Accuracy of diagnosis of salivary gland tumors with the use of ultrasonography, computed tomography, and magnetic resonance imaging: a meta-analysis

Ying Liu, MD,[a] Jia Li, MD,[b] Yi-ran Tan, MD,[a] Ping Xiong, MD,[b] and Lai-ping Zhong, MD, PhD[a]

Objective. To compare ultrasonography (US), computed tomography (CT), and magnetic resonance imaging (MRI) for clinical differential diagnosis in patients with salivary gland tumor (SGT).

Study Design. Six databases were used to search the literature published between 1982 and 2013. Histologic diagnosis was required as standard diagnosis. Pooled estimate for sensitivity, specificity, summary receiver-operating characteristic curve (SROC) and area under curve (AUC) were calculated and compared using STATA and Meta-Disc statistical software.

Results. Nineteen articles were included. Pooled sensitivity for US, CT, and MRI was 0.629 (95% confidence interval [CI] 0.52-0.73), 0.830 (95% CI 0.74-0.90), and 0.807 (95% CI 0.73-0.87), respectively; pooled specificity for US, CT, and MRI was 0.920 (95% CI 0.89-0.94), 0.851 (95% CI 0.79-0.90), and 0.886 (95% CI 0.85-0.92), respectively. The AUC under SROC for US, CT, and MRI was 0.934 ± 0.058, 0.912 ± 0.889, and 0.903 ± 0.045, respectively.

Conclusions. CT is recommended, as it is an effective imaging tool for differential diagnosis in patients with primary SGT, and MRI is suggested for differential diagnosis between benign and malignant GSTs because of its highest sensitivity and specificity. (Oral Surg Oral Med Oral Pathol Oral Radiol 2015;119:238-245)

Salivary gland tumors (SGTs) account for about 3% of head and neck tumors.[1] SGTs are clinically asymptomatic until they grow to a great volume or involve adjacent structures, such as nerves, ducts, or muscles. SGTs occur mostly in the parotid, submandibular, and sublingual glands. When SGTs are located superficially, they are usually easy to find; however, when the tumor is deep or at an early stage, it might be difficult to identify. Some imaging examinations, such as ultrasonography (US), computed tomography (CT), and magnetic resonance imaging (MRI), are necessary and are helpful for clinical diagnosis.[2] Although fine-needle aspiration biopsy (FNAB) is the most definitive tool to determine whether the lesion is benign or malignant, it is sometimes difficult to perform due to unusual location of the tumor or patients' unwillingness to undergo FNAB. In addition, FNAB is a more invasive procedure that usually requires local anesthesia as well as CT or US guidance.[3] FNAB could also modify the tumor structures and cause necrosis, hemorrhage, fibrosis, and squamous metaplasia thereby making the subsequent histologic evaluation more difficult.[4,5] The accuracy of the evaluation depends on the quality of the sample (quantity of tissue; avoidance of nonspecific areas, such as cystic changes or necrosis) and the pathologist's experience.[6] When FNAB is unavailable, imaging examination is helpful for establishing the clinical diagnosis and making the treatment plan.

The most common benign SGTs are pleomorphic adenoma, adenolymphoma, basal cell adenoma, oxyphilic adenoma, myoepithelioma, and papillary cystadenoma.[7] The most common malignant SGTs are adenoid cystic carcinoma, mucoepidermoid carcinoma, acinic cell carcinoma, and adenocarcinoma.[8] The common characteristics of benign SGTs delineated by CT and MRI are sharp margins, round shape, and uniform distribution of density; other characteristics of benign SGTs seen on MRI include a low-density signal with T1-weighted images and a high-density signal with T2-weighted images. The common characteristics of malignant SGTs seen on CT and MRI are irregularity and intraglandular extension.[9,10] Gadolinium-enhanced dynamic MRI and diffusion-weighted echo-planar imaging MRI with apparent diffusion coefficient

Statement of Clinical Relevance

Imaging examinations are very helpful for clinical diagnosis when fine-needle aspiration biopsy is difficult to perform due to unusual tumor location or patients' unwillingness. Ultrasonography (US), computed tomography (CT), and magnetic resonance imaging (MRI) are reliable methods for diagnosing salivary gland tumors (SGTs) clinically.

