1 (3)	1 (3)	0
High	0	0	0	0	0	0
Glandular						
None	4 (10)	3 (8)	14 (36)	12 (32)	18 (45)	15 (41)
Low	15 (38)	19 (49)	14 (36)	11 (30)	14 (35)	14 (38)
Moderate	21 (53)	17 (44)	11 (28)	14 (38)	8 (20)	8 (22)
Articular						
None	17 (43)	16 (41)	25 (64)	16 (43)	29 (73)	23 (62)
Low	5 (13)	5 (13)	3 (8)	5 (14)	0	2 (5)
Moderate	5 (13)	7 (18)	8 (21)	9 (24)	6 (15)	6 (16)
High	13 (33)	11 (28)	3 (8)	7 (19)	5 (13)	6 (16)
Cutaneous						
None	29 (73)	32 (82)	32 (82)	31 (84)	34 (85)	31 (84)
Low	2 (5)	4 (10)	1 (3)	2 (5)	0	3 (8)
Moderate	8 (20)	3 (8)	6 (15)	4 (11)	6 (15)	3 (8)
High	1 (3)	0	0	0	0	0
Pulmonary						
None	37 (93)	37 (95)	36 (92)	36 (97)	37 (93)	35 (95)

Domains	Baseline		Week 12		Week 24	
	Abatacept	Placebo	Abatacept	Placebo	Abatacept	Placebo
Low	0	0	0	0	0	0
Moderate	3 (8)	2 (5)	2 (5)	1 (3)	2 (5)	2 (5)
High	0	0	1 (3)	0	1 (3)	0
Muscular						
None	39 (98)	39 (100)	38 (97)	37 (100)	39 (98)	37 (100)
Low	1 (3)	0	1 (3)	0	1 (3)	0
Moderate	0	0	0	0	0	0
High	0	0	0	0	0	0
PNS						
None	38 (95)	34 (87)	37 (95)	32 (86)	38 (95)	33 (89)
Low	2 (5)	4 (10)	2 (5)	4 (11)	2 (5)	4 (11)
Moderate	0	1 (3)	0	1 (3)	0	0
High	0	0	0	0	0	0
Hematologic						
None	26 (65)	19 (49)	27 (69)	18 (50)	25 (63)	13 (35)
Low	12 (30)	18 (46)	10 (26)	17 (47)	12 (30)	22 (59)
Moderate	2 (5)	2 (5)	2 (5)	1 (3)	3 (8)	2 (5)
High	0	0	0	0	0	0
Biological						
None	8 (20)	9 (23)	8 (21)	10 (28)	10 (25)	10 (27)
Low	8 (20)	6 (15)	8 (21)	7 (19)	9 (23)	7 (19)
Moderate	24 (60)	24 (62)	23 (59)	19 (53)	21 (53)	20 (54)
Total ESSDAI						
Low	0	0	6 (15)	8 (22)	14 (35)	8 (22)
Moderate	19 (48)	23 (59)	24 (62)	15 (42)	14 (35)	19 (51)
High	21 (53)	16 (41)	9 (23)	13 (36)	12 (30)	10 (27)

Values are number (%) of patients. * N=36 for ESSDAI biological, hematologic and total score due to missing laboratory parameters. ESSDAI: EULAR Sjögren's syndrome disease activity index. PNS: peripheral nervous system.

Supplementary table 3.

Listing of serious adverse events

System organ class	Lower level term	Related to treatment ^a
Abatacept		
Injury, poisoning and procedural complications	Clavicle fracture	Doubtful
Placebo		
Immune system disorders	SLE flare	Doubtful
Infections and infestations	S. pneumoniae pneumonia	Possible
Nervous system disorders	Lumbar radiculopathy	Doubtful
Respiratory, thoracic and mediastinal disorders	Embolism lung	Doubtful

No suspected unexpected serious adverse reactions occurred during the study. ^a As assessed by the investigators during the trial according to Naranjo criteria. SLE: systemic lupus erythematosus.

Supplementary table 4.

Listing of adverse events per system organ class.

System organ class	Number of events		Number of patients (%) with events	
	Abatacept	Placebo	Abatacept (n=40)	Placebo (n=40)
Infections and infestations	46	49	29 (72.5)	28 (70)
General and administration site conditions	14	4	12 (30)	4 (10)
Gastrointestinal disorders	12	6	11 (27.5)	5 (12.5)
Nervous system disorders	8	4	7 (17.5)	3 (7.5)
Musculoskeletal and connective tissue disorders	7	5	7 (17.5)	5 (12.5)
Respiratory, thoracic and mediastinal disorders	4	1	4 (10)	1 (2.5)
Cardiac disorders	3	4	3 (7.5)	4 (10)
Skin and subcutaneous tissue disorders	3	1	3 (7.5)	1 (2.5)
Injury, poisoning and procedural complications	2	1	2 (5)	1 (2.5)
Investigations	2	2	2 (5)	2 (5)
Hepatobiliary disorders	1	1	1 (2.5)	1 (2.5)
Psychiatric disorders	1	1	1 (2.5)	1 (2.5)
Eye disorders	0	2	0	2 (5)
Immune system disorders	0	1	0	1 (2.5)
Metabolism and nutrition disorders	0	1	0	1 (2.5)
Reproductive system and breast disorders	0	3	0	3 (7.5)
Surgical and medical procedures	0	1	0	1 (2.5)

System organ classes in which no events occurred were not shown in this table.

Supplementary figure 1.

ESSDAI scores of individual abatacept and placebo patients.

ESSDAI: EULAR Sjögren's syndrome disease activity index.

Supplementary figure 2.

Secondary patient-reported efficacy outcomes. The upper parts of all figures show medians or means of the abatacept and placebo group. The lower parts shows the difference between groups (abatacept-placebo), adjusted for baseline values and previous DMARD use. Error bars represent IQR for medians, SD for means and 95% CI for adjusted differences. AD: Adjusted difference; EQ-5D-5L: five level version of the EuroQoL five dimensions health status questionnaire; FSFI: Female Sexual Function Index; GDA: global disease activity; MCS: Mental component summary; MFI: multidimensional fatigue inventory; PCS: physical component summary; SF-36: Short-Form 36.

Supplementary figure 3.

Laboratory safety parameters. ALAT: Alanine aminotransferase; ASAT: aspartate aminotransferase; CRP: C-reactive protein.

CHAPTER 9

One year in review: recent insights in the potential role of imaging modalities for diagnosing patients with primary Sjögren's syndrome

Esther Mossel¹, Suzanne Arends¹, Hendrika Bootsma¹

¹ Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
Edited version of Clin Exp Rheum 2020;38Suppl126:310-314.

Abstract

Within the last year, interesting developments regarding the assessment of salivary gland involvement in patients with clinical suspicion of, or diagnosed with primary Sjögren's syndrome (pSS) have been performed. In this review various topics will be discussed, starting with the use of salivary gland ultrasonography (SGUS) for the detection of glandular swelling. Furthermore, other imaging modalities, besides B-mode SGUS, that differentiate between pSS patients and healthy controls will be drawn to the attention. Moreover, storage of ultrasonographic images and videos will be discussed briefly, as will be some potential biases and pitfalls. Finally, efforts that have been made to make incorporation of SGUS into the most recent classification criteria possible will be discussed, as well as the important steps that have been taken to develop a new semi-quantitative scoring system for the assessment of salivary gland involvement in patients with suspected or confirmed pSS.

Primary Sjögren's syndrome (pSS) is a common, heterogeneous, systemic auto-immune disease, with an estimated prevalence of 0.04% in the general population.^{1,3} Above all, pSS is characterized by chronic inflammation of the salivary and lacrimal glands.⁴ The most common symptoms are dryness of the mouth and eyes, resulting in xerostomia and keratoconjunctivitis sicca, respectively. Furthermore, fatigue, arthralgia and various extraglandular manifestations can be present at any time during the disease course.^{1,4}

Currently, there is no golden standard to diagnose a patient with pSS, and therefore, diagnosis mainly relies on expert opinion, which is based on the interpretation of several tests and clinical observations. In addition, classification criteria have been developed with the intention to define homogenous study groups, rather than being used for diagnosing a patient, as for diagnosis not only specificity, but also sensitivity should be high.^{4,5} Nevertheless, diagnosis and classification often show quite some overlap, as both rely on the same components.

The most recent set of criteria applied for pSS are the American College of Rheumatology - European League Against Rheumatism (ACR-EULAR) classification criteria.^{6,7} The ACR-EULAR criteria actually are a new mixture of items that were already integrated in the previously developed 2002 American European Consensus Group (AECG) and 2012 ACR, and, endorsed by both, EULAR and ACR.⁸⁻¹⁰ The ACR-EULAR criteria assign three points for a positive biopsy (focus score ≥ 1) and positive anti-SSA/Ro antibodies, and one point for a decreased unstimulated whole salivary (UWS) flow, decreased Schirmer's test, or increased ocular staining score (OSS). Patients with a score of ≥ 4 are classified as pSS.^{6,7} An interesting critical note has recently been raised by van Nimwegen et al.¹⁰ As a matter of fact, the authors uncover that these new 2016 ACR-EULAR criteria are not optimally balanced.¹⁰ Whereas tear gland involvement is evaluated by the Schirmer's test, a functional test, as well as by the OSS, which is an imaging modality to measure structural ocular involvement, salivary gland involvement is solely evaluated by measuring the UWS flow, which can also be categorized as a functional test.¹⁰ An imaging modality to measure salivary gland involvement appears to be missing.

Interestingly, even after many years, salivary gland imaging is still an evolving field within pSS. Previously, the salivary glands were mainly visualized by sialography or scintigraphy, as both methods were also incorporated in the 2002 AECG criteria.^{8,11} Nowadays, these more invasive imaging modalities are, in daily clinical practice, mainly replaced by salivary gland ultrasonography (SGUS). With B-mode SGUS, which is easily accessible in the outpatient setting, repeated measurements can take place, as it is non-irradiating, not invasive, in-expensive, and well-tolerated by patients.^{10,12,13} In this review, we highlight interesting developments within

the last year, including studies investigating the use of SGUS for the detection of glandular swelling, the value of other imaging modalities besides B-mode SGUS to differentiate between pSS patients and healthy controls, storage of SGUS images and videos, some biases and pitfalls, the incorporation of SGUS into classification criteria, and the development of a novel semi-quantitative scoring system for the assessment of salivary gland involvement in patients with suspected or confirmed pSS.

The use of salivary gland ultrasonography for the detection of glandular swelling.

Besides sicca complaints and reduced salivary flow, another typical manifestation of salivary gland involvement in pSS is glandular swelling. Salivary gland enlargement, which is detected in about thirty percent of the patients, is considered a risk factor for the development of mucosa-associated lymphoid tissue (MALT)-lymphoma, especially in patients with persistent enlargements of the glands.¹⁴ Therefore, salivary gland involvement is one of the twelve domains of the EULAR Sjögren's Syndrome Disease Activity Index (ESSDAI), which is nowadays the most frequently used assessment of systemic disease activity.¹⁵ The domain measuring glandular involvement contains three levels of activity i.e., (i) No: absence of glandular swelling; (ii) Low: small glandular swelling with enlarged parotid (≤ 3 cm), or limited submandibular (≤ 2 cm) or lachrymal swelling (≤ 1 cm) and (iii) Moderate: major glandular swelling with enlarged parotid (> 3 cm), or important submandibular (> 2 cm) or lachrymal swelling (< 1 cm).¹⁵

The user guide, developed in 2015, states that 'parotid, submandibular or lachrymal gland swelling should be assessed by clinical examination and not by ultrasound'.¹⁵ Very recently, however, Marteau et al.¹⁶ came to a different conclusion by investigating the interobserver reproducibility of parotid and submandibular gland palpation, and comparing this clinical assessment with B-mode SGUS measurements. In this study, 34 patients diagnosed as pSS or non-SS sicca, according to the AECG criteria were included.¹⁶

Swelling of a parotid gland, measured on a binary scale, could reliably be detected by palpation if the gland is at least 3 cm in length. Of note, a low ESSDAI score on the glandular domain is the case when the parotid gland enlargement is up to 3 cm. Reliable detection of submandibular gland swelling, however, is a complete different story, as physical examination appeared not informative.¹⁶ Using ultrasonography, cut-off values for defining glandular hypertrophy were arbitrarily set at more than 5 cm² for the parotid glands, and more than 3 cm² for the submandibular glands. Herewith, for both investigators, a low agreement between glandular swelling measured by palpation and hypertrophy measured by SGUS was found.

Concerning clinical trials, among others, a complementary method to measure glandular swelling, like ultrasonography, might be warranted, as it appears that, even for pSS experts, physical examination regarding mild glandular swelling might be overestimated.¹⁶ Interestingly, Milic et al. found a significant correlation between B-mode SGUS score and the glandular domain of the ESSDAI in a large group of pSS patients, and suggest that SGUS might be a surrogate item for the glandular domain.¹⁷ Further research on the additional value of SGUS examination on top of clinical palpation of the salivary glands is needed, since the ESSDAI is a widely used measurement tool. Moreover, this information is also important for the development of future composite endpoints. The value of SGUS for early detection of MALT lymphomas has not been determined yet.

B-mode ultrasonography in relation to salivary flow.

Nowadays, there is sufficient evidence that SGUS enables us, at group level, to differentiate between pSS patients and healthy controls, and patients with other connective tissue diseases (CTDs), and even between pSS patients and non-SS sicca patients.¹⁸⁻²² Within the past year, these topics have been investigated further.²³⁻²⁵

Inanc et al. showed that patients with anti-SSA/Ro and/or anti-SSB/La antibodies, and those with a reduced UWS flow rate on average had higher B-mode SGUS scores in comparison with seronegative patients and patients with a normal UWS flow, respectively.²³ The former was also shown by La Paglia et al., who showed that SS patients had higher B-mode SGUS scores compared with other CTD patients with sicca complaints, but without a diagnosis of secondary SS.²⁴

In a recent review, Devauchelle et al.²⁶ emphasized the need to identify SGUS characteristics caused by glandular damage, and to differentiate between potentially reversible markers of active disease and probably irreversible damage. The importance to differentiate between active disease and disease damage was previously pointed out by Baldini et al.,¹¹ and recently underlined by Zabotti et al.²⁵ The latter authors aimed to identify the SGUS component that showed the best association with a decreased salivary flow in pSS patients, and compared 75 established pSS patients with 23 age- and sex-matched healthy controls. Reduced salivary flow showed significant associations with the detection of hyperechoic bands in the parotid and submandibular glands, of which the authors propose that this finding represents glandular damage. In addition, these hyperechoic bands were much rarer in the group of healthy controls, and less prominent.²⁵ More research needs to be done to fill current knowledge gaps, with perhaps also a role for other SGUS modalities.

