Salivary gland ultrasonography: can it be an alternative to sialography as an imaging modality for Sjögren's syndrome?

Yukinori Takagi, Yasuo Kimura, Hideki Nakamura, et al.

*Ann Rheum Dis* 2010 69: 1321-1324 originally published online May 24, 2010
doi: 10.1136/ard.2009.123836

Updated information and services can be found at:
http://ard.bmj.com/content/69/7/1321.full.html

These include:

**References**
This article cites 13 articles, 8 of which can be accessed free at:
http://ard.bmj.com/content/69/7/1321.full.html#ref-list-1

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To order reprints of this article go to:
http://ard.bmj.com/cgi/reprintform

To subscribe to *Annals of the Rheumatic Diseases* go to:
http://ard.bmj.com/subscriptions
Salivary gland ultrasonography: can it be an alternative to sialography as an imaging modality for Sjögren’s syndrome?

Yukinori Takagi,1 Yasuo Kimura,1 Hideki Nakamura,2 Miho Sasaki,1 Katsumi Eguchi,2 Takashi Nakamura1

ABSTRACT

Objectives Salivary ultrasonography (US) was evaluated as an alternative imaging modality to sialography for Sjögren’s syndrome (SS).

Methods Parotid sialography and parotid and submandibular US were performed in 360 patients (188 with SS and 172 non-SS) who were suspected of SS and had fulfilled the American-European criteria (AEC) for SS or had received ≥3 of the objective examinations that were considered minimal requirements for classifying non-SS and had undergone both the imaging examinations. The glands were considered positive for SS if they exhibited peripheral sialoectatic changes on sialography and/or hypoechoic areas, echogenic streaks and/or irregular gland margins on US. The images obtained were independently rated as SS-positive or SS-negative by three radiologists in a blind fashion and the final decision was made by consensus. Interobserver and intermodality agreement was evaluated using $\kappa$ values for sialography and parotid and/or submandibular US.

Results Average $\kappa$ values for the interobserver agreement were 0.81, 0.80 and 0.82 in sialography, parotid and submandibular US, respectively, indicating very good or good agreement. The $\kappa$ value for intermodality agreement between sialography and parotid US was 0.81 and between sialography and submandibular US was 0.76, indicating very good and good agreement, respectively. The diagnostic ability of parotid US was significantly lower than that of sialography ($p<0.001$, McNemar test). However, the diagnostic ability of submandibular US was comparable to that of sialography ($p=0.153$).

Conclusions Submandibular US is a promising technique that can be used as a practical alternative to sialography in the classification of SS.

INTRODUCTION

Sjögren’s syndrome (SS) is an autoimmune disease that affects exocrine systems, including the salivary and lacrimal glands. Its diagnosis is based on the oral and ocular symptoms; ocular signs as determined by the Schirmer test and/or ocular dye scores; histopathology of the labial or lacrimal glands; assessment of salivary gland involvement as determined by salivary flow test, parotid sialography or salivary scintigraphy; and serological test for SS-A or SS-B.1 However, to date, the ‘gold standard’ for SS diagnosis has not been established. Clinicians therefore usually diagnose the disease based on several combinations of the abovementioned criteria. The American-European criteria (AEC) for SS include sialography and scintigraphy for the assessment of salivary gland involvement.1 Recent studies have shown that salivary ultrasonography (US) is also useful for this purpose.2–4 Compared with the aforementioned two imaging techniques, US assessment of salivary gland involvement is non-invasive, of lower cost and does not require radiation. US could therefore be an alternative to the classic imaging techniques for the diagnosis of SS.5 However, the diagnostic accuracy of US relative to that of the established imaging techniques has not been fully investigated, and the consensus of employing US as an alternative imaging technique to sialography has not yet been achieved.