This study was supported by research grants from the National Natural Science Foundation of China (No. 81272979) and the Science and Technology Commission of Shanghai Municipality (No. 13 QH1401700).

Dr. Ying Liu and Dr. Jia Li contributed equally to this paper.

[a]Department of Oral and Maxillofacial—Head and Neck Oncology, Shanghai, China.

[b]Department of Ultrasound, Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

Received for publication Nov 17, 2013; returned for revision Jul 2, 2014; accepted for publication Oct 29, 2014.

2212-4403

http://dx.doi.org/10.1016/j.oooo.2014.10.020

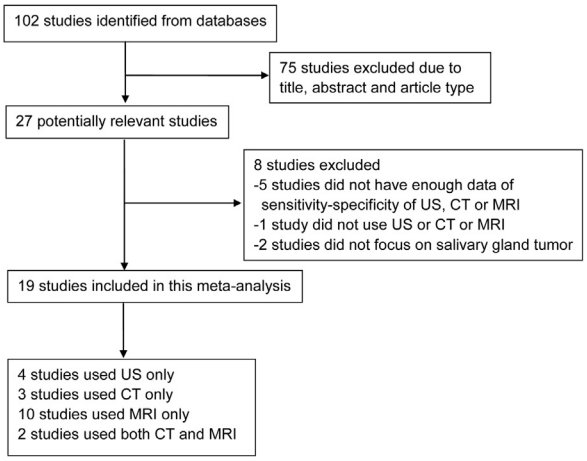


Fig. 1. Flowchart of articles included in this meta-analysis.

evaluation could both improve the effectiveness of MRI in distinguishing between benign and malignant parotid gland tumors.[11] The common US characteristics of parotid masses include shape, margin, echogenicity, echotexture, and vascularization. Some studies focus on the different criteria of these US characteristics for differential diagnosis of parotid tumors; for example, B-mode sonography and elastographic sonography have been investigated on the basis of these characteristics to differentiate between benign and malignant parotid tumors.[12] However, it is sometimes difficult to differentiate malignant SGTs from benign SGTs.

In this meta-analysis, we assessed the diagnostic capability of US, CT, and MRI and compared these findings with the standard pathologic results, with the aim of identifying the best imaging modality for diagnostic accuracy in SGT.

METHODS

Literature search

Five databases, including Embase, Pubmed, Springerlink, Sciencedirect, and Cochrane library databases, were searched for publications from September 1982 to April 2013. The data used were limited to those officially published in English. Key words included "salivary gland," "parotid gland," "submandibular gland," "sublingual gland," "salivary ducts," or "von Ebner glands"; "US," "ultrasound," "ultrasonography," "ultrasonic diagnosis," "CT," "computed tomography," "computerized tomography," "MR," "MRI," or "magnetic resonance imaging"; and "sensitivity," "specificity," or "accuracy". The article search steps are shown in Figure 1. All articles were required to have lesion origin, pathologic diagnosis, study type, and one of US, CT, or MRI results. True positive (TP), false positive (FP), true negative (TN), and false negative (FN) diagnostic results in differentiating malignant and benign tumor were also required to be reported in the articles. This study was exempt from approval by the ethics committee of the Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine.

Inclusion and exclusion criteria

The inclusion criteria were histologic diagnosis as final diagnosis, detailed description of each image examination, and specific regulation in differentiating malignant SGTs from benign SGTs. The exclusion criteria were study type being a review, case report, commentary, editorial, or outcome without raw data.

Data extraction

All data were extracted by two authors independently, and any lack of clarity or disagreement was resolved through discussion. The following items were deemed essential: description of population, such as age and gender ratios, publication year, study type, lesion number and location, study design, and imaging analysis related to our research. FP, TP, FN, and TN ratios were also recorded. A standard form was designed and followed to select potentially qualified articles. During data extraction, the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool was used as a guide line.[13] The QUADAS tool included 10 items to assess for risk of bias, source of variation, and reporting quality. The answer to each item was "yes," "no," or "unclear." When the answer was "yes," the item scored one point; when the answer was "no," the item scored minus one point; when the answer was "unclear," the item scored zero. The QUADAS chart is shown in Supplementary Figure S1. When the final score was higher than 7, the quality of the article was considered high; when the final score was 6 or 7, the quality of the article was considered medium; when the final score was less than 6, the quality of the article was considered low.