Imaging modalities other than B-mode ultrasonography.

Recently, also other ultrasound modalities or techniques have been investigated in pSS patients, as well as in healthy controls.²⁷⁻²⁹ Among these modalities are grey-scale histogram analysis, and (two-dimensional shear wave) elastography.²⁷⁻²⁹ With grey-scale histogram analysis of ultrasonographic images, a more objective method to provide quantitative information about the echogenicity of the salivary glands is applied, as a range of 0-255 grey-level pixels is used. Herewith, subjective, operator-dependent ratings of the normally applied B-mode ultrasound evaluation are eliminated.²⁷ With elastography, another ultrasonographic technique, the stiffness of the organ, in this case the salivary glands, can be evaluated²⁷ and two-dimensional shear wave elastography (2D SWE) enables a quantitative assessment of this tissue stiffness (in meters per second and kilopascals).^{29,30}

Caraba et al. found that ultrasonography and sonoelastography are valuable tools in assessing salivary glands involvement in pSS, since the ultrasonographic parameters showed significant differences between 32 pSS patients and 32 age and sex matched healthy controls.²⁸ Furthermore, Arslan et al. showed that 2D SWE values differed significantly between 53 pSS patients and 30 healthy controls, and concluded that salivary glands should be assessed by both, B-mode SGUS and 2D SWE.²⁹ Moreover, Sezer et al. found that the mean histogram values of parotid glands of 57 female pSS patients were significantly lower compared with the histogram values of 48 female healthy controls, whereas previous studies did not find a significant difference between normal and abnormal glands.^{27,31,32} However, the inclusion of healthy controls rather than non-SS sicca patients, makes it easier to detect differences, as it is expected that healthy individuals have normal salivary glands. It would be of interest to investigate whether these promising results could also be found between pSS patients and non-SS sicca controls. This, since an imaging modality should be able to differentiate between these groups, as both are referred to a rheumatologist, and healthy controls usually are not.

Storage and scoring of ultrasonographic images and videos.

Recently, Hammenfors et al. investigated the preferred projection, gland and storage format for post-examination evaluation of SGUS in a group of 32 pSS patients.³³ The authors concluded that there was a trend favoring longitudinal videos of the parotid gland, and that storing enables later image evaluation for diagnostics, second opinion and disease progression. Furthermore, the authors recommend regular calibration exercises among ultrasonographers, because in daily clinical practice, observer differences may not be detected easily.³³ These potential observer differences might be overcome with the use of image segmentation analysis,

providing automatically scored SGUS lesions based upon computer algorithms.¹¹ The application of artificial intelligence is also one of the aims of the HarmonicSS project of joint European Research, the Horizon 2020 project.³⁴ Unfortunately, it is not clear yet what different SGUS features truly represent in terms of histopathological parameters. Therefore, a study comparing SGUS evaluations with major salivary gland biopsies is eagerly awaited.

Incorporation of salivary gland ultrasonography into classification criteria.

Previously, it had already been suggested to add B-mode SGUS to the 2002 AECG criteria or ACR criteria, or to replace one of the current items by SGUS.^{22,35-39} Unfortunately, SGUS was not considered as a potential criteria item in the development of the ACR-EULAR criteria, as its methodological properties were not fully known.^{10,40}

It may be clear from the various efforts that have already been made to investigate incorporation of B-mode SGUS into the 2016 ACR-EULAR classification criteria, that this is eagerly awaited. Within the last year, two independent studies have been published,^{10,40} both investigated this topic, bringing the total number up to four.^{41,42} Despite the methodological differences, there are a few lessons can be learned to ensure that SGUS will become part of the classification criteria anytime soon. First, to keep the ACR-EULAR criteria applicable and easy to use, the weight of the original criteria items should stay the same.^{10,40} Furthermore, like the other items to measure tear- and salivary gland involvement, a positive SGUS should be assigned a weight of one point. Even though, consensus about the amount of points that are necessary to classify a patient must still be obtained, there is consensus that a patient should not be classified as pSS patient if only the minor criteria items are fulfilled.^{10,40} Replacement of serology or histopathology by SGUS resulted in a significant decrease in the performance of the criteria.¹⁰ However, SGUS can be added to the other minor items UWS, the OSS or the Schirmer's test.^{10,40} This improves feasibility in clinical practice, as not all individual tests may be accessible in every outpatient clinic. Previously it has been shown that presence of anti-SSA/Ro antibodies in combination with a positive SGUS is highly predictive of classifying a patient with pSS.¹⁸ Therefore, for classification purposes, the first steps could be determination of antibody status and performing an SGUS. Though for clinical diagnosis, a full work-up is advised, in order to make the best possible risk assessment for the individual patient.^{10,18}

However, before worldwide implementation of SGUS into classification criteria can take place, a few problems need to be solved. International consensus regarding which SGUS scoring system to use, with a validated cut-off point to differentiate between pSS patients and other CTD or non-SS sicca patients has main priority.

Necessarily, a high level of inter- and intra-observer reliability should be pursued. Within the past years, a sub-taskforce of the Outcome Measures in Rheumatology Clinical Trials (OMERACT) working group, created in 2016, has taken important steps in this procedure, as one of the main goals was to develop a standardized SGUS scoring procedure.⁴³ The first step was to achieve agreement on defining normal and abnormal SGUS findings is patients with probable or confirmed pSS. This resulted in the development of a new semi-quantitative scoring system for the assessment of the parotid and submandibular salivary glands, with excellent intra- and good inter-observer reliability.⁴³ During the development process, attention was also given to fatty replacement and fibrosis, both initially not included as potential items in the new scoring system, and it was suggested that both should be considered when the semi-quantitative system cannot be applied.⁴³ Herewith, significant progress has been made in the development of a reliable, semi-quantitative scoring system that can be used worldwide.

To conclude, there are a few key points to keep in mind. First, detection of mild glandular swelling by palpation appears to be overestimated. Therefore, a potential role for SGUS, especially when the ESSDAI is used as outcome parameter in clinical trials, warrants further investigation. Second, grey-scale histogram analysis and elastography should be evaluated in the target population, to further explore their potential value in diagnosing pSS patients. Third, important steps have been taken to develop a novel SGUS scoring system for the assessment of salivary glands in suspected or confirmed pSS. Herewith, inclusion of SGUS into the classification criteria seems even more justified, as this would rebalance the current classification criteria. SGUS examination should be part of the diagnostic work-up of potential pSS patients.

Take home messages

- SGUS examination should be part of the diagnostic work-up of potential pSS patients.
- Inclusion of SGUS into the classification criteria would rebalance the current criteria.
- Important steps have been taken by a sub-task force of the OMERACT working group to develop a novel semi-quantitative SGUS scoring system.
- Regular calibration exercises should be held among ultrasonographers to overcome potential observer bias.
- Grey-scale histogram analysis and elastography should be evaluated in the target population in order to explore their potential value.
- Detection of mild glandular swelling by palpation appears to be overestimated, a potential role for SGUS warrants further investigation.

References

1. Mariette X, Criswell LA. Primary Sjögren's syndrome. *N Engl J Med* 2018;378:931-9.
2. Qin B, Wang J, Yang Z, *et al.* Epidemiology of primary Sjögren's syndrome: a systematic review and meta-analysis. *Ann Rheum Dis* 2015;74:1983-9.
3. Brito-Zerón P, Baldini C, Bootsma H, *et al.* Sjögren syndrome. *Nat Rev Dis Prim* 2016;2:16047
4. Bootsma H, Spijkervet FKL, Kroese FGM, *et al.* Toward new classification criteria for Sjögren's syndrome? *Arthritis Rheum* 2013;65:21-3.
5. Aframian D, Konttinen Y, Carrozzo M, *et al.* Urban Legends series: Sjögren's Syndrome. *Oral Dis* 2013;19:46-58.
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;79 :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. 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.
9. Shiboski SC, Shiboski CH, Criswell LA, *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.
10. 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* 2020;72:583-90.
11. Baldini C, Zabotti A, Filipovic N, *et al.* Review Imaging in primary Sjögren's syndrome : the "obsolete and the new". *Clin Exp Rheumatol* 2018;36Sup-pl112:215-21.
12. Delli K, Dijkstra PU, Stel AJ, *et al.* Diagnostic properties of ultrasound of major salivary glands in Sjögren's syndrome: a meta-analysis. *Oral Dis* 2015;21:792-800.

13. Jousse-Joulin S, Milic V, Jonsson MV, *et al.* Is salivary gland ultrasonography a useful tool in Sjögren's syndrome? A systematic review. *Rheumatology (Oxford)* 2016;55:789-800.
14. Luciano N, Ferro F, Bombardieri S, *et al.* Advances in salivary gland ultrasonography in primary Sjögren's syndrome. *Clin Exp Rheumatol* 2018;36:S159-64.
15. Seror R, Bowman SJ, Brito-Zeron P, *et al.* EULAR Sjögren's syndrome disease activity index (ESSDAI): a user guide. *RMD open* 2015;1:e000022.
16. Marteau P, Cornec D, Gouillou M, *et al.* Assessment of major salivary gland size in primary Sjögren's syndrome: Comparison between clinical examination and ultrasonography. *Jt Bone Spine* 2019;86:627-32.
17. Milic V, Colic J, Cirkovic A, *et al.* Disease activity and damage in patients with primary Sjögren's syndrome: Prognostic value of salivary gland ultrasonography. *PLoS One* 2019;14:1-12.
18. Mossel E, Delli K, van Nimwegen JF, *et al.* Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1883-89.
19. 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. *Rheumatol (Oxford)* 2015;54:2198-204.
20. Hammenfors DS, Brun JG, Jonsson R, *et al.* Diagnostic utility of major salivary gland ultrasonography in primary Sjögren's syndrome. *Clin Exp Rheumatol* 2015;33:56-62.
21. Wernicke D, Hess H, Gromnica-Ihle E, *et al.* Ultrasonography of salivary glands – A highly specific imaging procedure for diagnosis of Sjögren's syndrome. *J Rheumatol* 2008;35:285-93.
22. 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;53:1604-07.
23. Inanc N, Şahinkaya Y, Mumcu G, *et al.* Evaluation of salivary gland ultrasonography in primary Sjögren's syndrome: does it reflect clinical activity and outcome of the disease? *Clin Exp Rheumatol* 2019;37Suppl118:140-5.
24. La Paglia GMC, Sanchez-Pernaute O, Alunno A, *et al.* Ultrasound salivary gland involvement in Sjögren's syndrome vs. other connective tissue diseases: is it autoantibody and gland dependent? *Clin Rheumatol* 2020;39:1207-15.

25. Zabotti A, Callegher SZ, Gandolfo S, *et al.* Hyperechoic bands detected by salivary glands ultrasonography are related to salivary impairment in established Sjögren's syndrome. *Clin Exp Rheumatol* 2019;37Suppl118:146-52.
26. Devauchelle-Pensec V, Zabotti A, Carvajal-Alegria G, *et al.* Salivary gland ultrasonography in primary Sjögren's syndrome: opportunities and challenges. *Rheumatology (Oxford)*; published online first 2019.
27. Sezer I, Erderm Toslak I, Yağci, B, *et al.* The role of real-time tissue elastography and gray-scale ultrasound histogram analysis in the diagnosis of patients with Sjögren's syndrome. *Arch Rheumatol* 2019;34:371-9.
28. Caraba A, Babalic FC, Lurcius S, *et al.* The utility of major salivary gland ultrasonographic parameters in the diagnosis of Sjögren's syndrome. *Dis Markers* 2019;1716848.
29. Arslan S, Durmaz MS, Erdogan H, *et al.* Two-dimensional shear wave elastography in the assessment of salivary gland involvement in primary Sjögren's syndrome. *J Ultrasound Med* 2020;39:949-56.
30. Ren W, Li X, He Y, *et al.* Two-dimensional shear wave elastography of breast lesions: comparison of two different systems. *Clin Hemorheol Microcirc* 2017;66:37-46.
31. Chikui T, Okamura K, Tokumori K, *et al.* Quantitative analyses of sonographic images of the parotid gland in patients with Sjögren's syndrome. *Ultrasound Med Biol* 2006;32:617-22.
32. Chikui T, Shimizu M, Kawazu T, *et al.* A quantitative analysis of sonographic images of the salivary gland: a comparison between sonographic and sialographic findings. *Ultrasound Med Biol* 2009;35:1257-64.
33. Hammenfors DS, Causevic H, Assmus J, *et al.* Assessment of major salivary gland ultrasonography in Sjögren's syndrome. A comparison between bedside and post-examination evaluations. *Clin Exp Rheumatol* 2019;37Suppl118:153-8.
34. Harmonization and integrative analysis of regional, national and international cohorts on primary Sjögren's syndrome (pSS) towards improved stratification, treatment and health policy making (HarmonicSS). *Horizon2020 project*. Grant agreement number 731944.
35. Cornec D, Jousse-Joulin S, Pers J-O, *et al.* Contribution of salivary gland ultrasonography to the diagnosis of Sjögren's syndrome: Toward new diagnostic criteria? *Arthritis Rheum* 2013;65:216-25.

36. Takagi Y, Kimura Y, Nakamura H, *et al.* Salivary gland ultrasonography: can it be an alternative to sialography as an imaging modality for Sjögren's syndrome? *Ann Rheum Dis* 2010;69:1321-4.
37. 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 (Oxford)* 2012;51:1081-5.
38. Yonetsu K, Takagi Y, Sumi M, *et al.* Sonography as a replacement for sialography for the diagnosis of salivary glands affected by Sjögren's syndrome. *Ann Rheum Dis* 2002;61:276-7.
39. Jonsson M, 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.
40. Jousse-Joulin S, Gatineau F, Baldini C, *et al.* Weight of salivary gland ultrasonography compared to other items of the 2016 ACR/EULAR classification criteria for primary Sjögren's syndrome. *J Intern Med* 2020;287:180-8.
41. 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.
42. 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.
43. 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.

CHAPTER 10

General discussion

General discussion

Primary Sjögren's syndrome (pSS) is a chronic, systemic autoimmune disease characterized by inflammation of the salivary and lacrimal glands, leading to xerostomia and keratoconjunctivitis sicca.^{1,2} There is not yet a golden standard to diagnose patients with pSS, and therefore, after a balanced interpretation of various tests and clinical observations, diagnosis is mainly based upon expert opinion. For research purposes, different sets of classification criteria have been developed in recent decades.² In daily clinical practice, however, there is a significant overlap between classification and diagnosis.