In this study we therefore directly compared the usefulness of salivary US with that of sialography by analysing the interobserver and intermodality agreement for parotid sialography and parotid and/or submandibular US in a large patient cohort comprising both patients with or without SS on the basis of predefined criteria. The $\kappa$ value is an index that compares the agreement against that which might be expected by chance. To test whether salivary US can replace sialography in classifying SS, we evaluated the interobserver agreement for sialography and parotid and/or submandibular US using $\kappa$ values, thereby evaluating the reliability of each imaging modality. We also evaluated the intermodality agreement between salivary US and sialography for assessing the clinical interchangeability between these techniques.

METHODS

Patients
Between June 1993 and May 2009, 1276 patients were referred to our outpatient clinic with suspected SS; 864 patients underwent both sialography and US, 360 of which were selected for the present retrospective study on the basis of the following selection criteria: patients who were diagnosed as having SS on the basis of AEC (SS group) and patients who underwent three or more of the objective examinations (oral, ocular, serological and pathological) but did not fulfil the AEC (non-SS group).1 Consequently, the study cohort comprised 188 patients with SS (168 women and 20 men; mean age 56±13 years) and 172 non-SS patients (146 women and 26 men; mean age 55±16 years). Of the 188 patients with SS, 134 were cases of primary SS and 54 were cases of secondary SS.
Ultrasonography
Grey-scale US was performed at 10 MHz using a Logiq 700 or Logiq 9 unit equipped with a wide band width (6–14 MHz) (General Electric Medical Systems, Milwaukee, Wisconsin, USA). We used this frequency (10 MHz) in order to obtain echoes from the deep as well as superficial portions of the glands. The performances of both the US units are similar. Patients were scanned in the supine position with their necks extended and head turned a little to the opposite side. The parotid glands were examined in both axial and coronal planes and the submandibular glands only in longitudinal planes along the inferior border of the mandible. The sonographers were blind to the diagnosis of SS or to the sialographic findings. The US diagnosis for the patients with SS was based on the internal echoes of the parotid and submandibular gland images and on the gland contour, and the obtained gland images were categorised into five grades (G0–G4) as described previously: glands positive for SS were determined by the presence of (1) irregular contour of the gland; (2) multiple, round or irregular hypoechoic spots or areas in the gland; and (3) multiple echoic bands in the gland (figure 1).6

Sialography
Sialography of the parotid glands was performed using a non-ionising contrast medium (Iopamiron; Schering AG, Berlin, Germany). The sialographic stages of SS were determined according to the criteria of Rubin and Holt, and the images categorised into five grades (G0–G4) as described previously (figure 1).7

Serological, salivary flow, ocular and pathological examinations
The results of serological (anti-SS-A or anti-SS-B antibodies), salivary flow (Saxon test), ocular (Schirmer test or ocular dye test) and pathological (lip biopsy) examinations were classified as positive or negative on the basis of the AEC.1 We made the following modifications to the original classification criteria. The AEC classification recommends the measurement of unstimulated whole saliva (≤1.5 ml in 15 min) for an objective evaluation of salivary gland involvement. However, we used the Saxon test (≤2.0 ml in 2 min) for the salivary flow examination.

Primary and secondary SS
Primary and secondary SS were classified according to the AEC.1

Image reading and data analysis
Ultrasonographs and sialographs were read by three radiologists (12–15 years’ experience in the field of head and neck radiology) blind to the diagnosis (SS or non-SS) or any results of serological, salivary flow, ocular or pathological examinations.

Interobserver and intermodality agreements were assessed using κ values. The κ value was calculated from the equation: 
κ value = (observed agreement − expected agreement)/(1 − expected agreement). A κ value in the range 0.81–1.00 was interpreted as being very good, in the range 0.61–0.80 as good, in the range 0.41–0.60 as moderate, in the range 0.21–0.40 as fair and in the range 0.00–0.20 as poor.8