Data analysis

Before merging raw data into the software, the likelihood ratio (I^2) index and Cochran Q test were used to quantify the heterogeneity of the enrolled articles. The percentage measure of the heterogeneity among the enrolled articles was calculated as I^2 index. When I^2 was greater than 25%, the random effects model was used to summarize the result of sensitivity; when I^2 was less than 25%, which meant little heterogeneity in the enrolled articles, the fixed effects model was used for data analysis. When using the Cochran Q test for likelihood ratio, if the P value was less than .05, the articles were deemed heterogeneous. Threshold effect was estimated by using the Meta-Disc software to evaluate the possible factors causing the heterogeneity in combining individual statistical data. The correlation

Table I. Summary of patient characteristics

References	Country (publish year)	Patient number	Study design	Male:Female	Mean age (years)	Measurement (US = 1, CT = 2, MRI = 3)
Eida et al.[14]	Japan (2007)	31	Unknown	1:1.4	63	3
Motoori et al.[15]	Japan (2005)	33	Unknown	1:0.3	60.8	3
Kurabayashi et al.[16]	Japan (2002)	30	Unknown	1:1.1	43.1	3
Takashima et al.[17]	Japan (2001)	72	Prospective	1:1.1	53	3
Takashima et al.[18]	Japan (1997)	53	Prospective	1:1.1	53	3
Inohara et al.[19]	Japan (2008)	81	Unknown	Unknown	Unknown	3
Arbab et al.[20]	Japan (2000)	22	Retrospective	1:1.4	Unknown	2, 3
Klintworth et al.[21]	Germany (2012)	57	Retrospective	1:1.1	53.3	1
Wu et al.[22]	China (2012)	189	Retrospective	1:1.1	42.3	1
Jin et al.[23]	China (2011)	51	Unknown	1:0.8	44	2
Lechner Goyault et al.[24]	France (2011)	60	Retrospective	1:0.9	59.4	3
Paris et al.[25]	France (2005)	86	Retrospective	Unknown	Unknown	3
Takashima et al.[26]	Japan (1999)	26	Unknown	1:0.5	56	3
Corr et al.[27]	Hong Kong (1993)	40	Prospective	Unknown	Unknown	1
Kim et al.[28]	South Korea (1998)	147	Retrospective	Unknown	Unknown	2, 3
Yabuuchi et al.[29]	Japan (2003)	42	Prospective	Unknown	Unknown	3
Gritzmann et al.[30]	Austria (1989)	289	Retrospective	Unknown	Unknown	1
Bryan et al.[31]	America (1982)	27	Retrospective	Unknown	Unknown	2
Park et al.[32]	Korea (2012)	67	Retrospective	1:0.4	61.1	2

US, ultrasonography; *CT*, computed tomography; *MRI*, magnetic resonance imaging.

coefficients of logit sensitivity and logit (1-specificity) were also calculated. When there was a positive correlation, which indicated a threshold effect, summary receiver-operating characteristic curve (SROC) and area under curve (AUC) were calculated. When there was a negative correlation, subgroup analysis was performed. Spearman correlation coefficient and *P* value were calculated for symmetry of SROC. When *P* was greater than .05, the Mantel-Haenszel model as well as both the DerSimonia-Laird and Moses-Shapiro-Littenber models were used to calculate diagnostic odds ratio (DOR) and SROC; when *P* was less than .05, the Moses-Shapiro-Littenber model was used.[14]

Sensitivity was calculated as TP/(FN+TP), specificity was calculated as TN/(FP+TN), and 95% confidence interval (CI) was also estimated; when calculating sensitivity and specificity for each article, all lesions were included. SROC was used to evaluate the overall diagnosis performance of determined groups. AUC was compared by using the Mann-Whitney U test. Q value was used to represent a global measure of test accuracy.[15] The DOR of US, CT, and MRI was calculated to illustrate positive likelihood ratio over negative likelihood. Meta-regression was used to test the potential source of heterogeneity, which was considered significant when the *P* value was less than .1. Publication bias was presented using a funnel plot, and Egger regression test was used to examine the asymmetry of the funnel.