The most recently developed classification criteria are the American College of Rheumatology – European League Against Rheumatism (ACR-EULAR) criteria.^{3,4} Within this criteria set, three points are given for a positive biopsy (i.e., a focus score ≥ 1), as well as for the presence of anti-SSA/Ro antibodies.^{3,4} Furthermore, one point is given for a decreased unstimulated whole salivary flow (UWSF ≤ 0.1 mL/min), a decreased Schirmer's test (≤ 5 mm/5 min in at least one eye), and an increased ocular staining score (OSS; i.e., ≥ 5 , in at least one eye). Patients with a total score of ≥ 4 points are classified as having pSS.^{3,4} Thus, for both the lacrimal and salivary glands, measurements to examine glandular function are included. Histopathologic evaluations, on the other hand, only concern biopsies from salivary glands, rather than biopsies from lacrimal glands, which might underestimate the inflammatory process.⁵ However, only for the lacrimal glands, a measurement to visualize glandular involvement (i.e., OSS) is included, and one could argue that salivary gland involvement might be underestimated without an imaging modality to measure it.

A promising imaging modality that has received a lot of attention in the last couple of years, when it comes to investigating major salivary gland involvement in pSS, is salivary gland ultrasonography (SGUS).⁶ Besides that SGUS is inexpensive and easily accessible, it is also non-invasive, non-irradiating and herewith patient-friendly.^{7,8} Previously, a meta-analysis identified 33 different scoring systems, with a large heterogeneity in which characteristics were evaluated, the type of glands that were examined and the different cut-off points that were applied.⁷ Initially, the Hocevar⁹ scoring system was implemented in the diagnostic work-up of patients with clinical suspicion of pSS at the University Medical Center Groningen (UMCG). By applying this scoring system, both parotid and submandibular glands are investigated and the following components are scored: (i) parenchymal echogenicity compared with the thyroid gland (graded 0-1); (ii) homogeneity (graded 0-3); (iii) presence of hypoechoic areas (graded 0-3); (iv) hyperechoic reflections

(graded 0-3 in parotid glands and 0-1 in submandibular glands); and (v) clearness of salivary gland border (graded 0-3).⁹ This in order to address all relevant aspects that can be measured by SGUS. At that time, the Hocevar scoring system was the most comprehensive and valid ultrasound scoring system available, and had previously been applied in various studies.⁹⁻¹²

Outcome Measures in Rheumatology Clinical Trials (OMERACT) filter

According to the OMERACT filter, there are different aspects, i.e., discrimination, truth and feasibility, that are considered important, when it comes to the applicability of a new diagnostic method.¹³

Discrimination, which captures issues of sensitivity to change and reliability, evaluates whether a measuring instrument, in this case SGUS, differentiates between situations of interest.¹⁴ Sensitivity to change is important when an instrument is considered to be (part of) an outcome measurement in clinical trials, and will be discussed in more detail at the end of the discussion. Reliability measures the degree to which an instrument is free of measurement error.¹⁴ In our center, the inter- and intra-observer reliability has been studied between three well-trained observers (AJS, KD, JvN) and Delli et al.¹⁰ concluded that SGUS, when applying the Hocevar⁹ scoring system, is a reliable imaging technique for patients clinically suspected with pSS. Furthermore, interobserver reliability was good to excellent for dividing pSS from non-SS sicca patients, but there were some discrepancies between the observers regarding scoring the severity of the SGUS findings.¹⁰ This is in line with the results of inter- and intra-observer reliability analyses performed by international SGUS experts.^{15,16}

Truth, which captures issues of face, content, construct, and criterion validity, questions whether an instrument measures what it is intended to measure.¹⁷ The aspect of criterion validity, i.e., the degree to which scores of the instrument are an adequate reflection of a gold standard, was evaluated in **chapter 2A**. Previously, no consensus was yet reached about the optimal cut-off point that should be used, and therefore receiver operating characteristics (ROC) analyses were performed to define the optimal cut-off point based upon our study population. When ACR-EULAR classification was applied as golden standard, optimal cut-off point of the Hocevar scoring system was defined as a total score of ≥ 15 (on a 0-48 scale).^{6,9} Overall, SGUS is considered a valid measuring instrument.

Feasibility, which captures an essential element in the selection of an instrument, and may be decisive in determining the success of an instrument, questions whether an instrument is easily accessible, given constraints in money, time, and interpretability.¹⁴ Since the comprehensive and valid Hocevar⁹ scoring system is re-

latively time-consuming, and time is precious, we investigated in **chapter 3** whether SGUS scoring of both parotid and submandibular glands, and all individual components of the Hocevar scoring system, is really needed for classifying pSS patients. In this chapter, we provided evidence that ultrasound scoring of hypoechogenic areas in one parotid and one submandibular gland is enough to predict whether or not a patient will fulfill the ACR-EULAR classification criteria for pSS. Herewith, SGUS is indeed considered to be a feasible instrument. However, we do recommend to compile a global view of all four glands (parotid and submandibular), but to score only one side (left or right) for classification. Although the role of SGUS for early detection of mucosa-associated lymphoid tissue (MALT) lymphoma has not yet been investigated, a global view of all four glands might rise suspicion of MALT lymphoma at an earlier stage, for instance in case of unilateral glandular swelling.

Overall, for diagnostic purposes, SGUS can be considered to be a reliable, valid and feasible imaging instrument, according to the OMERACT filter.

Incorporation of salivary gland ultrasonography into classification criteria

Previously, several studies have been performed in order to realize implementation of SGUS into either the 2002 AECG, or the 2012 ACR classification criteria.¹⁸⁻²³ Both, addition as well as replacement, had been investigated. However, before incorporation of SGUS into these criteria sets could have been realized, a new set of classification criteria, i.e., the ACR-EULAR criteria, already had been developed. Regrettably, the potential of SGUS was not fully known at the time the ACR-EULAR criteria were developed, and therefore SGUS was not considered as a possible criteria item.^{24,25} In **chapter 5**, we investigated the incorporation of SGUS into the ACR-EULAR classification criteria. Likewise, addition of SGUS as well as replacement of current criteria items, were investigated.²⁴ In our cohort, incorporation of SGUS did not lead to a substantial improvement (or decline) in the accuracy of the modified ACR-EULAR criteria, when SGUS was included as additional item, or when it replaced UWS, Schirmer's test, or OSS. However, the diagnostic accuracy of the modified ACR-EULAR classification criteria decreased substantially when SGUS replaced either the labial gland or the parotid gland biopsy, or when it replaced anti-SSA/Ro antibody status. Nevertheless, incorporation of SGUS leads to an increase in feasibility, since clinicians have a larger array of tests at their disposal. Thus far, three other studies have investigated the incorporation of SGUS into the ACR-EULAR classification criteria.²⁵⁻²⁷ There are, however, quite some methodological differences between the four studies, which complicates a direct comparison. Nevertheless, given the attention that incorporation of SGUS into the criteria is receiving, one could argue that this is eagerly awaited. Unfortunately, there are still a few hurdles to take,

before worldwide applicability of ultrasonography for classification purposes will be the case.

One of the problems to overcome is lack of international consensus regarding which SGUS scoring system to use and the optimal cut-off point that should be applied. Very recently, a sub-taskforce of the OMERACT working group has reached an international expert consensus on the definitions of normal and abnormal ultrasonographic features, resulting in the development of a new semi-quantitative scoring system for the assessment of major salivary glands.¹⁶ Fibrosis and fatty replacement were not included as potential items to score, but it was proposed to consider both items when the semi-quantitative system is not applicable. Thus, the first steps towards a novel semi-quantitative SGUS scoring system have been taken.¹⁶ Eventually, this scoring system should ideally be validated in different study populations from around the world, in order to confirm its potential as a reliable, valid and feasible scoring system, that should also be sensitive to change.

Measurements of salivary gland involvement, and their possible place in the diagnostic work-up

Nowadays, a salivary gland biopsy is one of the cornerstones in the classification of pSS patients, and such a biopsy can be taken from the parotid or labial salivary glands.²⁸ Performing a biopsy is, however, an invasive procedure, with a (small) risk of complications, like sensible nerve injury for labial gland biopsies, and temporary change in sensation of the incisional skin area for parotid gland biopsies.²⁸ Currently, a salivary gland biopsy is considered positive when there are ≥ 50 lymphocytes per 4 mm^2 of glandular tissue, which is considered a focus score (FS) of ≥ 1 .²⁹⁻³¹ However, when applying this definition to define whether a salivary gland biopsy is positive or negative, at least 18%-40% of pSS patients appear to have a negative labial gland biopsy.³²⁻³⁴ Furthermore, within the general population, 5% of the people have a positive parotid gland biopsy, whereas even 6%-15% of the people have a positive labial gland biopsy.^{32,35-37} Interestingly, it was shown previously that re-analyzing labial gland biopsies by another expert led to revision of the initial diagnosis in no less than 53% of the patients.³⁸ Therefore, for a more accurate assessment of the amount of lymphocytic infiltrate, the relative amount of CD45⁺ infiltrate can be measured, since the focus score is solely based upon the number, but not the size of the foci.^{29,39,40} Other histopathological features that could be present in biopsies of pSS patients include presence of lymphoepithelial lesions (LELs), a relative increase of IgG and IgM expressing plasma cells, compared to the number of IgA expressing plasma cells, and presence of germinal centers (GCs).⁴⁰⁻⁴⁷ Consensus concerning the best way to analyze and score salivary gland biopsies is, however, still work in progress.

In **chapter 2A**, parotid gland as well as labial gland biopsy outcome (i.e., negative or positive based upon FS) were compared with SGUS outcome (i.e., negative or positive based upon total Hoyer score). Upon a closer look at the patients with a negative SGUS, it was shown that the majority of these patients also had a negative parotid gland biopsy, whereas in about one quarter of these SGUS negative patients, the labial gland biopsy was positive. On the other hand, when solely looking at the patients with a positive SGUS, one fifth of the patients had a negative parotid gland biopsy, whereas most SGUS positive patients also had a positive labial gland biopsy. Evidently, although biopsies and SGUS outcome agree with each other to a large extent, there are discrepancies as well. Some possible suggestions for these discrepancies have been addressed in **chapter 2B** and **chapter 4**. First of all, salivary gland biopsies can be taken from the parotid and labial salivary glands, while with SGUS examination only the parotid glands, but not the labial glands, are visualized. Furthermore, it could be speculated that, overall, there might be a chronological order regarding the involvement of the various salivary glands in pSS, with the labial glands being affected earlier in comparison with the parotid glands. For some cases, this could explain why the SGUS outcome is negative, while the labial gland biopsy is (already) positive. Moreover, with ultrasonography, the entire (parotid) gland is visualized, whereas a (parotid gland) biopsy only consists of a small amount of tissue, which might not be representative for the remaining part of the gland. The possibility of sampling error could, to some extent, be reduced by multiple sectioning.²⁹

In **chapter 4**, a detailed comparison was made between parotid gland ultrasonography, parotid gland biopsy and parotid gland saliva production. Based on the analyses of the three measurements solely investigating parotid gland involvement, we concluded that these measurements are complementary to each other, and are all important for the evaluation of salivary gland involvement in pSS, as they assess different or at best partly related constructs. However, an invasive salivary gland biopsy does not necessarily have to be taken from every single patient with clinical suspicion of pSS. We have shown that the combination of a positive SGUS examination and presence of anti-SSA/Ro antibodies is highly predictive of pSS classification (**chapter 2A**).⁶ Although a full work-up is preferred for a good overall view, based upon our work, the following strategy can be recommended for outpatient rheumatologic clinics regarding the diagnostic work-up of patients with clinical suspicion of pSS. This work-up starts with an SGUS examination combined with the measurement of anti-SSA/Ro antibody status. The next step could be to assess the functional capacity of the exocrine glands, through UWS and a Schirmer's test. Hereafter, the degree of involvement of the target organs can be

obtained through the OSS and a salivary gland biopsy. Of note, in patients with negative serology, i.e., absence of auto-antibodies, which occurs in about 10%-50% of patients,⁴⁸⁻⁵⁰ a salivary gland biopsy is mandatory for classification. Training in major salivary gland ultrasonography is therefore recommended, in order for SGUS to be a reliable tool in the outpatient rheumatologic clinics as well. Furthermore, Hammenfors et al.⁵¹ recently recommended to perform regular calibration exercises between different ultrasonographers, as observer differences may not be detected easily in daily clinical practice.

The unsolved mystery of hypoechogenic areas

It was presupposed that presence of hypoechoic rounded multiple parenchymal lesions seen by ultrasonography, is a pathognomonic feature of pSS.^{52,53} In **chapter 3**, we underline this hypothesis, especially since homogeneity and hypoechogenic areas, sometimes used intertwined in ultrasound scoring systems other than the Hocevar⁹ scoring system, appear to correlate excellent with each other.⁵⁴ Thus far, however, it remains unclear what these hypoechogenic areas represent in terms of histopathological parameters, although various suggestions have been put forward.

Previously, it has been suggested that hypoechogenic areas represent enlarged glandular lobules that have been replaced by lymphocytic infiltration.^{53,55} To gain more insights regarding this topic, we directly compared parotid gland ultrasonography with histology of the same gland (**chapter 4**). We showed that there is a moderate association between presence of hypoechogenic areas in the parotid gland and relative amount of CD45⁺ infiltrate, and FS in parotid gland biopsies. This association might support the theory that hypoechogenic areas relate to areas with increased numbers of inflammatory cells. However, with great certainty, hypoechogenic areas are not directly represented by foci, since these lymphocytic infiltrates, which are located around striated ducts, are quite smaller in size than the size of the hypoechogenic areas. Therefore, it is more likely that hypoechogenic areas are more indirectly associated with these foci. Alternatively, these hypoechogenic areas could represent areas in the salivary gland that are composed of a number of foci, closely located to each other.

Wang et al.⁵⁶ hypothesized that presence of hypoechogenic areas could be caused by leakage of saliva into the periductal infiltrate and eventually into the salivary gland parenchyma, due to dysfunction of tight junctions between striated ductal cells. Pro-inflammatory cytokines are able to reduce the function of tight junctions.⁵⁷ In salivary glands of pSS patients, these cytokines can be secreted by activated lymphocytes that are located with the periductal infiltrates, or even within the ductal epithelium. Since epithelial cells can also produce cytokines,⁵⁸

cytokines secreted by ductal epithelial cells may also be responsible for the dysfunction of tight junctions. This is an interesting suggestion that warrants further investigation. However, it should be mentioned that histopathological evaluation of parotid gland biopsies does not reveal large amounts of extracellular liquid in the periductal areas.

