Sonoelastography for differential diagnosis between malignant and benign parotid lesions: a meta-analysis

Yun-Fei Zhang¹ · Hong Li² · Xue-Mei Wang¹ • Yun-Fei Cai³

Received: 24 April 2018 / Revised: 3 June 2018 / Accepted: 15 June 2018 / Published online: 10 July 2018
© The Author(s) 2018

Abstract

Objectives To assess the performance of sonoelastography for differential diagnosis between malignant and benign parotid lesions using a meta-analysis.

Methods An independent literature search of English medical databases, such as PubMed, Embase and Medline (Embase.com), Web of Science, Cochrane Library and Ovid was performed. The diagnostic accuracy of sonoelastography was compared with that of histopathology and/or cytology, which was used as reference standard. The pooled sensitivity, specificity, diagnostic odds ratio (DOR) and area under the curve (AUC) were calculated to evaluate the accuracy of sonoelastography. A meta-regression analysis evaluating imaging mechanisms, shear wave elastography techniques, assessment methods and QUADAS scores was performed.

Results Ten eligible studies that included a total sample of 711 patients with 725 parotid lesions were included. Sonoelastography showed a pooled sensitivity of 0.67 (95% CI 0.59–0.74), specificity of 0.64 (95% CI 0.60–0.68), DOR of 8.00 (95% CI 2.96–21.63) and an AUC of 0.77. The results of the meta-regression analysis revealed that no heterogeneity was due to the imaging mechanism (p = 0.119), shear wave elastography technique (p = 0.473) or QUADAS score (p = 0.462). However, the assessment method was a significant factor that affected the study heterogeneity (p = 0.035). According to the subgroup analysis, quantitative and semiquantitative methods performed better than qualitative ones.

Conclusion Overall, sonoelastography has a limited value for differential diagnosis between malignant and benign parotid lesions. Quantitative and semiquantitative methods perform better than qualitative ones.

Key Points

• Overall, sonoelastography has a limited value for differential diagnosis between malignant and benign parotid lesions.
• Quantitative and semiquantitative assessment methods perform better than qualitative ones.
• Semiquantitative and quantitative methods are automatically calculated by an ultrasound machine and are thus less operator-dependent.

Keywords Parotid neoplasms · Elasticity imaging techniques · Ultrasonography

Abbreviations

DOR Diagnostic odds ratio
FN False negative
FNAC Fine-needle aspiration cytology
FP False positive
QUADAS Quality Assessment of Diagnostic Accuracy Studies
SROC Summary receiver operating characteristic
TN True negative
TP True positive
**Introduction**

Ultrasonography, CT and MRI are non-invasive imaging methods that are commonly used for the evaluation of parotid tumours. However, these conventional imaging methods are less accurate owing to the overlap in the appearance of parotid tumours. Some malignancies that contain a large amount of serous and mucoid contents are well defined with a homogeneous appearance and resemble benign lesions. In addition, haemorrhage and calcification in benign tumours may result in a heterogeneous appearance that resembles a malignancy [1–5]. Although ultrasound-guided fine-needle aspiration cytology (FNAC) is considered the gold standard for preoperative diagnosis [6], it is an invasive method and, as a general rule, non-invasive methods are preferred when the results are similar [7].

Sonoelastography is an innovative diagnostic imaging tool that assesses tissue stiffness [8]. Since malignant tissues are generally stiffer than benign components, sonoelastography has been used in many organs, such as the breast, thyroid and prostate, for differential diagnosis between malignant and benign lesions [9–14]. Recently, numerous studies have been published on the role of sonoelastography for differentiating between malignant and benign parotid lesions. However, there are large differences in the results, with a sensitivity ranging from 40% to 100% and a specificity ranging from 26% to 97% [7, 8, 15–20]. Therefore, this study aimed to assess the performance of sonoelastography for differential diagnosis between malignant and benign parotid lesions using a meta-analysis.

**Materials and methods**

**Literature search**

The study complied with the PRISMA recommendations [21, 22]. An independent literature search of English medical databases including PubMed, Embase and Medline (Embase.com), Web of Science, Cochrane Library and Ovid was performed to identify all studies evaluating differential diagnosis between malignant and benign parotid lesions. The strategies are shown in Table 1. Duplicated articles were excluded manually. Unpublished relative data were considered as well, but no suitable studies were identified for inclusion. The study was performed by two independent researchers. This literature search was updated until 30 October 2017 and a beginning date limit was not used.