Statistical analysis was performed with STATA statistical software (Version 11.0, StataCorp LP, College Station, TX) and Meta-Disc software (Version 1.4, Madrid, Spain). When the *P* value was less than .05, the difference was considered statistically significant.

RESULTS

Literature evaluation

One hundred and two articles were identified in the literature databases, and 73 articles were excluded after reading their abstracts. According to the inclusion and exclusion criteria, 10 articles were excluded, and only 19 articles could be used for analysis,[16-32] as described in detail in Figure 1. With the QUADAS tool, 8 articles were evaluated as high-quality articles, 10 articles were deemed medium quality, and only 1 article was of low quality. There were 784 patients with 792 SGTs enrolled in this analysis. The male-to-female ratio was 1:1.05. The patients' ages ranged from 42 to 63 years, with a mean of 52.4 ± 7.9 years. There were 12 articles evaluating MRI, 5 articles evaluating CT, and 4 articles evaluating US (Table I).

Publication bias and heterogeneity

Because there were only 5 and 4 articles evaluating CT and US, respectively, the sample size was too small for statistical analysis when the funnel plot was used to test diagnostic effect; 12 articles evaluating MRI were used to test diagnostic effect using the funnel plot. Information from each patient was incorporated into the funnel plot, the x-axis was the DOR and the y-axis was the inverse of the effective sample size (1/ESS). Consequently, a regression line and a significant regression coefficient (−13.39; 95% CI = −47.62-20.83; $P = .393$) could be obtained, and the funnel plot was symmetric (Supplementary Figure S2). Meta-regression was used to analyze the relationship between the DOR and the composite variables; unfortunately, no significant relationship was found ($P > .05$). The Spearman correlation coefficients for MRI, CT, and US