Taken together, no clear-cut answer can be given to the burning question about which histopathologic structures are present in areas with ultrasonographic abnormalities, which, in the future, might be helpful to realize a more personalized medicine approach. A technique that might help unravel this question (a little further), is ultrasound-guided core needle biopsy (CNB), a technique very recently proven to be useful in the evaluation of suspected salivary gland lymphoma in pSS.⁵⁹ Although less tissue is collected in comparison with an open biopsy, CNB provides biopsies with a preserved architecture.⁵⁹ Herewith, a biopsy can be taken specifically from the location of a hypoechoic area, rather than from superficial parotid gland tissue located below the earlobe.²⁸ Another possible option to unravel this question could be performing a post-mortem study. However, it still remains to be proven whether either CNB or post-mortem studies will provide the answers that are so eagerly awaited.

Stratification of patients with primary Sjögren's syndrome

Primary Sjögren's syndrome is a very heterogeneous disease, which raises the possibility that it might be possible to define subgroups of patients within the broad spectrum of the disease. Thus far, it remains challenging to predict which patients are at risk of a more severe disease outcome. Prospective cohort studies are becoming increasingly important in the search to unravel these issues.⁶⁰ Within the University Medical Center Groningen (UMCG), there is a unique collaboration between different departments. Here, the REgistry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) cohort has been set up (**chapter 6**). This observational cohort combines up-to-date quality of care with gathering long-term prospective follow-up data in a large cohort of patients, in order to determine and predict the longitudinal course of the disease.

Recently, Tarn et al.⁶¹ were able to determine four different subgroups of pSS patients, a so-called symptom based stratification, using the patient-reported outcomes dryness, fatigue, pain and anxiety and depression. The following subgroups were identified, i.e., low symptom burden (LSB), high symptom burden (HSB), dryness dominant with fatigue (DDF), and pain dominant with fatigue (PDF).⁶¹ Significant differences were found in various serological parameters like anti-SSA/Ro and anti-SSB/La positivity, and serum IgG levels, and results were similar in the

independent validation cohorts.⁶¹ Unfortunately, ultrasonography was not included in their analysis.

In **chapter 6**, we investigated SGUS abnormalities in relation to clinical phenotype and patient characteristics, disease activity and disease damage in pSS patients included in the RESULT cohort. We showed that SGUS positive patients, i.e., patients with a total Hocevar score of ≥ 15 , had a distinct clinical phenotype compared with SGUS negative patients. This holds for all aspects of the disease, i.e., clinical, functional, serological and patient-reported outcome measurements (PROMs). Regrettably, the Hospital Anxiety and Depression Scale (HADS) questionnaire was not part of the initial set of questionnaires that were implemented in RESULT. Therefore, we were thus far unable to truly verify whether the applied SGUS score, i.e., the total Hocevar score,⁹ is also able to differentiate between the four different subgroups defined by Tarn et al.⁶¹ In view of the current developments, an amendment has recently been accepted to include the HADS into the ongoing RESULT cohort. In our cohort, the spread in the three domains of the ESSPRI, i.e., dryness, fatigue and pain, was less pronounced compared with the spread in ESSPRI domains that was shown in the study of Tarn et al.,⁶¹ and therefore, we were not able to clearly define the same subsets of patients. Although it is promising that there are clear differences between SGUS negative and SGUS positive patients in our cohort, the potential to use ultrasonography for patient stratification warrants further investigations and is still at the beginning. The applicability of SGUS as a self-contained entity to stratify patients, for instance for clinical trials might, however, be a bit too optimistic. This, at least partly, because the different SGUS components are scored categorically, which could lead to a sub-optimal sensitivity to change, as differences in the amount of ultrasound abnormalities may be too subtle to lead to a measurable difference.

In the future, this might be overcome with the use of image segmentation analysis, which provides automatically scored ultrasound lesions based upon computer algorithms.⁶² This form of artificial intelligence (AI) is one of the pillars in the HarmonicSS project of joint European research, the Horizon 2020 project.⁶³ Within the HarmonicSS research project,⁶³ as well as within the innovative medicines initiative (IMI) project 'NECESSITY', an innovative collaboration funded by the European Union,⁶⁴ initiatives are being taken to improve patient stratification, and this also includes the use of ultrasonography.

Systemic therapy and its effect measured by using ultrasonography

As for any disease, the development of a curative treatment, available (and effective) for every single patient, is the ultimate goal. Unfortunately, thus far this appears to be unprocurable. Systemic therapy, able to achieve disease remission, and approved by the Food and Drug Administration (FDA) and/or European Medicines Agency (EMA), therefore is second-best. Thus far, however, treatment mainly relies upon symptom relief. Although initial open-label studies of rituximab, a humanized anti-CD20 monoclonal antibody, and abatacept, a fully human soluble fusion protein of cytotoxic T lymphocyte antigen coupled to the Fc tail of IgG, showed promising effects, four large randomized, double-blind, placebo-controlled trials investigating the effectiveness of rituximab (n=2) and abatacept (n=2) failed to meet their primary endpoint (**chapter 8**).⁶⁵⁻⁷⁶ Currently, however, other potential drugs for systemic therapy are under development.⁷⁷⁻⁷⁹

Interestingly, Fisher et al.⁸⁰ demonstrated a significant improvement in total ultrasound score (TUS) after rituximab treatment compared with placebo at weeks 16 and 48, in 52 of the 133 pSS patients originally included in the TRACTISS study, a randomized double blind placebo controlled Trial of Anti-B-Cell Therapy In patients with primary Sjögren's syndrome. This, despite the fact that the TRACTISS study failed to meet its primary endpoint, i.e., a reduction of at least 30% in either fatigue or oral dryness on a visual analogue scale (VAS).⁷⁵

Previously, Jousse-Joulin et al. already demonstrated, in a study that included 28 of the 122 pSS patients originally included in the Tolerance and Efficacy of Rituximab in Primary Sjögren's Syndrome (TEARS) trial, that the parotid gland echostructure score improved in a significantly larger number of rituximab-treated patients compared with placebo-treated patients.⁸¹ Noteworthy, this again occurred, despite the fact that the original trial failed to meet its primary endpoint, i.e., an improvement of at least 30-mm (on a 100-mm scale) on at least two of the four VAS scores (i.e., global disease, dryness, fatigue and pain).^{76,81}

Delli et al.³⁹ previously demonstrated that prior to rituximab treatment, clinical responders had higher numbers of CD20⁺ B cells/mm² of parenchyma of parotid gland tissue compared with non-responders. Furthermore, after 12 weeks of rituximab treatment, there was a significant decrease in amount of CD45⁺ infiltrate, while no change was observed in the placebo group.³⁹ The studies of Fisher et al.⁸⁰ and Jousse-Joulin et al.⁸¹ demonstrate that SGUS is able, to a certain extent, to measure the effects of rituximab that are observed at a histopathological level.

Whether abatacept-treated patients also show significant ultrasonographic improvement compared with placebo-treated patients is currently under investigation. However, given that, on a histopathological level, treatment with abatacept

solely resulted in a disappearance of germinal centers, but did not lead to a decrease in the relative amount of CD45⁺ infiltrate,⁸² we hypothesize that abatacept treatment does not lead to a significant decrease in ultrasonographic abnormalities.

Salivary gland ultrasonography as part of a new composite endpoint for clinical trials

In order to demonstrate clinical efficacy of systemic treatment in pSS, it is crucial that a primary endpoint is reliable, valid, feasible, sensitive to change, and able to discriminate between active treatment and placebo. Regarding the trials investigating the efficacy of rituximab and abatacept, the choice of the primary endpoint certainly played a role in the possibility of achieving it in the first place.^{65,67} The minimal clinically important improvement (MCII) for the ESSPRI score was set at a decrease of only 1 point or 15% from baseline.⁸³ Within this light, expecting a decrease of at least 30% on the VAS score of comparable patient-reported outcomes, appears to be a bit optimistic. Therefore one should not be surprised that trials that expect these (unrealistically) high levels of efficacy, come out negative. A negative trial, however, does not necessarily mean that the treatment fails for every single patient. Furthermore, the ESSDAI was relatively new when it was chosen as primary endpoint of various RCTs,^{67,68,84} although previous studies showed that the ESSDAI was a reliable tool to measure systemic disease activity and was also sensitive to change.^{85,86} However, after a few years of ESSDAI experience, not all domains of the ESSDAI appear to be as reliable or sensitive to change as was initially expected. This gives the impression that the search for the most suitable endpoint in pSS is still work in progress. The development of new valid and sensitive composite endpoints, which is another aim of the IMI project 'NECESSITY',⁶⁴ hopefully help to realize the personalized medicine approach that is so eagerly warranted in pSS.

It would be appropriate to also consider ultrasonography as a possible item when new composite endpoints are being developed. Considering that pSS is a very heterogeneous disease, a composite endpoints that consists of items covering different disease aspects, might be better to show clinical efficacy.⁸⁷ Very recently, the UMCG launched a promising new composite endpoint. This composite endpoint is called CRESS, which is an acronym for Composite of Relevant Endpoints in Sjögren's Syndrome. The CRESS consists of the following components: systemic disease activity measured by clinESSDAI,⁸⁸ patient-reported symptoms measured by ESSPRI⁸⁹ tear gland involvement measured by OSS or Schirmer's test, salivary gland involvement measured by UWS or SGUS, and serological activity measured by IgG or rheumatoid factor.⁸⁷ Total CRESS response was defined as response on at least 3 out of the 5 items.⁸⁷

Interestingly, when the ASAPIII study data are re-analyzed based upon the CRESS, it was shown that abatacept-treated patients showed a significant difference in CRESS response compared with placebo-treated patients, i.e., 65% versus 18%, respectively.⁸⁷ These preliminary results are promising, and the CRESS appears to be valid and sensitive to change, as it seems able to differentiate between active treatment and placebo. Within NECESSITY, further validation analyses will take place in other recently performed RCTs. Currently, the SGUS component in the CRESS is represented by the total Hocevar⁸⁹ score, since it is not yet clear which ultrasound items are most important to evaluate treatment response.⁹⁰ Feasibility of this item, and herewith applicability in daily clinical practice warrants further improvement. Ideally, (parts of) the recently developed semi-quantitative scoring system¹⁶ should be implemented in the CRESS, in order to increase its use all around the globe.

Clinical implications and future perspectives

The potential value of SGUS has been elucidated even further in recent years. This is, however, still an ongoing process. At this point, we know that overall, SGUS is a valid and reliable imaging modality. Its feasibility depends on the purpose for which it is being applied, i.e., for classification purposes feasibility is good, as it suffices to apply a simple scoring system. Considering other potential purposes, like stratification, or to monitor treatment efficacy, still more comprehensive scoring systems are required.

The development of a universal scoring system is a prerequisite for ultrasonography to reach its true potential. Important steps have already been taken by a sub-taskforce of the OMERACT working group to develop a semi-quantitative scoring system. Ultimately, this should be implemented in daily rheumatology practice. It would be a major step forward if every center applied the same method. In the future, there might also be an important role for artificial intelligence in this.

SGUS has proven its value for diagnostic purposes. It would therefore be more than justified if SGUS, when there actually is a universal scoring system, becomes part of the classification criteria for pSS.

Concerning the development of new composite endpoints for the measurement of treatment response in pSS patients, important steps have recently been taken with the development of the CRESS. Regarding the fact that SGUS is being part of the CRESS, again ideally the Hocevar scoring system should be replaced by a universally accepted scoring system. This would also make a fair comparison between various studies easier and more reliable. Hopefully, ultrasound will eventually get the role it deserves.

References

1. Mariette X, Criswell LA. Primary Sjögren's syndrome. *N Engl J Med* 2018;378:931-9.
2. Bootsma H, Spijkervet FKL, Kroese FGM, *et al.* Toward new classification criteria for Sjögren's syndrome? *Arthritis Rheum* 2013;65:21-3.
3. 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 patients cohorts. *Ann Rheum Dis* 2017;76:9-16.
4. 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 patients cohorts. *Arthritis Rheumatol* 2017;69:35-45.
5. Luemsamran P, Rootman J, White VA, *et al.* The role of biopsy in lacrimal gland inflammation: a clinicopathologic study. *Orbit (London)* 2017;36:411-8.
6. Mossel E, Delli K, Van Nimwegen JF, *et al.* Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1883-9.
7. Delli K, Dijkstra PU, Stel AJ, *et al.* Diagnostic properties of ultrasound of major salivary glands in Sjögren's syndrome: a meta-analysis. *Oral Dis* 2015;21:792-800.
8. Jousse-Joulin S, Milic V, Jonsson M, *et al.* Is salivary gland ultrasonography a useful tool in Sjögren's syndrome? A systematic review. *Rheumatology (Oxford)* 2015;55:1-12.
9. Hocevar A, Ambrozic A, Rozman B, *et al.* Ultrasonographic changes of major salivary glands in primary Sjögren's syndrome. Diagnostic value of a novel scoring system. *Rheumatology (Oxford)* 2005;44:768-72.
10. Delli K, Arends S, Van Nimwegen JF, *et al.* Ultrasound of the major salivary glands is a reliable imaging technique in patients with clinically suspected primary Sjögren's syndrome. *Ultraschall Med* 2018;39:328-33.
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. 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.

13. Boers M, Brooks P, Strand CV, *et al.* The OMERACT filter for Outcome Measures in Rheumatology. *J Rheumatol* 1998;25:198-9.
14. Wells G, Beaton DE, Tugwell P, *et al.* Updating the OMERACT filter: discrimination and feasibility. *J Rheumatol* 2014;41:1005-10.
15. 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.
16. 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.
17. Tugwell P, Boers M, D'Agostino MA, *et al.* Updating the OMERACT filter: implications of filter 2.0 to select outcome instruments through assessment of 'truth': content, face, and construct validity. *J Rheumatol* 2014;41:1000-4.
18. 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.
19. Cornec D, Jousse-Joulin S, Pers J-O, *et al.* Contribution of salivary gland ultrasonography to the diagnosis of Sjögren's syndrome: Toward new diagnostic criteria? *Arthritis Rheum* 2013;65:216-25.
20. Takagi Y, Kimura Y, Nakamura H, *et al.* Salivary gland ultrasonography: can it be an alternative to sialography as an imaging modality for Sjögren's syndrome? *Ann Rheum Dis* 2010;69:1321-4.
21. 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. *Rheumatol (Oxford)* 2012;51:1081-5.
22. Jonsson M, 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.
23. Yonetsu K, Takagi Y, Sumi M, *et al.* Sonography as a replacement for sialography for the diagnosis of salivary glands affected by Sjögren's syndrome. *Ann Rheum Dis* 2002;61:276-7.
24. 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* 2020;72:583-90.