**Inclusion and exclusion criteria**

All the articles were assessed independently by two researchers. The inclusion criteria for the studies were as follows: (1) The study was approved by an ethics committee or institutional review board. (2) The diagnostic performance of sonoelastography for the differential diagnosis between malignant and benign parotid lesions was evaluated in the study. (3) Postoperative pathology and/or fine-needle aspiration cytology (and/or histology) results were used as the reference standard in the study. (4) Complete reported data were available to calculate the true positive (TP), false positive (FP), false negative (FN) and true negative (TN) cases. The exclusion criteria for the studies were as follows: (1) Reviews, case reports, letters, conference reports, editorial comments and articles that were not published in English were excluded. (2) In studies with insufficient data, the corresponding authors were contacted and requested to provide the missing data via e-mail. The studies were excluded if the author did not reply within 15 days. (3) When two or more studies were performed by the same department, the study that was older or that had the smaller number of patient samples was excluded. All the disagreements were resolved by consensus.

**Data extraction**

Two investigators extracted the data independently. All relevant data including first author, country where the study was performed, published year, patient age, proportion of male and female patients, number of patients, number of lesions, reference standard, type of lesions, ultrasound system, sonoelastography index, cut-off value and number of TPs, FPs, FNs and TNs were extracted. The cut-off value was defined according to the Youden method if it was not clearly provided by the author. Disagreements were resolved by consensus.

**Quality assessment**

The methodological qualities of primary studies were assessed with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) criteria [23]. The defined questions were answered as yes, no or unclear, and ultimately, a maximum score of 14 was used to estimate the quality of each article. Two researchers completed all the items and disagreements were resolved by consensus.

**Data analysis**

The statistical software Meta-Disc (Version 1.4, Unit of Clinical Biostatistics team of the Ramón y Cajal Hospital), STATA (Version 12.0, Stata Corporation) and SPSS Statistics (Version 17.0, SPSS Inc.) were used in this study. The Spearman correlation coefficient was used to analyse the threshold effect. The heterogeneity was evaluated by the Cochran Q statistic and the F test. A random effects model was used when the p value of heterogeneity was less than 0.05
or the I² was at least 50%, otherwise a fixed effects model was
used. The pooled sensitivity, specificity, diagnostic odds ratio
(DOR), area under the curve (AUC) and $Q^*$ index were cal-
culated using Meta-Disc. Potential sources of heterogeneity
were explored with a meta-regression analysis. Deeks’
funnel plot was generated in STATA to analyse the potential publica-
tion bias, with a $p<0.05$ indicating potential publication bias.

Interobserver agreement was analysed with Cohen’s $\kappa$
analysis using SPSS software while screening articles and applying
the QUADAS criteria.

Results

Literature search and characteristics of included
studies

Ten relative studies including 711 patients with 725 parotid
lesions were included in the meta-analysis after literature
search, which were published from 2012 to 2017 [6–8,
15–20, 24] (Fig. 1). The main characteristics of the included
studies are summarised in Table 2. Controversies occurred
between two observers in the step when the records were
excluded by title and abstract. However, it showed an ex-
cellent interobserver agreement ($\kappa = 0.86$; 95% CI 0.72–
0.99). Ultimately, all the controversial articles were
included in this step. There was no disagreement in other
steps of screening ($\kappa = 1$).

Quality assessment

Quality assessment of each study is shown in Table 3. Most of
the indexes were adequate and resulted in a high QUADAS
score. However, it was unclear if the pathologist was blinded
to the sonoelastography results in all the studies. In one study,
only pleomorphic adenomas were identified in the benign
group [16]. In one study, it was unclear if the radiologist was
blinded to the pathology [20], and in another study, the ultra-
sound examiners were aware of the histological properties of
the respective lesions [8]. The interobserver agreement was
good ($\kappa = 0.77$; 95% CI 0.60–0.93).

Diagnostic accuracy for differential diagnosis
between malignant and benign parotid lesions

No heterogeneity was identified by analysis of the diagnostic
threshold, with a Spearman correlation coefficient of 0.389 ($p$
= 0.266). The diagnostic accuracy of sonoelastography for
differential diagnosis between malignant and benign parotid
lesions was computed on the basis of a pooled sensitivity of
0.67 (95% CI 0.59–0.74), specificity of 0.64 (95% CI 0.60–
0.68) and DOR of 8.00 (95% CI 2.96–21.63) (Fig. 2). An
overall moderate degree of accuracy was identified by the summary receiver operating characteristic (SROC) curve with an AUC of 0.77 ($Q^* = 0.71$) (Fig. 3).