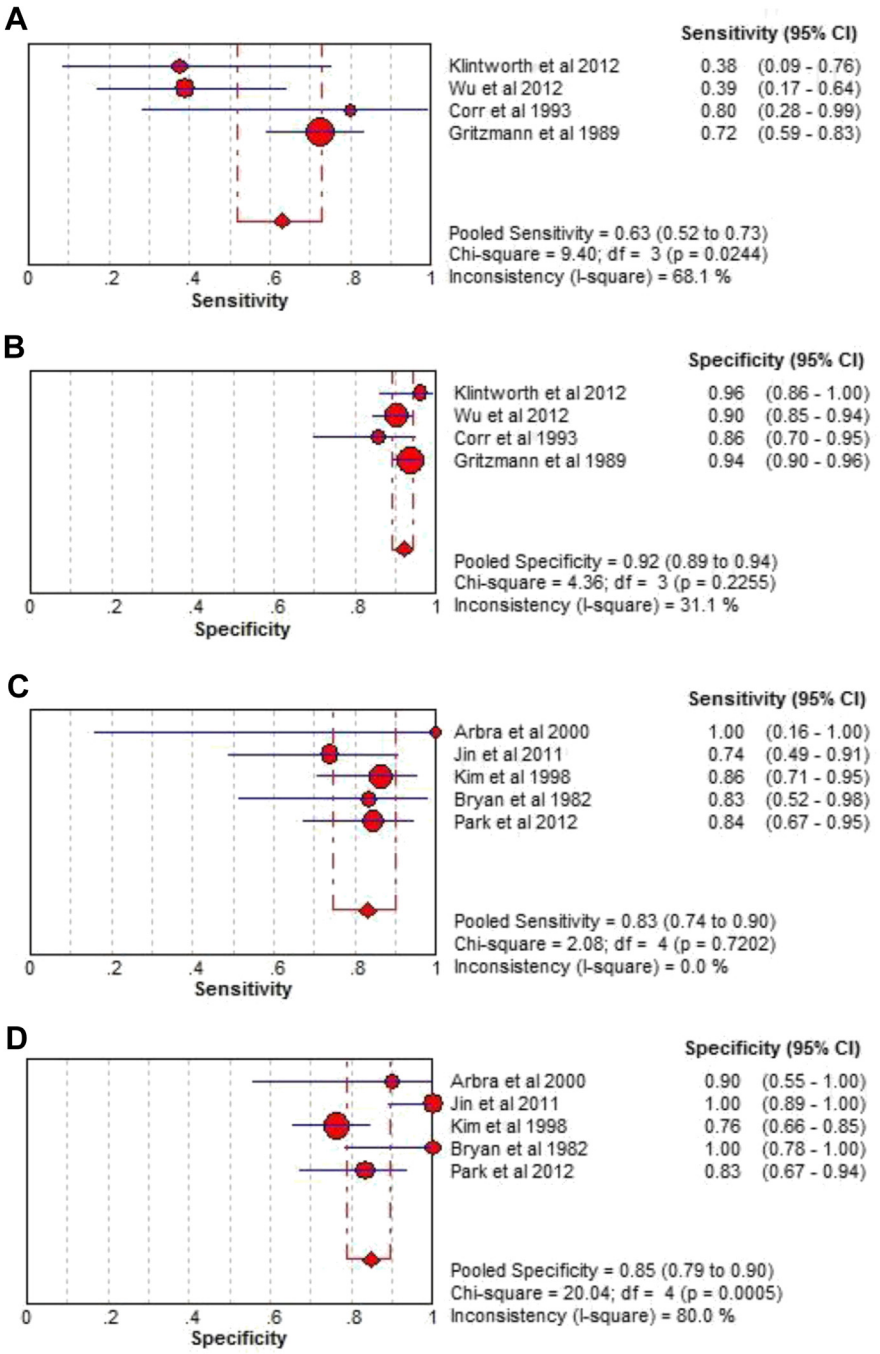


Fig. 2. Forest plot (random effects model) of pooled sensitivity and specificity for differential diagnosis between benign and malignant salivary gland tumors with ultrasonography (**A**, **B**), computed tomography (**C**, **D**), and magnetic resonance imaging (**E**, **F**), respectively.

were −0.27 ($P = .397$), 1 ($P < .001$), and 0.800 ($P = .200$), respectively.

Diagnostic sensitivity and specificity of ultrasonography

When US was used to differentiate malignant SGTs from benign SGTs, for sensitivity calculation, the I^2 index was 68.1%, and the Cochran Q test was 9.4 (df = 3; $P = .024$); a random effects model was used, with a pooled sensitivity of 63% (95% CI 52%-73%). For specificity calculation, the I^2 index was 31.1%, and the Cochran Q test was 92.0 (df = 3; $P = .225$); a fixed effects model was used, with a pooled specificity of 92% (95% CI 89%-94%) (Figure 2, A and B).