Figure 1 Ultrasoundography and sialography of the salivary gland in patients with Sjögren’s syndrome (SS). (A–E) Parotid ultrasonography. Parotid glands in (A) grade 0, (B) grade 1, (C) grade 2, (D) grade 3 and (E) grade 4. Grade 0 gland shows homogeneous echogenicity (A). Glands in higher grades (grades 1–4) show multiple hypoechoic areas and echogenic streaks (B–E). The number of hypoechoic areas increases with increasing disease grade (up to grade 3). In grade 4 glands, echogenic streaks are more evident than hypoechoic areas (E). (F–J) Submandibular ultrasonography. Submandibular glands in (F) grade 0, (G) grade 1, (H) grade 2, (I) grade 3 and (J) grade 4. Submandibular glands in patients with SS show ultrasonographic changes similar to parotid glands. However, hypoechoic areas are less noticeable in submandibular glands than in parotid glands. (K–O) Parotid sialography. Parotid glands in (K) grade 0, (L) grade 1, (M) grade 2, (N) grade 3 and (O) grade 4. Grade 0 gland (K) shows normal intraglandular ramifications; sialoectatic changes are not evident. Parotid sialography shows multiple punctuate (L), globular (M), cavitory (N) and destructive (O) sialoectasia. Panels A, F and K were obtained from a 49-year-old woman, panels B, G and L from a 62-year-old woman, panels C, H and M from a 73-year-old woman, panels D, I and N from a 57-year-old woman and panels E, J and O from a 63-year-old woman.
Diagnostic accuracy was assessed by calculating sensitivity, specificity, accuracy and positive and negative predictive values. The AEC was used as the gold standard. The significance of the difference in diagnostic accuracy (ability) between sialography and US was determined by the McNemar test.

RESULTS
Serological, salivary flow, ocular and pathological examinations
Table 1 summarises the rates of positive results obtained for serological, salivary flow, ocular and pathological examinations in the 360 patients included in the study. Serological, Saxon and Schirmer tests were performed in most of the 360 patients. These examinations yielded sensitivity and specificity of 82% and 65% (serological), 81% and 41% (Saxon test) and 87% and 50% (Schirmer test), respectively. However, lip biopsy was performed in only 108 (30%) of the 360 patients, yielding 88% sensitivity and 42% specificity. Ocular staining test and scintigraphy were rarely performed.

Interobserver agreement
Interobserver reliability of using the imaging criteria for US and sialography in classifying SS was evaluated using the $\kappa$ values. We found that the average $\kappa$ values indicated good agreement between sialography and parotid US and very good agreement between sialography and submandibular US (table 2).

Intermodality agreement
We considered that, in the absence of a gold standard for the diagnosis of SS, analysis of the agreement between sialography and US may be a crucial step for the confirmation of the usefulness of US as an alternative to sialography.

The $\kappa$ values indicated very good agreement between sialography and parotid US and good agreement between sialography and submandibular US (table 3).

Table 1 Summary of serological, oral, ocular, pathological and imaging examinations in 360 patients with or without SS

<table>
<thead>
<tr>
<th>Examinations</th>
<th>SS (188) Rates of positive/total</th>
<th>Non-SS (172) Rates of positive/total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS-A or SS-B</td>
<td>146/177</td>
<td>60/171</td>
</tr>
<tr>
<td>Saxon test</td>
<td>132/163</td>
<td>86/147</td>
</tr>
<tr>
<td>Schirmer test</td>
<td>107/123</td>
<td>74/149</td>
</tr>
<tr>
<td>Ocular staining</td>
<td>8/10</td>
<td>5/18</td>
</tr>
<tr>
<td>Lip biopsy</td>
<td>69/78</td>
<td>30/52</td>
</tr>
<tr>
<td>Scintigraphy</td>
<td>4/5</td>
<td>2/4</td>
</tr>
<tr>
<td>SG</td>
<td>148/188</td>
<td>31/172</td>
</tr>
<tr>
<td>PG US</td>
<td>128/188</td>
<td>31/172</td>
</tr>
<tr>
<td>SMG US</td>
<td>154/188</td>
<td>47/172</td>
</tr>
</tbody>
</table>

PG, parotid gland; SG, sialography; SMG, submandibular gland; SS, Sjögren’s syndrome; US, ultrasonography.