**Heterogeneity results**

The Cochran $Q$ test and the $I^2$ test revealed significant heterogeneity with $p < 0.001$ and $I^2 = 77.2\%$. To further explore the sources of heterogeneity, a meta-regression analysis evaluating imaging mechanisms (group 1, strain elastography (SE); group 2, shear wave elastography (SWE)), shear wave elastography techniques (group 1, supersonic shear imaging (SSI) with a SuperSonic Imagine Aixplorer; group 2, acoustic radiation force impulse imaging (ARFI) with a Siemens S2000), assessment methods (group 1, qualitative; group 2, quantitative or semiquantitative) and QUADAS scores was performed. The results indicated that no heterogeneity was due to the imaging mechanism ($p = 0.119$), shear wave elastography technique ($p = 0.473$) or QUADAS score ($p = 0.462$). However, the assessment method was a significant factor that affected the study heterogeneity ($p = 0.035$). Compared with qualitative assessment methods, quantitative and semiquantitative methods performed better (Table 4).

**Evaluation of publication bias**

Publication bias was explored with a Deeks’ funnel plot and no significant differences were detected in this meta-analysis ($p = 0.143$) (Fig. 4).

**Discussion**

Our current meta-analysis demonstrated that sonoelastography showed a pooled sensitivity of 0.67 (95% CI 0.59–0.74) and specificity of 0.64 (95% CI 0.60–0.68) for differential diagnosis between malignant and benign parotid lesions. The pooled DOR was 8.00 (95% CI 2.96–21.63) and the AUC was 0.77. The meta-regression analysis results revealed that the assessment method was a significant factor.
| Author                   | Country       | Year | Age (avg. or median) | Male/female | Number of patients | Number of lesions | Reference standard                                                                 |
|--------------------------|---------------|------|----------------------|-------------|--------------------|-------------------|-------------------------------------------------------------------------------------|
| Klintworth et al [8]     | Germany       | 2012 | 53.3                 | 27/30       | 57                 | 57                | Postoperative pathology                                                              |
| Yerli et al [24]         | Turkey        | 2012 | 47.4                 | 15/15       | 30                 | 30                | Postoperative pathology ($n = 33$) or clinical indications, fine-needle aspiration biopsy ($n = 3$) postoperative pathology |
| Badia et al [16]         | Romania       | 2013 | NA                   | 15/5        | 20                 | 20                | Postoperative pathology                                                              |
| Celebi and Mahmutoglu [17]| Turkey        | 2013 | 47.19                | 36/39       | 75                 | 81                | Postoperative pathology ($n = 73$), FNAC ($n = 6$), both ($n = 2$)                   |
| Wierzbicka et al [19]    | Poland        | 2013 | 54                   | 27/16       | 43                 | 43                | Postoperative pathology                                                              |
| Yu et al [20]            | China         | 2016 | 45                   | NA          | 51                 | 51                | Postoperative pathology                                                              |
| Altmphas et al [15]      | Turkey        | 2017 | 60.01                | 26/20       | 46                 | 54                | Fine-needle aspiration biopsy (cytological and histological results)                 |
| Cantisani et al [7]      | Italy         | 2017 | 56                   | NA          | 63                 | 63                | Postoperative pathology and FNAC                                                     |
| Herman et al [6]         | Czech Republic| 2017 | 60                   | 66/58       | 124                | 124               | Postoperative pathology                                                              |
| Mansour et al [18]       | Germany       | 2017 | 58.6                 | NA          | 202                | 202               | Postoperative pathology                                                              |