25. Jousse-Joulin S, Gatineau F, Baldini C, *et al.* Weight of salivary gland ultrasonography compared to other items of the 2016 ACR/EULAR classification criteria for primary Sjögren's syndrome. *J Intern Med* 2020;287:180-8.
26. 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.
27. 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.
28. Delli K, Vissink A, Spijkervet FKL. Salivary gland biopsy for Sjögren's syndrome. *Oral Maxillofac Surg Clin North Am* 2014;26:22-33.
29. Fisher BA, Jonsson R, Daniels T, *et al.* Standardisation of labial salivary gland histopathology in clinical trials in primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1161-8.
30. Chisholm DM, Mason DK. Labial salivary gland biopsy in Sjögren's disease. *J Clin Pathol* 1968;21:656-60.
31. Greenspan JS, Daniels TE, Talal N, *et al.* The histopathology of Sjögren's syndrome in labial salivary gland biopsies. *Oral Surgery, Oral Med Oral Pathol* 1974;37:217-29.
32. Pijpe J, Kalk WWI, Van der Wal JE, *et al.* Parotid gland biopsy compared with labial biopsy in the diagnosis of patients with primary Sjögren's syndrome. *Rheumatology (Oxford)* 2007;46:335-41.
33. Vitali C, Moutsopoulos HM, Bombardieri S. The European Community Study Group on diagnostic criteria for Sjögren's syndrome. Sensitivity and specificity of tests for ocular and oral involvement in Sjögren's syndrome. *Ann Rheum Dis* 1994;53:637-47.
34. Vitali C, Tavoni A, Simi U, *et al.* Parotid sialography and minor salivary gland biopsy in the diagnosis of Sjögren's syndrome. A comparative of 84 patients. *J Rheumatol* 1988;15:262-7.
35. Daniels TE. Labial salivary gland biopsy in Sjögren's syndrome assessment as a diagnostic criterion in 362 suspected cases. *Arthritis Rheum* 1984;27:147-56.
36. Segerberg-Konttinen M, Konttinen YT, Bergroth V. Focus score in the diagnosis of Sjögren's syndrome. *Scand J Rheumatol Suppl* 1986;61:47-51.
37. Radfar L, Kleiner DE, Fox PC, *et al.* Prevalence and clinical significance of lymphocytic foci in minor salivary glands of healthy volunteers. *Arthritis Rheum* 2002;47:520-4.

38. Vivino FB, Gala I, Hermann GA. Change in final diagnosis on second evaluation of labial minor salivary gland biopsies. *J Rheumatol* 2002;29:938-44.
39. Delli K, Haacke EA, Kroese FGM, *et al.* Towards personalised treatment in primary Sjögren's syndrome: baseline parotid histopathology predicts responsiveness to rituximab treatment. *Ann Rheum Dis* 2016;75:1933-8.
40. Kroese FGM, Haacke EA, Bombardieri M. The role of salivary gland histopathology in primary Sjögren's syndrome: promises and pitfalls. *Clin Exp Rheumatol* 2018;36:222-33.
41. Risselada AP, Looije MF, Kruize AA, *et al.* The role of ectopic germinal centers in the immunopathology of primary Sjögren's syndrome: a systematic review. *Semin Arthritis Rheum* 2013;42:368-76.
42. Ihrler S, Zietz C, Sendelhofert A, *et al.* Lymphoepithelial duct lesions in Sjögren-type sialadenitis. *Virchows Arch* 1999;434:315-23.
43. Ihrler S, Adam P, Guntinas-Lichius O, *et al.* Pattern recognition in the differential diagnosis of salivary gland lymphoepithelial lesions. *Pathologe* 2009;30:432-41.
44. Van Ginkel MS, Haacke EA, Bootsma H, *et al.* Presence of intraepithelial B-lymphocytes is associated with the formation of lymphoepithelial lesions in salivary glands of primary Sjögren's syndrome patients. *Clin Exp Rheumatol* 2019;37Suppl118:42-8.
45. Zandbelt M, Wentink J, De Wilde P, *et al.* The synergistic value of focus score and IgA% score of sublabial salivary gland biopsy for the accuracy of the diagnosis of Sjögren's syndrome: a 10-year comparison. *Rheumatology (Oxford)* 2002;41:819-23.
46. De Wilde P, Vooy's G, Baak J, *et al.* Quantitative immunohistologic and histomorphometric diagnostic criteria for Sjögren's syndrome. *Pathol Res Pr* 1989;185:778-80.
47. Bodeutsch C, De Wilde P, Kater L, *et al.* Quantitative immunohistologic criteria are superior to the lymphocytic focus score criterion for the diagnosis of Sjögren's syndrome. *Arthritis Rheum* 1992;35:1075-87.
48. Peri Y, Agmon-Levin N, Theodor E, *et al.* Sjögren's syndrome, the old and the new. *Best Pract Res Clin Rheumatol* 2012;26:105-17.
49. Ramos-Casals M, Solans R, Rosas J, *et al.* Primary Sjögren's syndrome in Spain. *Medicine (Baltimore)* 2008;87:210-9.
50. Oni C, Mitchell S, James K, *et al.* Eligibility for clinical trials in primary Sjögren's syndrome: lessons from the UK Primary Sjögren's Syndrome Registry. *Rheumatology (Oxford)* 2016;55:544-52.

51. Hammenfors DS, Causevic H, Assmus J, *et al.* Assessment of major salivary gland ultrasonography in Sjögren's syndrome. A comparison between bedside and post-examination evaluations. *Clin Exp Rheumatol* 2019;37Sup-
pl118:153-8.
52. Theander E, Mandl T. Primary Sjögren's syndrome: diagnostic and prognostic value of salivary gland ultrasonography using a simplified scoring system. *Arthritis Care Res* 2014;66:1102-7.
53. Wernicke D, Hess H, Gromnica-Ihle E, *et al.* Ultrasonography of salivary glands – A highly specific imaging procedure for diagnosis of Sjögren's syndrome. *J Rheumatol* 2008;35:285-93.
54. 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.
55. Takashima S, Morimoto S, Tomiyama N, *et al.* Sjögren syndrome: comparison of sialography and ultrasonography. *J Clin Ultrasound* 1992;20:99-109.
56. Wang X, Bootsma H, Terpstra J, *et al.* Progenitor cell niche senescence reflects pathology of the parotid salivary gland in primary Sjögren's syndrome. *Rheumatology (Oxford)* 2020;59:3003-13.
57. Chai Q, He WQ, Zhou M, *et al.* Enhancement of blood-brain barrier permeability and reduction of tight junction protein expression are modulated by chemokines/cytokines induced by rabies virus infection. *J Virol* 2014;88:4698-710.
58. Koch S, Nusrat A. The life and death of epithelia during inflammation: lessons learned from the gut. *Annu Rev Pathol Mech Dis* 2012;7:35-60.
59. Baer A, Grader-Beck T, Antiochos B, *et al.* Ultrasound-guided biopsy of suspected salivary gland lymphoma in Sjögren's syndrome. *Arthritis Care Res* 2020; published online first: 5 April 2020.
60. Rischmueller M, Tieu J, Lester S. Primary Sjögren's syndrome. *Best Pract Res Clin Rheumatol* 2016;30:189-220.
61. Tarn JR, Howard-Tripp N, Lendrem DW, *et al.* Symptom-based stratification of patients with primary Sjögren's syndrome: multi-dimensional characterization of international observational cohorts and reanalyses of randomized clinical trials. *Lancet Rheumatol* 2019;1:e85-94.
62. Baldini C, Zabotti A, Filipovic N, *et al.* Review Imaging in primary Sjögren's syndrome : the "obsolete and the new". *Clin Exp Rheumatol* 2018;36Sup-
pl112:215-21.

63. Harmonization and integrative analysis of regional, national and international cohorts on primary Sjögren's syndrome (pSS) towards improved stratification, treatment and health policy making (HarmonicSS). *Horizon2020 project*. Grant agreement number 731944.
64. New clinical endpoints in primary Sjögren's syndrome: an international trial based on stratifying patients (NECESSITY). *Innovative Medicines Initiative*. Grant agreement number 806975.
65. Verstappen GM, Van Nimwegen JF, Vissink A, *et al*. The value of rituximab treatment in primary Sjögren's syndrome. *Clin Immunol* 2017;182:62-71.
66. Meiners PM, Vissink A, Kroese FGM, *et al*. Abatacept treatment reduces disease activity in early primary Sjögren's syndrome (open-label proof of concept ASAP study). *Ann Rheum Dis* 2014;73:1393-6.
67. Van Nimwegen JF, Mossel E, van Zuiden GS, *et al*. Abatacept treatment for patients with early active primary Sjögren's syndrome: a single-centre, randomized, double-blind, placebo-controlled, phase 3 trial (ASAP-III study). *Lancet Rheumatol* 2020;9913:1-11.
68. Baer A, Gottenberg J-E, St. Clair EW, *et al*. Efficacy and safety of abatacept in active primary Sjögren's syndrome: results of a phase III, randomized, placebo-controlled trial. *Ann Rheum Dis* 2020; published online first: 9 November 2020.
69. Adler S, Körner M, Förger F, *et al*. Evaluation of histologic, serologic, and clinical changes in response to abatacept treatment of primary sjögren's syndrome: a pilot study. *Arthritis Care Res* 2013;65:1862-8.
70. Machado AC, Dos Santos LC, Fidelix T, *et al*. Effectiveness and safety of abatacept for the treatment of patients with primary Sjögren's syndrome. *Clin Rheumatol* 2020;39:243-8.
71. Meijer JM, Meiners PM, Vissink A, *et al*. Effectiveness of rituximab treatment in primary Sjögren's syndrome: a randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2010;62:960-8.
72. Meiners PM, Arends S, Brouwer E, *et al*. Responsiveness of disease activity indices ESSPRI and ESSDAI in patients with primary Sjögren's syndrome treated with rituximab. *Ann Rheum Dis* 2012;71:1297-302.
73. Dass S, Bowman SJ, Vital EM, *et al*. Reduction of fatigue in Sjögren's syndrome with rituximab: Results of a randomised, double-blind, placebo-controlled pilot study. *Ann Rheum Dis* 2008;67:1541-4.
74. Carubbi F, Cipriani P, Marrelli A, *et al*. Efficacy and safety of rituximab treatment in early primary Sjögren's syndrome: A prospective, multi-center, follow-up study. *Arthritis Res Ther* 2013;15:R172.

75. Bowman SJ, Everett CC, O'Dwyer JL, *et al.* Randomized controlled trial of rituximab and cost-effectiveness analysis in treating fatigue and oral dryness in primary Sjögren's syndrome. *Arthritis Rheumatol* 2017;69:1440-50.
76. Devauchelle-Pensec V, Mariette X, Jousse-Joulin S, *et al.* Treatment of primary Sjögren's syndrome with rituximab – a randomized trial. *Ann Intern Med* 2014;160:233-42.
77. Fisher BA, Szanto A, Ng WF, *et al.* Assessment of anti-CD40 antibody iscalimab in patients with primary Sjögren's syndrome: a multicenter, randomised, double-blind, placebo-controlled, proof of concept study. *Lancet Rheumatol* 2020;2:e142-52.
78. Dörner T, Posch MG, Li Y, *et al.* Treatment of primary Sjögren's syndrome with ianalumab (VAY736) targeting B cells by BAFF receptor blockade coupled with enhanced, antibody-dependent cellular cytotoxicity. *Ann Rheum Dis* 2019;78:641-7.
79. Radstake TRDJ, Van der Heijden EHM, Moret FM, *et al.* Clinical efficacy of leflunomide/hydroxychloroquine combination therapy in patients with primary Sjögren's syndrome: results of a placebo-controlled double-blind randomized clinical trial [abstract]. *Arthritis Rheumatol* 2018;70:suppl110.
80. Fisher BA, Everett CC, Rout J, *et al.* Effect of rituximab on a salivary gland ultrasound score in primary Sjögren's syndrome: results of the TRACTISS randomized double-blind multicentre substudy. *Ann Rheum Dis* 2018;77:412-6.
81. Jousse-Joulin S, Devauchelle-Pensec V, Cornec D, *et al.* Brief report: Ultrasonographic assessment of salivary gland response to rituximab in primary Sjögren's syndrome. *Arthritis Rheumatol* 2015;67:1623-8.
82. Haacke EA, van der Vegt B, Meiners PM, *et al.* Abatacept treatment of patients with primary Sjögren's syndrome results in a decrease of germinal centres in salivary gland tissue. *Clin Exp Rheumatol* 2017;35:317-20.
83. Seror R, Bootsma H, Saraux A, *et al.* Defining disease activity states and clinically meaningful improvement in primary Sjögren's syndrome with EULAR primary Sjögren's syndrome disease activity (ESSDAI) and patient-reported indexes (ESSPRI). *Ann Rheum Dis* 2016;75:382-9.
84. Felten R, Meyer N, Duffaut P, *et al.* IL-6 receptor inhibition in primary Sjögren's syndrome: results from a randomized multicenter academic double blind placebo-controlled trial of tocilizumab in 110 patients [abstract]. *Arthritis Rheumatol* 2019;71:suppl10.
85. Moerman RV, Arends S, Meiners PM, *et al.* EULAR Sjögren's syndrome disease activity index (ESSDAI) is sensitive to show efficacy of rituximab treatment in a randomized controlled trial. *Ann Rheum Dis* 2014;73:472-4.