Table 2 Interobserver agreement

<table>
<thead>
<tr>
<th>Observer</th>
<th>PG US</th>
<th>SMG US</th>
<th>SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>A vs B</td>
<td>0.83</td>
<td>0.84</td>
<td>0.80</td>
</tr>
<tr>
<td>B vs C</td>
<td>0.78</td>
<td>0.79</td>
<td>0.78</td>
</tr>
<tr>
<td>C vs A</td>
<td>0.81</td>
<td>0.82</td>
<td>0.84</td>
</tr>
<tr>
<td>Average</td>
<td>0.80</td>
<td>0.82</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Three radiologists (observers A, B and C) independently classified US and SG images into SS or non-SS. PG, parotid gland; SG, sialography; SMG, submandibular gland; SS, Sjögren’s syndrome; US, ultrasonography.

Diagnostic abilities of US and sialography
Given the very good or good agreement between salivary US and sialography in the SS classification, respectively, we determined the diagnostic abilities of these imaging techniques (table 4). The diagnostic ability of parotid US was slightly but significantly lower than that of sialography. However, the diagnostic ability of submandibular US was comparable to that of sialography. The combined assessment of parotid and submandibular glands did not significantly improve the diagnostic abilities of the single assessment.

DISCUSSION
In this study we found very good and good interobserver agreement for parotid and submandibular US, respectively, in the classification of SS. In addition, intermodality agreement between parotid US and sialography and that between submandibular US and sialography was very good and good, respectively. Although the ability of parotid US in classifying SS was slightly lower than that of sialography, the diagnostic ability of submandibular US was comparable to that of sialography. Taken together, these results suggest that submandibular US could be a practical alternative to sialography in classifying SS.

The US changes in the salivary glands of patients with SS were characterised in detail and were correlated with the sialographic

Table 3 Intermodality agreement

| SG vs PG US | 0.81 (0.75 to 0.85) |
| SG vs SMG US | 0.76 (0.69 to 0.80) |
| SG vs PG and SMG US | 0.81 (0.74 to 0.84) |
| SG vs PG or SMG US | 0.76 (0.70 to 0.80) |

$\kappa$ values were calculated between SG and PG US, SMG US or PG and/or SMG US on the basis of consensus results by three radiologists. PG and/or SMG US; US results were considered positive if PG and/or SMG were classified as SS-positive. PG, parotid gland; SG, sialography; SMG, submandibular gland; SS, Sjögren’s syndrome; US, ultrasonography.

Table 4 Comparison of diagnostic abilities of sialography and ultrasonography in 360 patients in reorganised cohort

<table>
<thead>
<tr>
<th>Diagnostic ability</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>146</td>
</tr>
<tr>
<td>PG</td>
<td>128</td>
</tr>
<tr>
<td>SMG</td>
<td>154</td>
</tr>
<tr>
<td>PG or SMG</td>
<td>154</td>
</tr>
</tbody>
</table>

TP 146 128 154 154 128
TN 141 141 125 122 143
FP 31 31 47 50 29
FN 42 42 34 34 60
Total 360 360 360 360 360
Sensitivity (%) 78 68 82 82 68
Specificity (%) 82 82 73 71 83
Accuracy (%) 80 75 78 77 75
PPV (%) 82 81 77 75 82
NPV (%) 77 70 79 78 70

Diagnostic abilities were calculated on the basis of consensus SG and US results obtained by three radiologists. PG and/or SMG; US results were classified as SS-positive if PG and/or SMG results were positive. p Values <0.05 using the McNemar test were considered statistically significant. McNemar analysis was performed by comparing the numbers of patients with true results by SG and false results by US with the numbers of patients with false results by SG and true results by US. This table therefore compares the accuracy between US and SG in diagnosing SS glands.

FN, false negative; FP, false positive; PG, parotid gland; SG, sialography; SMG, submandibular gland; SS, Sjögren’s syndrome; TN, true negative; TP, true positive; US, ultrasonography.

and histological findings. These studies showed that the US findings of the parotid glands, such as hypoechoic spots and/or areas and hyperchoic streaks, are characteristic of the gland involvement in SS. Furthermore, multiple logistic regression analysis showed that, among the diagnostic criteria examined, including salivary flow rate by the Saxon test, the Schirmer test and serological tests (SS-A or SS-B), only sialographic and US results were positively correlated with the diagnosis of SS.