| Type of lesions (number of lesions) | Ultrasound system | Index of elastography | Cut-off value | TP | FP | FN | TN |
|-------------------------------------|--------------------|-----------------------|---------------|----|----|----|----|
| 1 Squamous cell carcinoma (3), mucoepidermoid carcinoma (1), salivary duct carcinoma (1), basal cell carcinoma (1), carcinoma ex pleomorphic adenoma (1), metastatic malignant melanoma (1); pleomorphic adenoma (22), Warthin tumour (20), cystadenoma (2), lipoma (2), cyst (3) | Siemens Ltd., Healthcare Sector, Acuson S2000 | Garland sign | Garland sign (+) | 3 | 2 | 5 | 47 |
| 2 Adenocarcinomas (2), mucoepidermoid carcinomas (2), adenoid cystic carcinomas (1), malignant lymphoma (1); Warthin tumours (10), pleomorphic adenomas (10), lymphadenopathies (3), abscesses (1) | Hitachi Medical Systems, EUB-7000 | ES (4-point) | ES3 | 4 | 8 | 2 | 16 |
| 3 Parotid carcinomas (6), metastases (2); all pleomorphic adenomas (12) | General Electric GE 7, GE 8, GE 9; Phillips, IU22; Siemens S 2000 | Tactile Elastography or ARFI | NA | 8 | 6 | 0 | 6 |
| 4 Lymphoma (9), mucoepidermoid carcinoma (6), adenoid cystic carcinoma (5), metastases (3), myoepithelial malignant tumour (1), pleomorphic adenocarcinoma (2), salivary duct carcinoma (2), acinic cell carcinoma (2) and basal cell carcinoma (2); pleomorphic adenoma (28), Warthin tumour (10), lymphadenopathy (9), cystic adenoma, infected cyst rupture (2) | Siemens Medical Solutions, Siemens S2000 | CSES (4-point) | CSES3 | 19 | 19 | 13 | 30 |
| 5 Adenocarcinoma (3), malignant lymphoma (2), clear cell carcinoma (1), non-differentiated cancer (1), squamous cell carcinoma (1), cancer from pleomorphic adenoma (1) and salivary duct cancer (1); pleomorphic adenoma (23), Warthin tumour (5), monomorphic adenoma (2), neurofibroma (1), cyst (1), basal cell adenoma (1) | Supersonic Imagine, Aixplorer | ES (5-point) | ES4 | 4 | 1 | 6 | 32 |
| 6 Mucoepidermoid carcinomas (5), acinic cell carcinomas (3), salivary duct carcinomas (3), basal cell carcinomas (2), adenoid cystic carcinoma (1), adenocarcinoma (1), carcinoma ex pleomorphic adenoma (1); pleomorphic adenoma (16), adenolymphomas (15), basal cell | Siemens Medical Solutions, Acuson S2000 | SWV | 2.76 m/s | 11 | 1 | 5 | 34 |
| Type of lesions (number of lesions) | Ultrasound system | Index of elastography | Cut-off value | TP  | FP | FN | TN |
|-----------------------------------|-------------------|-----------------------|---------------|-----|----|----|----|
| adenomas (2), multiple nodular acidophil adenoma (1), myoepitheloma (1) | GE HealthCare, Logiq S7 Expert | E-index | 3 | 7 | 15 | 3 | 29 |
| Malignant melanoma (3), salivary duct carcinoma (2), squamous cell carcinoma (1), adenoid cystic carcinoma (1), low-grade adenocarcinoma (1), mucoepidermoid carcinoma (1), Merkel cell carcinoma (1); Warthin tumour (18), plasmacytoma adenoma (8), reactive lymphoid hyperplasia (6), lipoma (3), granulomatous inflammation (2), basal cell adenoma (2), lymphoepithelial cyst (1), sialadenitis and abscess (1), sialadenosis (1), primary benign oncocytic neoplasm (1), Rosai-Dorfman disease (1) | Samsung Medison, Accuvix A30, RS 80 A | ECI | 3.5 | 15 | 5 | 1 | 42 |
| Mucoepidermoid carcinoma (5), malignant lymphoma (3), squamous cell carcinoma (2), acinic cell carcinoma (2), salivary duct carcinoma (2) and squamous cell carcinoma metastasis (2); pleomorphic adenoma (18), Warthin tumour (17), oncocytoma (5), lymphoepithelial cyst (3), ductal cyst (2), benign vascular tumour (2). | SuperSonic Imagine, Aixplorer | CSV | 1025 | 18 | 17 | 10 | 79 |
| Squamous cell carcinoma (8), low grade salivary tumour (6), high grade salivary tumour (7), lymphoma (3), melanoma (2), sarcoma (1), neuroendocrine carcinoma (1); plasmacytoma adenoma (49), Warthin tumour (33), other benign lesions (14) including oncocytic adenomas, lipomas, lipomatosis, basal cell adenoma, non-sebaceous lymphadenoma, branchiogenic cyst, and chronic inflammation | Siemens Healthcare, Acuson S2000 | ES (3-point) | ES2 | 22 | 126 | 10 | 44 |