86. Seror R, Mariette X, Bowman S, *et al.* Accurate detection of changes in disease activity in primary Sjögren's syndrome by the European league against rheumatism Sjögren's syndrome disease activity index. *Arthritis Care Res* 2010;62:551-8.
87. Arends S, De Wolff L, Van Nimwegen JF, *et al.* Composite of relevant endpoints for Sjögren's syndrome (CRESS) [abstract]. *Arthritis Rheumatol* 2020;72:suppl10.
88. Seror R, Meiners P, Baron G, *et al.* Development of the ClinESSDAI: A clinical score without biological domain. A tool for biological studies. *Ann Rheum Dis* 2016;75:1945-50.
89. Seror R, Ravaud P, Mariette X, *et al.* EULAR Sjögren's syndrome patients reported index (ESSPRI): development of a consensus patient index for primary Sjögren's syndrome. *Ann Rheum Dis* 2011;70:968-72.
90. Arends S, De Wolff L, Van Nimwegen JF, *et al.* Development of the Composite of Relevant Endpoints in Sjögren's Syndrome (CRESS): a comprehensive tool for clinical trials. *Submitted manuscript*.

A decorative border surrounds the central text area, featuring a pattern of teal and white polka dots of varying sizes. The border is thicker on the left and right sides and thinner at the top and bottom.

APPENDICES

Summary

Samenvatting

Dankwoord

Curriculum vitae

List of publications

List of abbreviations

Summary

Primary Sjögren's syndrome (pSS) is a chronic, systemic autoimmune disease characterized by inflammation of the salivary and lacrimal glands, leading to xerostomia and keratoconjunctivitis sicca. When it comes to investigating major salivary gland involvement in patients (suspected) with pSS, salivary gland ultrasonography (SGUS) seems to be a promising imaging modality. Thus far, however, its true potential still appears to be underestimated. At the multidisciplinary Sjögren's expertise center within the University Medical Center Groningen (UMCG), however, SGUS evaluations were already performed as part of the diagnostic work-up of patients suspected with pSS. The SGUS scoring system of Hocevar and colleagues was implicated and investigates five domains i.e., parenchymal echogenicity, homogeneity, presence of hypoechogenic areas, hyperechogenic reflections and clearness of the salivary gland border in both parotid and submandibular salivary glands. The overall aim of the research described in this PhD thesis was to investigate whether SGUS is a valid and feasible tool, applicable in the diagnostic work-up of (suspected) pSS patients. Furthermore, incorporation in the ACR-EULAR criteria was investigated, as well as its potential to be used for stratification of previously classified pSS patients. Moreover, a detailed comparison between ultrasonography, histopathology and saliva production was performed, all collected from the same parotid gland. Finally, the efficacy and safety of treatment with abatacept was investigated.

In **chapter 2A**, we assessed the validity of SGUS in a representative population of 103 patients clinically suspected with pSS. SGUS was compared with parotid and labial gland biopsy outcome, sialometry, anti-SSA/Ro antibody status, and American European Consensus Group (AECG) criteria, American College of Rheumatology (ACR) criteria and ACR – European League Against Rheumatism (EULAR) classification criteria. We showed that, although the accuracy of SGUS to predict parotid gland and labial gland biopsy outcome was good, SGUS is not fully able to replace the invasive salivary gland biopsy. However, we did show that a combination of a positive SGUS and presence of anti-SSA/Ro antibodies is highly suggestive of pSS classification. In **chapter 2B**, we furthermore emphasized that a parotid gland biopsy should not be ignored and that the parotid gland is an important organ in the evaluation of pSS.

In **chapter 3**, we investigated whether ultrasonographic scoring of (i) both parotid and submandibular glands and (ii) all five components of the Hocevar scoring system, are needed for classifying patients as pSS. These analyses were performed in

a population of 204 patients clinically suspected with pSS. We showed that there is excellent correlation between the ultrasound scores of both parotid glands, as well as between the ultrasound scores of both submandibular glands, but that the associations between both glands located on either the right or the left side of the face (i.e., parotid vs. submandibular glands), were slightly lower. Compared with the original Hocevar score (four glands and five components), comparable accuracy of SGUS to predict fulfillment of the ACR-EULAR classification for pSS was achieved, when only one parotid and one submandibular gland were scored. Furthermore, only the two components parenchymal echogenicity and hypoechogenic areas significantly predicted ACR-EULAR classification. When also taking the reliability of the components into account, our results indicate that ultrasound scoring of hypoechogenic areas in one parotid and one submandibular gland is enough to predict whether or not a patient will fulfill the ACR-EULAR classification criteria for pSS.

In **chapter 4**, a detailed comparison was performed between three different measurements to assess salivary gland involvement in pSS, i.e., ultrasonography, histopathology and collection of saliva in the same parotid gland. These analyses regarding construct validity were performed in a group of 41 pSS patients, classified according to the ACR-EULAR classification criteria. We showed that there is a moderate association between hypoechogenic areas in the parotid gland and amount of CD45⁺ infiltrate. Furthermore, presence or absence of hypoechogenic areas does not predict very well whether a biopsy contains a plasma cell shift, lymphoepithelial lesions, or germinal centers. Moreover, we found poor associations between parotid gland ultrasonography and parotid flow, which implies that ultrasonographic aberrations are partly reflected by reduced salivary secretion. Besides, changes measured by different evaluation modalities may depend upon the current disease phase a patient is in. Overall, we concluded that all three measurements that assess parotid gland involvement provide other information for the evaluation of pSS, and are therefore complementary.

In **chapter 5**, we assessed whether the addition of SGUS or replacement of current criteria items by SGUS influences the performance of the ACR-EULAR criteria for pSS. We performed a large clinical cohort study, based on a study population that consisted of all eligible consecutive patients with complete data on all ACR-EULAR items, who underwent an SGUS examination at the multidisciplinary Sjögren's expertise center at the UMCG (n=243). An average SGUS score for presence of hypoechogenic areas of ≥ 1.5 was considered positive. Furthermore, performance of the ACR-EULAR

criteria was best when a positive SGUS was assigned a weight of one point, and when the optimal cut-off point to discriminate between pSS and non-SS patients remained ≥ 4 . We showed that the addition of SGUS to the ACR-EULAR criteria resulted in negligible changes in the performance of the criteria, but it does improve the feasibility of the criteria, as clinicians can choose from a wider array of tests. When SGUS replaced one of the current criteria items, we showed that SGUS can replace the ocular staining score, Schirmer's test, or unstimulated whole salivary flow, without major changes in the performance of the criteria. However, SGUS cannot completely replace a salivary gland biopsy or the measurement of anti-SSA/Ro antibodies, since this led to a considerable decrease in the performance of the criteria.

In **chapter 6**, we investigated SGUS abnormalities in relation to patient characteristics, disease activity and disease damage in 172 patients with pSS. All patients were included in the REgistry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) cohort, which is a prospective observational cohort and was set up in order to identify biomarkers and clinical parameters that determine and predict the longitudinal course of pSS.

We showed that SGUS positive patients, here defined as a total Hocevar score of ≥ 15 , had a distinct clinical phenotype compared with SGUS negative patients. These differences were found in all aspects of the disease, i.e., clinical, functional, serological and regarding patient-reported outcome measurements. Herewith, a full SGUS evaluation could be considered as a tool to stratify pSS patients. However, results were not completely similar when only hypoechoogenic areas were taken into account. This underlines that the applicability of SGUS for diagnostic purposes is essentially different from the use of SGUS to stratify patients. For the latter purpose, a more comprehensive scoring system may be preferred.

In **chapter 7**, we emphasized that the baseline histopathological characteristics of the parotid gland should be used to predict the effect of rituximab therapy. This could be of help for the physician to make treatment decisions based upon a personalized medicine approach.

In **chapter 8**, we investigated the efficacy and safety of abatacept treatment in the Abatacept Sjögren Active Patients phase III (ASAP-III) study, a randomized, double-blind, placebo-controlled trial. The study population consisted of 80 patients, who were classified as pSS patients according to the AECG criteria, were over 18 years of age, had a time from diagnosis of 7 years or less, and an European League

Against Rheumatism Sjögren's Syndrome Disease Activity Index (ESSDAI) score of at least 5. Patients were randomly allocated (1:1) for a treatment with subcutaneous administration of abatacept or placebo for 24 weeks. PSS patients tolerated abatacept well. Unfortunately, the primary outcome, i.e., the difference in ESSDAI score between the abatacept and placebo-treated patients at week 24, was not met. However, abatacept did show beneficial effects on the biological domain, i.e., a significant difference in IgG and rheumatoid factor level between abatacept and placebo-treated patients. At baseline and at week 24, SGUS evaluations were performed as well. Analyses on whether Abatacept treatment leads to significant reduction of ultrasonographic abnormalities are currently being performed. Future research on this highly heterogeneous disease could evaluate whether patients with specific clinical or biological manifestations might benefit from abatacept treatment.

In the review described in **chapter 9**, we discussed various topics about interesting developments within the last year regarding the assessment of salivary gland involvement in patients with clinical suspicion of, or diagnosed with pSS. Based upon the discussed literature, a few conclusions could be drawn. One of the items addressed was that detection of mild glandular swelling by palpation appears to be overestimated and that the potential role of SGUS to fill the gap warrant further investigation. Furthermore, the true potential of grey-scale histogram analysis and elastography should be evaluated in the target population. Moreover, important steps have recently been taken to develop a novel semi-quantitative scoring system, although the advice is to perform regular calibration exercises to overcome potential observer bias. Finally, we concluded that SGUS examination should be part of the diagnostic work-up of potential pSS patients and that incorporation of SGUS into the current classification criteria would rebalance them.

Samenvatting

Het primaire syndroom van Sjögren is een chronische, systemische auto-immuun aandoening, die wordt gekenmerkt door ontsteking van de speeksel- en traanklieren. Dit uit zich in droge ogen en droge mond.

Patiënten kunnen geclassificeerd worden aan de hand van verschillende criteria, waarvan de ACR-EULAR het meest recent zijn. De ACR-EULAR criteria bestaan uit verschillende onderdelen, namelijk een speekselklier bipt, anti-SSA/Ro antilichamen, twee testen om betrokkenheid van de ogen in kaart te brengen (ocular staining score; OSS en Schirmer's test) en één test om de betrokkenheid van de speekselklieren in kaart te brengen (verzameling van ongestimuleerd speeksel; UWS). Het is op dit moment niet mogelijk om te genezen van het syndroom van Sjögren.

Echografie van de grote speekselklieren lijkt een veelbelovende methode te zijn om de betrokkenheid van de speekselklieren te onderzoeken bij (potentiële) Sjögren patiënten. Binnen het multidisciplinaire Sjögren Expertise Centrum in het Universitair Medisch Centrum Groningen (UMCG) wordt echografie van de grote speekselklieren al enige tijd gebruikt als onderdeel van de diagnostiek die patiënten ondergaan als er verdenking is op het syndroom van Sjögren. Hiervoor wordt het scoresysteem van Hoyer gebruikt. Middels dit internationaal gebruikte scoresysteem worden vijf verschillende aspecten van de speekselklieren onderzocht: echogeniciteit, homogeniciteit, aanwezigheid van hypo-echogene gebieden, hyper-echogene reflecties en de mate waarop de randen van de speekselklier afgebakend zijn van het omliggende weefsel. Zowel beide parotiden (oor speekselklieren) als beide submandibulaire (kaak) speekselklieren worden middels speekselklier echografie geëvalueerd.

Het voornaamste doel van de onderzoeken beschreven in dit proefschrift betreft het onderzoeken of echografie van de grote speekselklieren een goede (valide) en toepasbare methode is om te worden gebruikt in de diagnostiek van (potentiële) Sjögren patiënten.

Verder is er onderzoek gedaan naar het toevoegen van speekselklier echografie aan de ACR-EULAR classificatie criteria voor het syndroom van Sjögren, evenals naar het gebruik van speekselklier echografie om reeds geclassificeerde Sjögren patiënten onder te verdelen in subgroepen (stratificeren).

Daarnaast is er een gedetailleerde analyse gedaan naar drie verschillende methoden, die elk op hun eigen wijze de (mate van) betrokkenheid van de parotiden in kaart kunnen brengen, namelijk een bipt van de speekselklier (histopathologie), het meten van de speeksel productie (sialometrie) en echografie.

Tot besluit is er onderzoek gedaan naar de effectiviteit en veiligheid van behandeling met abatacept.

In **hoofdstuk 2A** is de validiteit van speekselklier echografie onderzocht in een studiepopulatie van 103 patiënten, bij wie er een klinische verdenking was op het syndroom van Sjögren. Echografie van de grote speekselklieren is vergeleken met de uitkomsten van het parotisbipt en het lipbipt, sialometrie, de aan- of afwezigheid van anti-SSA/Ro antilichamen, en de AECG, ACR en ACR-EULAR classificatie criteria.

Hoewel speekselklier echografie over het algemeen de uitkomsten van parotis en lipbipten goed kan voorspellen, blijkt echografie van de grote speekselklieren niet in staat om de bijdrage van het bipt in zijn geheel te vervangen. De combinatie van een positieve echo van de grote speekselklieren en aanwezigheid van anti-SSA/Ro antilichamen voorspelt echter goed welke patiënten geclassificeerd worden als Sjögren patiënten.

In **hoofdstuk 2B** is het belang van het parotisbipt benadrukt en is onderstreept dat het een belangrijk orgaan is in de evaluatie van het syndroom van Sjögren.

In **hoofdstuk 3** is onderzocht of het voor het classificeren van Sjögren patiënten noodzakelijk is om alle vier de grote speekselklieren (2 parotiden en 2 submandibulaire speekselklieren) en alle vijf aspecten van het Hocevar score systeem te onderzoeken. Hiervoor is een studiepopulatie van 204 patiënten gebruikt, bij wie er een klinische verdenking was op het syndroom van Sjögren. De statistische samenhang (correlatie) tussen echografie uitkomsten van de beide parotiden is uitstekend en datzelfde geldt voor de correlatie tussen de echografie uitkomsten van beide submandibulaire speekselklieren. De correlatie tussen de echografie uitkomsten van de parotis en submandibulaire speekselklier aan één zijde van het lichaam is echter iets lager. Hierbij maakt het niet uit of er naar de klieren aan de linkerzijde of rechterzijde van het gelaat wordt gekeken.

Vergeleken met de originele Hocevar score (vier speekselklieren en vijf aspecten) blijkt het scoren van twee componenten (echogeniciteit en aanwezigheid van hypo-echogene gebieden) in één parotis en één submandibulaire speekselklier genoeg om te voorspellen of patiënten voldoen aan de ACR-EULAR classificatie criteria. Als je vervolgens ook nog de betrouwbaarheid van de verschillende echografische aspecten in ogenschouw neemt, tonen de resultaten aan dat je, om te voorspellen of een patiënt zal voldoen aan de ACR-EULAR criteria, kunt volstaan met het scoren van hypo-echogene gebieden in één parotis en één submandibulaire speekselklier.