Quantitative texture analysis in US was introduced for the diagnosis and grading of the salivary glands affected by SS. The high diagnostic ability by this study may be attributable to the use of healthy volunteers as controls. However, the post-imaging processing is time-consuming and may be unsuitable for routine application.

Many researchers have recently used the US scoring system, which involves the combined assessment of parotid and submandibular glands to differentiate between SS and non-SS glands of patients with sicca symptoms. These studies showed that, although US yielded a high specificity, it could detect only half of the patients with SS; for example, the sum of the US scores of the parotid and submandibular glands differentiated between patients with and without SS with 59% sensitivity and 99% specificity. Wernicke et al reported sensitivity of 65% (for primary SS) or 64% (for secondary SS) and 99% specificity. In the present study, on the other hand, the submandibular gland US yielded 82% sensitivity and 73% specificity (table 4). We found that the diagnostic ability of US of both the parotid and submandibular glands was similar to that of US of either of the two glands (table 4). The lower sensitivities and higher specificities of the previous studies compared with those of the present study might be attributable to the inclusion of a large number of patients without symptoms or healthy controls in the former studies.

Considering that the assessment of both the parotid and submandibular glands would be time-consuming, we conclude that submandibular US alone may be sufficient for the assessment of salivary gland involvement in SS.

The intermodality agreement between sialography and parotid US in this study was higher than that between sialography and submandibular US (table 3). On the other hand, the diagnostic ability of submandibular US was significantly higher than that of parotid US (table 4). These results suggest that US detected changes in the submandibular glands affected by SS, which were characteristic of SS but different from those of the parotid glands.

Tzioufas and Moutsopoulos raised several issues that need to be addressed before US can be considered a reliable diagnostic tool for the assessment of SS. Of these issues, a flaw in the usefulness of US in the discrimination between glands in patients with SS and those in patients with other diseases affecting salivary function is of particular importance; for example, US of the salivary glands in patients with hepatitis C virus or HIV infection or sarcoidosis may show multiple hypoechoic foci or cystic areas, mimicking the US imaging features of the glands in patients with SS. Therefore, US alone cannot be sufficient to effectively differentiate between SS and other diseases mimicking SS. In this regard, MRI could be a useful tool for differentiating SS from diseases that mimic SS.

The US and sialographic findings were highly specific for SS. The AEC includes sialographic findings as relevant along with the findings of salivary flow and scintigraphic examinations for the estimation of salivary gland involvement. However, the findings of salivary flow and scintigraphic examinations are not specific for SS but are objective evidence of impaired salivary flow. It therefore appears that US or sialography can be included as a diagnostic modality independent of the other examinations for the assessment of salivary involvement.

Many non-SS patients were positive for SS-A/SS-B or lip biopsy (table 1). These patients included those with systemic lupus erythematosus (n = 19), rheumatoid arthritis (n = 16), human T lymphotrophic virus type 1 (HTLV-1) infection (n = 7), anti-phospholipid antibody syndrome (n = 5), HTLV-1-associated myelopathy (n = 3), mixed connective tissue disease (n = 3), CREST syndrome (n = 2) and systemic sclerosis (n = 2). It is therefore plausible that the autoimmune backgrounds caused positive results for serology and pathology in non-SS patients.

In this study we proposed US as a substitute for sialography in the imaging diagnosis of SS. However, such a type of change in the classification criteria must be carefully validated with large multidepartment studies.

CONCLUSION

The present results, together with the previous findings, suggest that the diagnostic accuracy of submandibular US is comparable to that of parotid sialography in classifying patients with SS. It is concluded that, since submandibular US is inexpensive, easy to use, non-invasive and satisfactorily efficient for the diagnosis of SS, it can be used as an alternative to the classic x-ray sialography for the classification of SS.

Ethics approval None.

Provenance and peer review Not commissioned; externally peer reviewed.

Patient consent Obtained.

REFERENCES