Avg. average, FNAC fine needle aspiration cytology, SWV shear wave velocity, CSV coefficient of stiffness variability (maximum of stiffness/minimum of stiffness), CSES consensus sonoelastography scores, ES elastographic scoring, ECI elasticity contrast index, ARFI acoustic radiation force impulse, TP true positive, FN false negative, FP false positive, TN true negative, NA not available

*Missing data provided by corresponding author via e-mail

* Another study [25] from the same department was omitted
| QUADAS questionnaire                                                                 | Klintworth et al 2012 [8] | Ye et al 2012 [24] | Badea et al 2013 [16] | Celebi and Mahmutoglu 2013 [17] | Wierzbicka et al 2013 [19] | Yu et al 2016 [20] | Altunbas et al 2017 [15] | Cantisani et al 2017 [7] | Herman et al 2017 [6] | Mansour et al 2017 [18] |
|-------------------------------------------------------------------------------------|--------------------------|-------------------|------------------------|---------------------------------|--------------------------|------------------------|-------------------------|--------------------------|-------------------------|-------------------------|
| 1: Was the spectrum of patient representative of the patients who will receive the test in practice? | Yes                      | No                | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 2: Were selection criteria clearly described?                                         | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 3: Is the reference standard likely to correctly classify the target condition?        | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 4: Is the time period between reference standard and index test short enough to be sure that the target condition did not change between the two tests? | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 5: Did the whole sample, or a random selection of the sample, receive verification using a reference standard of diagnosis? | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 6: Did patients receive the same reference standard regardless of the index test result? | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 7: Was the reference standard independent of the index test (i.e., the index test did not form part of the reference standard)? | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 8: Was the execution of the index test described in sufficient detail to permit replication of the test? | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 9: Was the execution of the reference standard described in sufficient detail to permit replication? | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 10: Were the index test results interpreted without knowledge of the results of the reference standard? | No                       | Yes               | Yes                     | Yes                              | Yes                      | Unclear                | Yes                     | Yes                      | Yes                     | Yes                     |
| 11: Were the reference standard results interpreted without knowledge of the results of the index test? | Unclear                  | Unclear           | Unclear                 | Unclear                          | Unclear                  | Unclear                | Unclear                 | Unclear                  | Unclear                 | Unclear                 |
| 12: Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 13: Were un-interpretable/intermediate test results reported?                         | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| 14: Were withdrawals from the study explained?                                        | Yes                      | Yes               | Yes                     | Yes                              | Yes                      | Yes                    | Yes                     | Yes                      | Yes                     | Yes                     |
| QUADAS score                                                                         | 12.5                     | 13.5              | 12.5                    | 13.5                             | 13.5                     | 13                     | 13.5                    | 13.5                     | 13.5                    | 13.5                    |
Fig. 2 Forest plots of the pooled sensitivity (a) and specificity (b) of sonoelastography for differentiating between malignant and benign parotid lesions.

| Study                | Sensitivity (95% CI) | Specificity (95% CI) |
|----------------------|----------------------|----------------------|
| Klintworth et al. 2012 | 0.38 (0.09 - 0.76) | 0.96 (0.86 - 1.00)   |
| Yerli et al. 2012    | 0.67 (0.22 - 0.96)  | 0.67 (0.45 - 0.84)  |
| Badea et al. 2013   | 1.00 (0.63 - 1.00) | 0.67 (0.46 - 0.75)  |
| Celebi et al. 2013  | 0.59 (0.41 - 0.78)  | 0.61 (0.21 - 0.79)  |
| Wierzbicka et al. 2013 | 0.40 (0.12 - 0.74) | 0.97 (0.84 - 1.00)  |
| Yu et al. 2016      | 0.69 (0.41 - 0.89)  | 0.97 (0.65 - 1.00)  |
| Altinbas et al. 2017 | 0.70 (0.35 - 0.93) | 0.66 (0.50 - 0.80)  |
| Cantisani et al. 2017 | 0.73 (0.57 - 0.96) | 0.89 (0.77 - 0.96)  |
| Herman et al. 2017  | 0.64 (0.44 - 0.81)  | 0.82 (0.73 - 0.89)  |
| Mansour et al. 2017 | 0.69 (0.50 - 0.84)  | 0.26 (0.19 - 0.33)  |