In **hoofdstuk 4** is er een gedetailleerde vergelijking gemaakt tussen drie verschillende methoden om betrokkenheid van de parotis speekselklier in kaart te brengen, namelijk histopathologie, sialometrie en echografie. De analyses naar de constructvaliditeit tussen deze verschillende modaliteiten zijn gedaan in een groep van 41 Sjögren patiënten, die allemaal voldoen aan de ACR-EULAR classificatie criteria.

Er is enige relatie tussen de aanwezigheid van hypo-echogene gebieden in de parotis en het percentage infiltraat wat positief is voor CD45. Verder blijkt dat de aan- of afwezigheid van hypo-echogene gebieden niet voorspelt of er nog andere specifieke histologische kenmerken van een biop, zoals plasma cel shift, lymfoepitheliale laesies of kiemcentra aanwezig zijn. Daarnaast zijn er slechte relaties gevonden tussen echografie van de parotis en parotis speeksel, wat impliceert dat echografische afwijkingen slechts ten dele worden weerspiegeld door een verminderde speeksel secretie. De fase van de ziekte waarin een patiënt zich op dat moment bevindt zou ook van belang kunnen zijn.

Op basis hiervan kan geconcludeerd worden dat alle drie methoden om betrokkenheid van de parotis speekselklier bij Sjögren patiënten in kaart te brengen andere (unieke) informatie oplevert en ze hierdoor complementair zijn aan elkaar.

In **hoofdstuk 5** is onderzocht of het toevoegen van speekselklier echografie aan de huidige classificatie criteria voor het syndroom van Sjögren, of het vervangen van één van de huidige criteria items door speekselklier echografie, invloed heeft op de nauwkeurigheid van de ACR-EULAR criteria. Er is een grote klinische cohort studie verricht met 243 patiënten die binnen de multidisciplinaire setting van het Sjögren Expertise centrum in het UMCG volledig in kaart zijn gebracht. Een gemiddelde speekselklier echografie score van ≥ 1.5 op basis van aanwezigheid van hypo-echogene gebieden wordt als positief beschouwd.

De ACR-EULAR criteria zijn het meest accuraat wanneer er één punt wordt toegekend aan een positieve echo. Verder blijft het optimale afkappunt om onderscheid te maken tussen Sjögren en non-Sjögren ≥ 4 punten.

Het toevoegen van speekselklier echografie aan de ACR-EULAR criteria zorgt voor een vergelijkbare accuraatheid van de criteria. Het vergroot echter de toepasbaarheid van de criteria, aangezien klinici uit een breder arsenaal aan onderzoeken kunnen kiezen. Wanneer OSS, de Schirmer's test, of UWS worden vervangen door echografie van de grote speekselklieren, leidt dit niet tot grote veranderingen in de accuraatheid van de criteria. Speekselklier echografie kan echter niet het biop of de bepaling van anti-SSA/Ro antilichamen vervangen.

In **hoofdstuk 6** is onderzocht in hoeverre er een relatie is tussen echografische afwijkingen van de grote speekselklieren en patiënt karakteristieken, ziekte activiteit en schade door het syndroom van Sjögren. Dit is gedaan in een groep van 172 Sjögren patiënten die zijn geïncludeerd in het REgistry of Sjögren syndrome in Umcg – LongiTudinal (RESULT) cohort. Het RESULT cohort is een prospectief observationeel cohort en heeft als doel om biomarkers en klinische parameters van patiënten te identificeren, die het beloop van de ziekte op de lange termijn zouden kunnen voorspellen.

Patiënten met een positieve echo van de grote speekselklieren, hier gedefinieerd als totale Hocevar score ≥ 15 , hebben een ander klinisch fenotype in vergelijking met patiënten met een negatieve echo. Binnen alle aspecten van de ziekte worden deze verschillen gevonden. Dit geldt niet alleen voor de klinische, functionele en serologische aspecten van de ziekte, maar ook op het gebied van de patiënt-geschatte uitkomsten zijn er verschillen gevonden tussen patiënten met een positieve en patiënten met een negatieve echo.

Echografie van de grote speekselklieren zou hierdoor mogelijk een geschikt instrument kunnen zijn om patiënten met Sjögren te stratificeren. Echter, de resultaten zijn niet exact hetzelfde wanneer alleen het aspect hypo-echogene gebieden wordt meegenomen. Dit laat zien dat het gebruik van speekselklier echografie voor diagnostische doeleinden een wezenlijk andere toepassing is dan het gebruik voor stratificatie, waarvoor een uitgebreider score systeem wenselijk is.

In **hoofdstuk 7** wordt benadrukt dat baseline karakteristieken van de parotis speekselklier gebruikt kunnen worden om te voorspellen welke patiënten baat zullen hebben bij behandeling met rituximab. Dit kan de clinicus helpen om een therapie te kiezen die past bij de specifieke patiënt.

In **hoofdstuk 8** is de effectiviteit en veiligheid van behandeling met abatacept onderzocht in een gerandomiseerde, dubbel-blinde, placebo-controleerde studie genaamd ASAP-III (Abatacept Sjögren Active Patients III). De studie populatie bestond uit 80 patiënten, allen 18 jaar of ouder, geclassificeerd als Sjögren op basis van de AECG criteria, bij wie de diagnose Sjögren maximaal 7 jaar geleden gesteld was en die actieve ziekte hadden (gedefinieerd als een score van 5 of hoger), gemeten met de internationaal gebruikte index voor systemische ziekteactiviteit, de EULAR Sjögren's Syndrome Disease Activity Index (ESSDAI). Patiënten werden willekeurig ingedeeld in twee groepen, waarbij de ene groep gedurende 24 weken behandeling kreeg middels subcutane abatacept injecties en de andere groep middels subcutane placebo injecties.

Abatacept wordt goed verdragen, er zijn nauwelijks patiënten die ernstige bijwerkingen krijgen. Helaas is er geen statistisch significant verschil gevonden in de primaire uitkomstmaat van de studie, namelijk de ESSDAI score na 24 weken tussen patiënten die behandeld werden met abatacept versus placebo. Ook is er geen verschil gevonden in ESSPRI score (een internationaal gebruikte index voor patiënt-gerapporteerde symptomen) tussen beide behandelgroepen, al is er in de abatacept groep een hoger percentage patiënten die een klinisch relevante verbetering laat zien ten opzichte van de placebo groep. Abatacept laat wel gunstige effecten zien op het biologische domein van de ESSDAI, namelijk een significant verschil in IgG en reumafactor tussen beide behandelgroepen.

In deze studie is er op baseline en na 24 weken een echo van de grote speekselklieren gemaakt. De analyses naar het eventuele effect van abatacept behandeling op echografische afwijkingen worden momenteel gedaan.

Verder is er een samengestelde uitkomstmaat in ontwikkeling. Deze maat neemt verschillende klinische aspecten in ogenschouw, waaronder echografie van de speekselklieren. Deze nieuwe ontwikkelingen kunnen klinici helpen om te bepalen welke patiënten mogelijk baat kunnen hebben bij behandeling met abatacept.

In het review in **hoofdstuk 9** zijn verschillende interessante ontwikkelingen uit het jaar 2019-2020 bediscussieerd met betrekking tot speekselklier betrokkenheid bij patiënten gediagnosticeerd of met een klinische verdenking op het syndroom van Sjögren.

Naar aanleiding van de literatuur zijn er een aantal conclusies getrokken. Allereerst lijkt de mogelijkheid om milde klierzwellen middels palpatie te detecteren te worden overschat. De potentiële rol die echografie van de grote speekselklieren hierbij zou kunnen spelen moet nog verder worden onderzocht. In het afgelopen jaar zijn er tevens belangrijke stappen gezet om een nieuw semi-kwantitatief score systeem te ontwikkelen. Echter blijft het belangrijk om echo's met enige regelmaat samen met collega's te scoren, om te voorkomen dat de een structureel hoger scoort dan de ander.

Uiteindelijk is er geconcludeerd dat echografie van de grote speekselklieren onderdeel zou moeten zijn van de diagnostische work-up van mogelijke Sjögren patiënten en dat het toevoegen van speekselklier echografie aan de huidige ACR-EULAR classificatie criteria deze meer in balans zou brengen.

Dankwoord

Het is volbracht! Na een aantal jaren met veel plezier aan dit proefschrift te hebben gewerkt is het nu dan ook echt af. Zonder hulp van een veelzijdig team was dit natuurlijk niet gelukt en ik wil enkele mensen hiervoor dan ook speciaal bedanken. Voordat ik hiermee begin wil ik eerst alle Sjögren patiënten bedanken, want zonder patiënten geen (klinisch) onderzoek.

Mijn eerste promotor, **prof. dr. H. Bootsma**. Beste Hendrika, ik kan mij onze eerste ontmoeting nog goed herinneren. Als M1 coassistent had ik een afspraak gemaakt, omdat ik graag mijn wetenschapsstage en deel van mijn semi-arts bij de afdeling Reumatologie en Klinische Immunologie wilde doen. Tijdens dit gesprek werden zelfs de eerste plannen voor een promotietraject en de opleiding tot reumatoloog uitgestippeld, want er was vanaf het eerste moment een klik. Ondanks je volle agenda kon er altijd tijd worden vrijgemaakt om, vaak samen met Frans, de vele presentaties voor nationale en internationale congressen tot in de puntjes (letterlijk) voor te bereiden. Bedankt dat ik de kansen kreeg om ook zelf te spreken tijdens deze congressen en daarnaast dank dat ik mijzelf de afgelopen jaren zowel op het werk, als ook op persoonlijk vlak heb mogen ontwikkelen.

Mijn tweede promotor, **prof. dr. F.G.M. Kroese**. Beste Frans, al tijdens de bachelor wist je indruk te maken met je boeiende en enthousiasmerende manier van lesgeven. In de flat zorg je voor een prettige en veilige werkomgeving. Je deur stond altijd voor mij open, niet alleen zodat ik even snel mijn zoveelste nieuwe gemaakte figuur kon laten zien, maar ook voor een goed gesprek of het voorbereiden van een presentatie, dank hiervoor.

Mijn co-promotor, **dr. S. Arends**. Beste Suzanne, onbewust hadden we elkaar al ontmoet, voordat ik tijdens mijn wetenschappelijke stage aan jou gekoppeld werd. Vanaf het eerste moment kon ik bij jou terecht als ik vastliep met de analyses. Tijdens het schrijven van de manuscripten heb ik veel aan je feedback gehad en met je kritische blik wist je de kleinste ongeregelheden, zoals een vergeten komma, op te sporen. Bedankt voor je fijne begeleiding en je oprechte interesse ook buiten het werk om.

Prof. dr. A. Vissink, beste Arjan, dank voor je immer snelle reactie en suggesties hoe een manuscript in te korten. Verder was je altijd beschikbaar voor vragen. Bedankt voor je begeleiding.

Ook **prof. dr. F.K.L. Spijkervet** wil ik bedanken voor de begeleiding binnen het Sjögren Expertise team. Zonder parotis en lipbiopten zou er waardevolle informatie ontbreken, en niet alleen in dit proefschrift.

Furthermore, I would like to thank the members of the reading committee: **prof. A.W.J.M. Glaudemans, prof V. Devauchelle-Pensec, and prof. J.M. Van Laar**. Thank you for spending some of your precious time on reading and judging the manuscript of my thesis.

Graag wil ik ook mijn twee paranimfen **Greetje & Marthe** bedanken.

Lieve Greetje, dank voor alle gezellige momenten en de fijne samenwerking. Het hoogtepunt was toch wel het concert van Celine Dion waar we samen geweest zijn. We hebben veel gelachen en ik heb, zeker op communicatief gebied, veel van jou mogen leren. Je heldere uitleg zorgt ervoor dat eenieder je kan volgen en in combinatie met je empathische vermogen voelen patiënten (maar ook collega's) zich gehoord.

Lieve Marthe, dank voor je gezellige aanwezigheid, je humor, maar bovenal ook de prettige samenwerking. Je hebt me er bij het laatste hoofdstuk er echt doorheen gesleept. Zonder jouw inzet was het nu nog lang niet af. Je timing voor het uitdelen van lekkere versnaperingen, meestal in de vorm van chocolade, is ook een kunst op zich. Dank voor je luisterende oor en veel succes met het afronden van je studie en proefschrift.

Beste collega's van de Sjögren/SLE groep: dank voor de samenwerking, jullie expertise en de gezelligheid. Ik mag terugdenken aan een hele waardevolle tijd in de flat, waar er altijd wel een goede reden was voor lekkernijen. Werden er even geen artikelen geaccepteerd, dan waren daar de Kerstbrunch, het Chinese dumpling avontuur of een afscheidsbrunch, om maar eens wat te noemen.

Bedankt dat ik met vragen altijd bij jou terecht kon **Jolien**, ook bij het verkennen van Washington en San Diego wist jij wat de 'highlights' waren die we zeker gezien moesten hebben. **Konstantina**, samen in Madrid op avontuur, jij wist leuke locaties te vinden om te eten en ik wist de weg weer terug. Bedankt voor de mooie herinneringen en dat ik je paranimf mocht zijn. **Sarah**, bedankt dat ik bij jou terecht kon als ik feedback nodig had op de suggesties die ik kreeg op mijn Engelse schrijven. We hebben veel gelachen, als jij een gezegde toch wel heel grappig interpreteerde, of ik met de zoveelste mango kwam aanzetten bij de lunch. **Gwenny**, bedankt voor je enthousiasme en immer positieve energie. **Xiaoyan**, thank you for bringing a little bit of China to Groningen. **Robin, Erlin, Rada** en **Taco**, dank voor jullie exper-

tise, prettige gesprekken en wijze raad. **Annie, Silvia en Janneke**, bedankt voor de gezelligheid en de zomerse wandelingen door het mooie Groningen. Verder wil ik **Uzma, Rick en Liseth** de komende jaren veel succes wensen met het onderzoek.

Dr. A.J. Stel, beste Alja, bedankt dat je mij hebt geleerd hoe ik de grote speekselkieren middels echografie goed in beeld kan brengen en kan beoordelen.

Ik wil onder andere **Janita, Ellis, Ragonda en Eefke** bedanken voor de voor de hulp bij het RESULT cohort, zowel logistiek als op de polikliniek. Verder wil ik **Janny, Marjolein en Kiki** bedanken voor de administratieve hulp.