Pooled Sensitivity = 0.67 (0.59 to 0.74)
Chi-square = 20.12; df = 9 (p = 0.0172)
Inconsistency (I-square) = 55.3%

Pooled Specificity = 0.64 (0.60 to 0.68)
Chi-square = 209.31; df = 9 (p = 0.0000)
Inconsistency (I-square) = 95.7%

Fig. 3 Summary receiver operating characteristic (SROC) curve on sonoelastography for differentiating between malignant and benign parotid lesions. The middle curve is the SROC curve. The upper and lower curves show the 95% confidence intervals.
affecting study heterogeneity ($p = 0.035$). However, the summary estimates did not differ between SE and SWE ($p = 0.119$) or between ARFI and SSI ($p = 0.473$).

Recently, several original studies have focused on the value of sonoelastography for differentiating between malignant and benign parotid lesions. Sonoelastography is a novel ultrasonographic technique for assessing tissue elasticity and stiffness. Theoretically, malignant parotid tumours should be stiffer than benign ones. However, the situation seems complicated. Some authors have described the great performance of sonoelastography for differentiating between malignant and benign lesions with a high sensitivity of 94% and a specificity of 89% [7]. Some have described a relatively lower but still clear performance of sonoelastography, with a sensitivity of 70% and specificity of 66% [15]. However, others have described that there was no benefit of sonoelastography for differentiating between malignant and benign tumours; only cystic lesions or cystic areas within a lesion were reliably identified [18]. Our meta-analysis ultimately revealed a pooled sensitivity of 67% and a pooled specificity of 64% for differentiating between malignant and benign parotid lesions. Therefore, we believe that the overall value of sonoelastography for differential diagnosis was limited and not satisfactory.

Heterogeneity was revealed in our study. Therefore, a meta-regression analysis was performed to further explore the potential sources. The results showed that there was no difference between SE and SWE or between ARFI and SSI ($p = 0.473$).

Table 4 Results of the meta-regression and subgroup analysis for differential diagnosis between malignant and benign parotid lesions

| Subgroup               | Number of studies | Pooled sensitivity (95% CI) | Pooled specificity (95% CI) | Pooled DOR (95% CI) | AUC | p value |
|------------------------|-------------------|-----------------------------|-----------------------------|---------------------|-----|---------|
| Mechanism              |                   |                             |                             |                     |     |         |
| SWE                    | 6                 | 0.61 (0.52–0.70)            | 0.62 (0.57–0.66)            | 6.39 (1.82–22.35)   | 0.67| 0.119   |
| SE                     | 3                 | 0.81 (0.64–0.93)            | 0.76 (0.67–0.83)            | 11.67 (1.56–87.41)  | 0.64|         |
| SWE technique          |                   |                             |                             |                     |     |         |
| ARFI                   | 4                 | 0.63 (0.52–0.73)            | 0.51 (0.45–0.57)            | 4.91 (0.96–25.05)   | 0.65|         |
| SSI                    | 2                 | 0.58 (0.41–0.74)            | 0.86 (0.79–0.92)            | 9.50 (3.99–22.63)   | NA  |         |
| Assessment method      |                   |                             |                             |                     |     |         |
| Qualitative            | 5                 | 0.59 (0.48–0.69)            | 0.52 (0.46–0.58)            | 3.38 (1.08–10.57)   | 0.63|         |
| Qualitative or         |                   |                             |                             |                     |    |         |
| semiquantitative       | 4                 | 0.73 (0.61–0.83)            | 0.83 (0.77–0.88)            | 18.64 (4.51–77.07)  | 0.88|         |
| QUADAS score           |                   |                             |                             |                     |    |         |
| 13.5                   | 7                 | 0.66 (0.58–0.74)            | 0.59 (0.54–0.63)            | 5.41 (1.81–16.13)   | 0.74|         |
| ≤ 13                   | 3                 | 0.69 (0.50–0.84)            | 0.91 (0.83–0.96)            | 26.56 (6.91–102.11) | 0.90|         |

SE strain elastography, SWE shear wave elastography, SSI supersonic shear imaging, ARFI acoustic radiation force impulse imaging, NA not available
*Meta-regression, $p < 0.05$

Fig. 4 Funnel plot for evaluating potential publication bias. Each solid circle represents a study in the meta-analysis. The line is the regression line.