Er is een multidisciplinaire samenwerking tussen de afdeling Reumatologie en Klinische Immunologie en onder andere de afdelingen Kaakchirurgie, Oogheelkunde en Pathologie, wat van groot belang niet alleen bij het doen van onderzoek, maar zeker ook voor de patiëntenzorg. Een aantal mensen wil ik hierbij in het bijzonder bedanken: **Judith Baldi, Sitske Oort, Jenny van den Akker** en alle mondhygiënist, bedankt voor jullie bijdrage aan de logistiek van de speekselafnames. **Drs. Nicole Sillevius Smit – Kamminga, dr. Leonie Los, dr. Jelle Vehof en Lisette Olie**, bedankt voor jullie bijdrage met betrekking tot het verzamelen van de oogheelkundige data en hulp bij de interpretatie hiervan. **Dr. Bert van der Vegt**, bedankt voor het beoordelen van de vele speekselklier bipten.

Dr. K. Hoogenberg en dr. B. Kuenen, beste Klaas en Bart, opleiders Interne Geneeskunde in het Martini ziekenhuis. Bedankt voor de tijd en ruimte die ik gekregen heb om mijn proefschrift af te maken. Dank ook aan alle andere internisten en collega arts-assistenten in het Martini ziekenhuis voor de prettige sfeer en voor het creëren van een werkplek waar ik mijzelf mag ontplooien, als dokter en als mens.

Ook wil ik mijn vriendinnen van buiten de onderzoekswereld bedanken voor jullie steun, gezelligheid en immer luisterende oor. **Lieve middelbare school, psychologie en geneeskunde vriendinnen**, ook al zien we elkaar niet zo vaak, telkens weer is het als vanouds vertrouwd. Bedankt voor jullie waardevolle vriendschap. Lieve **Tamara**, nog steeds ben ik blij dat jij opnieuw besloot om met World Servants naar Afrika te gaan en we hierover aan de praat raakten. Dit bleek het eerste begin van een inmiddels jarenlange waardevolle vriendschap, dank hiervoor. Lieve **Pietsje**, inmiddels zijn we samen vele herinneringen en ervaringen rijker en op dit moment nog samen in verwachting ook. Ik kan bij jou terecht als ik een luisterend oor nodig heb, maar ook voor een gezellige (fanatieke) spelletjesavond kan ik/ kunnen wij bij jou en Frank aankloppen. Ook jij bedankt voor je waardevolle vriendschap.

Lieve **Ane**, **Wietske** en **Henk**, **Richtsje** en **Lodewijk**, **Maaïke** en **Marco**. Vanaf begin af aan heb ik mij ontzettend welkom gevoeld bij jullie. Bedankt voor jullie interesse en steun en voor het kritisch doorlezen van de Nederlandse samenvatting.

Lieve **heit** en **mem**, **Tamarah** en **Roger** en **Sjadirah**. Het is nu toch echt af, dat ‘ingewikkelde’ proefschrift. Heit en mem bedankt dat jullie mij de kans hebben gegeven mij te ontplooiën en geneeskunde te gaan studeren en mem natuurlijk ook dank voor het vele oppassen op Josefien, zodat ik kon afmaken waar ik aan begonnen was. Gelukkig was er ook genoeg tijd voor ontspanning, bijvoorbeeld in de vorm van een spelletje, al gaat dat er (uiteraard) altijd bloedfanatiek aan toe. Bedankt voor jullie steun en interesse.

Lieve **Jelle**, mijn lieve man en steun en toeverlaat. Vaak genoeg heb jij je afgevraagd wat ik toch hele dagen aan het doen was daar in die flat, maar desalniettemin stond je altijd voor mij klaar. Vele taxiritjes van en naar het ziekenhuis, maar ook als ik naar het buitenland op congres ging hielp jij met het voorbereiden van de route, zodat verdwalen geen optie meer was. Samen met onze lieve Josefien (en baby broertje of zusje) ben jij een verrijking voor mijn leven. Lieve **Josefien**, wat ben ik blij dat ik jouw mama mag zijn. Ik geniet van onze momenten samen, je armpjes die mij lief omhelzen en je kusjes en knuffels, maar ook van je liefde voor boekjes en spelletjes. Ik kijk uit naar de mooie avonturen met jullie die komen gaan.

I'll draw the years all passing by, so much to learn so much to try.

I offer what you cannot buy, devoted love until we die.

– Celine Dion

Curriculum vitae

Esther-Jellina Mossel was born in Hurdegaryp on October 15th, 1990. After finishing high school in 2009 at 't Ichtus College in Kampen, she studied psychology at the University of Groningen. In 2010 she started her medicine study at the University of Groningen. In the final year of her medical education, she followed clinical internships in Rheumatology, and wrote a thesis on the validity of major salivary gland ultrasonography in patients suspected with primary Sjögren's syndrome. After obtaining her medical degree in 2016, she continued as a PhD candidate under the guidance of prof. dr. H. Bootsma, prof. dr. F.G.M. Kroese, and dr. S. Arends, at the department of Rheumatology and Clinical Immunology of the University Medical Center Groningen. The overall aim of her PhD thesis was to assess the potential role of salivary gland ultrasonography in patients diagnosed with or suspected for primary Sjögren's syndrome. She presented her work at several national and international conferences including the American College of Rheumatology (ACR), the European League Against Rheumatism (EULAR), the International Symposium on Sjögren's syndrome (ISSS), the Dutch Society for Rheumatology (NVR) annual meeting, and the Systemic Diseases Symposium. In 2019, she enrolled in the training to become a rheumatologist at the University Medical Center Groningen, starting as a resident in Internal Medicine at the Martini hospital in Groningen.

List of publications

Van Ginkel MS, Glaudemans AWJM, van der Vegt B, **Mossel E**, Kroese FGM, Bootsma H, Vissink A. Imaging in primary Sjögren's syndrome. *J Clin Med* 2020; aug3 9: E2492.

Hammenfors DS, Valim V, Bica BERG, Pasoto SG, Lilleby V, Nieto-González JC, Silva CA, **Mossel E**, Pereira RMR, Coelho A, Bootsma H, Thatayatikom A, Brun JG, Jons-son MV. Juvenile Sjögren's syndrome: clinical characteristics with focus on salivary gland ultrasonography. *Arthritis Care Res (Hoboken)* 2020;72:78-87.

Moerman RV, Arends S, **Mossel E**, Kroese FGM, Vissink A, Bootsma H. 10-year follow-up of patients with rheumatoid arthritis and secondary Sjögren syndrome or sicca symptoms in daily clinical practice. *Clin Exp Rheumatol* 2020;38Suppl126:64-72.

Mossel E, Delli K, van Nimwegen JF, Stel AJ, Kroese FGM, Spijkervet FGM, Vissink A, Arends S, Bootsma H, EULAR US-pSS study group. Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome. *Ann Rheum Dis* 2017;76:1883-89.

Mossel E, Delli K, van Nimwegen JF, Stel AJ, Haacke EA, Kroese FGM, Spijkervet FKL, Vissink A, Arends S, Bootsma H. The parotid gland connection: ultrasound and biopsies in primary Sjögren's syndrome. *Ann Rheum Dis* 2018;77:e38.

Mossel E, Arends S, van Nimwegen JF, Delli K, Stel AJ, Kroese FGM, Spijkervet FKL, Vissink A, Bootsma H, EULAR US-pSS study group. 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.

Mossel E, Delli K, Arends S, Haacke EA, van der Vegt B, van Nimwegen JF, Stel AJ, Spijkervet FKL, Vissink A, Kroese FGM, Bootsma H. Can ultrasound of the major salivary glands assess histopathological changes induced by treatment with rituximab in primary Sjögren's syndrome? *Ann Rheum Dis* 2019;78:e27.

Mossel E, van Nimwegen JF, Stel AJ, Wijnsma RF, van Zuiden GS, Olie L, Vehof J, Los LI, Vissink A, Kroese FGM, Arends S, Bootsma H. Clinical phenotyping of primary

Sjögren's patients using salivary gland ultrasonography – data from the RESULT cohort. *The Journal of Rheumatology* 2020;published online first October 1st 2020.

Mossel E, Arends S, Bootsma H. Recent insights in the potential role of imaging modalities for diagnosing patients with primary Sjögren's syndrome. *Clin Exp Rheumatol* 2020;38Suppl126:310-4.

Mossel E, van Ginkel MS, Haacke EA, Arends S, Liefers S, Delli K, van Nimwegen JF, Stel AJ, Spijkervet FKL, van der Vegt B, Vissink A, Kroese FGM, Bootsma H. Histopathology, salivary flow and ultrasonography of the parotid gland: three complementary measurements in primary Sjögren's syndrome. *Submitted manuscript*.

Van Nimwegen JF, **Mossel E**, Delli K, van Ginkel MS, Stel AJ, Kroese FGM, Spijkervet FKL, Vissink A, Arends S, Bootsma H. Incorporation of salivary gland ultrasonography into the American College of Rheumatology/European League Against Rheumatism criteria for primary Sjögren's syndrome. *Arthritis Care Res (Hoboken)* 2020;72:583-90.

Van Nimwegen JF, **Mossel E**, van Zuiden GS, Wijnsma RF, Delli K, Stel AJ, van der Vegt B, Haacke EA, Olie L, Los LI, Verstappen GM, Pringle SA, Spijkervet FKL, Kroese FGM, Vissink A, Arends S, Bootsma H. Abatacept treatment for patients with early active primary Sjögren's syndrome: a single-centre, randomised, double-blind, placebo-controlled, phase 3 trial (ASAP-III study). *Lancet Rheumatology* 2020;9913:1-11.

Verstappen GM, Moerman RV, van Nimwegen JF, van Ginkel MS, Bijzet, J, **Mossel E**, Vissink A, Hazenberg BPC, Arends S, Kroese FGM, Bootsma H. Serum immunoglobulin free light chains are sensitive biomarkers for monitoring disease activity and treatment response in primary Sjögren's syndrome. *Rheumatology (Oxford)* 2018;57:1812-21.

Verstappen GM, Nakshbandi U, **Mossel E**, Haacke EA, van der Vegt B, Vissink A, Bootsma H, Kroese FGM. Is the T follicular regulatory: follicular helper T cell ratio in blood a biomarker for ectopic lymphoid structure formation in Sjögren's syndrome? Comment of the article by Fonseca et al. *Arthritis Rheumatol* 2018;70:1354-55.

List of abbreviations

ACR	American College of Rheumatology
AD	Adjusted difference
AE	Adverse event
AECG	American European Consensus Group
AI	Artificial intelligence
ALAT	Alanine aminotransferase
ANA	Anti-nuclear antibody
ASAPIII	Abatacept Sjögren active patients phase III trial
ASAT	Aspartate aminotransferase
AUC	Area under the ROC curve
BDMARD	Biological disease-modifying anti-rheumatic drug
BMI	Body mass index
CI	Confidence interval
CLINESSDAI	Clinical EULAR Sjögren's Syndrome Disease Activity Index
CNB	Core needle biopsy
CNS	Central nervous system
CRESS	Composite of Relevant Endpoints in Sjögren's Syndrome
CRP	C-reactive protein
CTDS	Connective tissue diseases
DAS-28	28-joint activity score
DDF	Dryness dominant with fatigue
DIA	Digital image analysis
DMARD	Disease-modifying anti-rheumatic drug
EMA	European Medicines Agency
EQ-5D-5L	5-level EuroQol five dimensions health status questionnaire
ESR	Erythrocyte sedimentation rate
ESSDAI	EULAR Sjögren Syndrome Disease Activity Index
ESSPRI	EULAR Sjögren Syndrome Patient Reported Index
EULAR	European League against rheumatism
FDA	Food and drug administration
FS	Focus score
FSFI	Female sexual function index
GCS	Germinal center
GDA	Global disease activity
GEE	Generalized estimating equations

H&E	Hematoxylin and eosin
HADS	Hospital anxiety and depression scale
HSB	High symptom burden
IGA	Immunoglobulin A
IGG	Immunoglobulin G
IMI	Innovative medicines initiative
IQR	Interquartile range
ITT	Intention to treat
LELS	Lymphoepithelial lesions
LSB	Low symptom burden
MALT	Mucosa-associated lymphoid tissue
MCII	Minimally clinically important improvement
MCS	Mental component summary
METC	Medical Ethics Committee
MFI	Multidimensional fatigue inventory
NA	Not available/applicable
NPV	Negative predictive value
NRS	Numeric rating scale
OMERACT	Outcome Measures in Rheumatology Clinical Trials
OR	Odds ratio
OSS	Ocular staining score
PASS	Patient acceptable symptom state
PCS	Physical component summary
PDF	Pain dominant with fatigue
PGUS	Parotid gland ultrasonography
PGUS-hyper	Parotid gland ultrasonography – hyperechogenic reflections
PGUS-hypo	Parotid gland ultrasonography – hypoechogenic areas
PGUS-total	Parotid gland ultrasonography – total score
PNS	Peripheral nervous system
PPV	Positive predictive value
PRO	Patient reported outcome
PROM	Patient-reported outcome measurement
PSS	Primary Sjögren's syndrome
RCT	Randomized controlled trial
RESULT	REgistry of Sjögren syndrome in Umcg – LongiTudinal
RF	Rheumatoid factor
ROC	Receiver operating characteristic
SAE	Serious adverse event

SCLE	Subacute cutaneous lupus erythematosus
SD	Standard deviation
SF-36	36-item short form health survey
SGUS	Salivary gland ultrasonography
SICCA	Sjögren's International Collaborative Clinical Alliance
SLE	Systemic lupus erythematosus
SPF	Stimulated parotid flow
SS	Sjögren's syndrome
SSA	Sjögren's syndrome-related antigen A (Ro)
SSB	Sjögren's syndrome-related antigen B (La)
SSDAI	Sjögren's syndrome disease activity index
SSDDI	Sjögren's syndrome disease damage index
SSRI	Sjögren's syndrome responder index
SUSAR	Suspected unexpected serious adverse reaction
SWE	Shear wave elastography
SWS	Stimulated whole saliva
SWSF	Stimulated whole salivary flow
TBUT	Tear break-up time
TEARS	Tolerance and Efficacy of Rituximab in Primary Sjögren's Syndrome
TFH CELL	T follicular helper cell
TRACTISS	TRial of Anti-B-Cell Therapy in patients with primary Sjögren's Syndrome
TUS	Total ultrasound score
UMCG	University Medical Center Groningen
UPF	Unstimulated parotid flow
UWS	Unstimulated whole saliva
UWSF	Unstimulated whole salivary flow
VAS	Visual analogue scale
WPAI	Work participation and activity impairment
%CD45	Percentage of CD45 ⁺ infiltrate