Another potential source of heterogeneity might be the histopathological variety in malignant and benign parotid lesions. Celebi and Mahmutoglu [17] indicated that the diagnostic value of sonoelastography for evaluating pleomorphic adenomas, Warthin tumours, adenoid cystic carcinoma and high-grade tumours was low, whereas the diagnostic rates for low-grade tumours, such as mucoepidermoid carcinoma, acinic cell carcinoma and metastases of basal cell carcinoma, were better. Pleomorphic adenomas contained variable
proportions of chondroid and/or myxoid matrix, which contained different amounts of fluid. Warthin tumours contained different amounts of lymphatic, cellular, mucous and fluid components. Thus, the two types of benign tumours could be solid, solid and cystic, or completely cystic, which resulted in a wide variety in stiffness. In a small sample study of 20 patients with only pleomorphic adenomas included in the benign group, 50% (6/12) of the adenomas were misdiagnosed as malignancies [16]. We tried to analyse whether sonoelastography could differentiate between low-grade parotid tumours and high-grade and benign ones. We also tried to analyse the effects of the different components in pleomorphic adenomas and Warthin tumours on sonoelastography. However, both of these analyses were not accomplished because, in most of the studies, the data were not recorded.

A strict procedure was carried out to screen the articles and ultimately 10 relative studies were identified. Deeks’ funnel plots showed no significant publication bias. Most of the studies were high quality according to the QUADAS questionnaire. A meta-regression revealed that the QUADAS score was not a significant factor affecting study heterogeneity. However, the QUADAS score seemed to perform better in relatively lower quality studies, as shown in Table 4. In one study [20], it was unclear whether the observers knew the histopathological results before analysing the images. In another study, the observers were aware of histological properties before reviewing the images and videos [8]. These unblinded studies probably had better performance and influenced the results. In addition, in all the studies it was unclear whether the histopathology reviewer knew the results of sonoelastography evaluations, which probably caused heterogeneity and influenced the results as well. To the best of our knowledge, this is the first meta-analysis to assess the diagnostic value of sonoelastography merely for differentiating between malignant and benign parotid lesions, except for salivary gland masses [26].

There are some limitations in our study. First, relatively few studies were included (i.e. ten). Second, we failed to acquire unpublished data and language limitations might have affected the reliability of the results. Third, postoperative pathology was used as a reference standard for tumour detection in most of the studies in this meta-analysis; however, in one study [15], only cytological and histological results from ultrasound-guided fine needle aspiration biopsy were used as reference standards, and in another two studies, cytology results from ultrasound-guided fine needle aspiration were used in six cases [17] and two cases [24], respectively. Although cytology and histology of fine-needle aspiration biopsy are suggested diagnostic methods for most parotid tumours, these methods have variable success with sensitivity ranging from 57% to 98%, specificity ranging from 56% to 100% and accuracy ranging from 78% to 98% [7].

In conclusion, this meta-analysis shows that sonoelastography has a limited value for differential diagnosis between malignant and benign parotid lesions. Quantitative and semiquantitative methods performed better than qualitative ones. Further large-sample, prospective, multicentre studies evaluating these two assessment methods are needed to confirm the findings. In addition, more studies should focus on the correlation between sonoelastography and corresponding histopathological changes in the future.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Xue-Mei Wang.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors (Yun-Fei Cai) has significant statistical expertise.

Informed consent Written informed consent was not required for this study because of the nature of our study, which was a systemic review and meta-analysis.

Ethical approval Institutional review board approval was not required because of the nature of our study, which was a systemic review and meta-analysis.

Study subjects or cohorts overlap All the study subjects have been previously reported in literature because this is a systemic review and meta-analysis.

Methodology

• Systemic review
• Meta-analysis performed at one institution

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Białek EJ, Jakubowski W, Zajkowski P, Szopinski KT, Osmolski A (2006) US of the major salivary glands: anatomy and spatial relationships, pathologic conditions, pitfalls. Radiographics 26: 745–763
2. Habermann CR, Arndt C, Graessner J et al (2009) Diffusion-weighted echo-planar MR imaging of primary parotid gland tumors: is a prediction of different histologic subtypes possible? AJNR Am J Neuroradiol 30:591–596
3. Bozzato A, Zenk J, Greess H et al (2007) Potential of ultrasound diagnosis for parotid tumors: analysis of qualitative and quantitative parameters. Otolaryngol Head Neck Surg 137:642–646
4. Bhatia KS, Rasalkar DD, Lee YP et al (2010) Evaluation of real-time qualitative sonoelastography of focal lesions in the parotid and submandibular glands: Applications and limitations. Eur Radiol 20:1958–1964
5. Lee YY, Wong KT, King AD, Ahuja AT (2008) Imaging of salivary gland tumours. Eur J Radiol 66:419–436
6. Herman J, Sedlackova Z, Vachutka J et al (2017) Differential diagnosis of parotid gland tumors: role of shear wave elastography. Biomed Res Int 2017:9234672
7. Cantisani V, David E, De Virgilio A et al (2017) Prospective evaluation of quasi-static ultrasound elastography (USE) compared with baseline US for parotid gland lesions: preliminary results of elasticity contrast index (ECI) evaluation. Med Ultrasound 19:32–38
8. Klintworth N, Mantsopoulos K, Zenk J, Psychogios G, Iro H, Bozzato A (2012) Sonoelastography of parotid gland tumours: Initial experience and identification of characteristic patterns. Eur Radiol 22:947–956
9. Gietka-Czernel M, Kochman M, Bujalska K, Stachlewka-Nasfeter E, Zgliczynski W (2010) Real-time ultrasound elastography - a new tool for diagnosing thyroid nodules. Endokrynol Pol 61:652–657
10. Bhatia KS, Tong CS, Cho CC, Yuen EH, Lee YY, Ahuja AT (2012) Shear wave elastography of thyroid nodules in routine clinical practice: preliminary observations and utility for detecting malignancy. Eur Radiol 22:2397–2406
11. Evans A, Whelehan P, Thomson K et al (2012) Differentiating benign from malignant solid breast masses: value of shear wave elastography according to lesion stiffness combined with greyscale ultrasound according to BI-RADS classification. Br J Cancer 107:224–229
12. Pinto F, Totaro A, Calarco A et al (2011) Imaging in prostate cancer diagnosis: present role and future perspectives. Urol Int 86:373–382
13. Seitz M, Strittmatter F, Roosen A, Tilki D, Gratzeke C (2010) Current status of ultrasound imaging in prostate cancer. Panminerva Med 52:189–194
14. Mahdavi SS, Moradi M, Wen X, Morris WJ, Salcudean SE (2011) Evaluation of visualization of the prostate gland in vibro-elastography images. Med Image Anal 15:589–600
15. Altinbas NK, Anamurluoglu EG, Oz II et al (2017) Real-time sonoelastography of parotid gland tumors. J Ultrasound Med 36:77–87
16. Badea AF, Bran S, Tamas-Szora A, Floare A, Badea R, Baciat G (2013) Solid parotid tumors: an individual and integrative analysis of various ultrasonographic criteria. A prospective and observational study. Med Ultrason 15:289–298
17. Celebi I, Mahmutoglu AS (2013) Early results of real-time qualitative sonoelastography in the evaluation of parotid gland masses: A study with histopathological correlation. Acta Radiol 54:35–41
18. Mansour N, Bas M, Stock KF, Strassen U, Hofauer B, Knopf A (2017) Multimodal ultrasonographic pathway of parotid gland lesions. Ultraschall Med 38:166–173
19. Wierzbicka M, Kahuzy J, Szczepanek-Panulska E et al (2013) Is sonoelastography a helpful method for evaluation of parotid tumors? Eur Arch Otorhinolaryngol 270:2101–2107
20. Yu J, Du Y, Lu Y et al (2016) Application of DTI and ARFI imaging in differential diagnosis of parotid tumours. Dentomaxillofac Radiol 45:20160100
21. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6:e1000097
22. Liberati A, Altman DG, Tetzlaff J et al (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 339:b2700
23. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 3:25
24. Yerli H, Eski E, Korucuk E, Kaskati T, Agildere AM (2012) Sonoelastographic qualitative analysis for management of salivary gland masses. J Ultrasound Med 31:1083–1089
25. Mansour N, Stock KF, Chaker A, Bas M, Knopf A (2012) Evaluation of parotid gland lesions with standard ultrasound, color duplex sonography, sonoelastography, and acoustic radiation force impulse imaging - a pilot study. Ultraschall Med 33:283–288
26. Li C, Zhang C, Li N, Li J (2016) Compression real-time elastography for evaluation of salivary gland lesions: a meta-analysis. J Ultrasound Med 35:999–1007