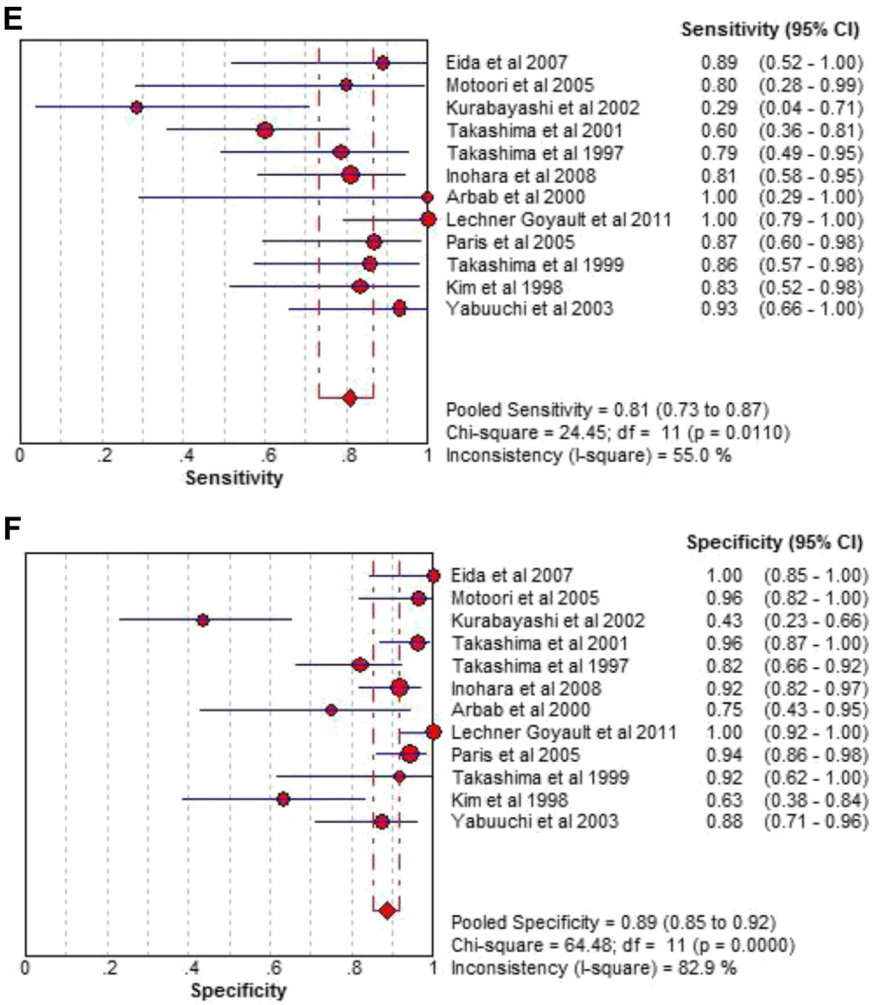


Fig. 2. (*continued*).

Diagnostic sensitivity and specificity of computed tomography

For calculation of the sensitivity of CT, the I^2 index was 0, and the Cochran Q test was 2.1 (df = 4; $P = .720$); a fixed effects model was used, with a pooled sensitivity of 83% (95% CI 74%-90%). For specificity calculation, the I^2 index was 80%, and the Cochran Q test was 20.4 (df = 4; $P < .001$); a random effects model was used, with a pooled specificity of 85% (95% CI 79%-90%) (see Figure 2, C and D).

Diagnostic sensitivity and specificity of magnetic resonance imaging

For calculation of the sensitivity of MRI, the I^2 index was 55.0%, and the Cochran Q test was 24.45 (df = 11; $P = .011$); a random effects model was used, with a pooled sensitivity of 81% (95% CI 73%-87%). For specificity calculation, the I^2 index was 82.9%, and the Cochran Q test was 64.5 (df = 11; $P < .001$); a random effects model was used, with a pooled specificity of 89% (95% CI 85%-92%) (see Figure 2, E and F).

Area under curve and diagnostic odds ratio

For US, the AUC under SROC was 0.934 ± 0.058, and the Q index was 0.870 ± 0.072 (Figure 3, A). For CT, the AUC under SROC was 0.912 ± 0.889, and the Q index was 0.844 ± 0.025 (see Figure 3, B). For MRI, the AUC under SROC was 0.903 ± 0.045, and the Q index was 0.834 ± 0.049 (see Figure 3, C). The pooled DORs for US, CT, and MRI were 16.46 (95% CI 5.40-50.15; $P = .048$), 28.81 (95% CI 13.58-61.12; $P = .590$), and 34.94 (95% CI 11.08-110.24; $P < .001$), respectively. There was no significant difference among these three groups (Supplementary Figure S3).

According to the SROC analysis, there was no significant difference among these three groups. However, there was statistical difference in sensitivity between the US and CT modalities (P value = .027, Kruskal-Wallis test) and between the US and MRI modalities (P value = .045, Kruskal-Wallis test). The pooled sensitivity of CT and MRI was higher than that of US for clinical diagnosis of SGTs.

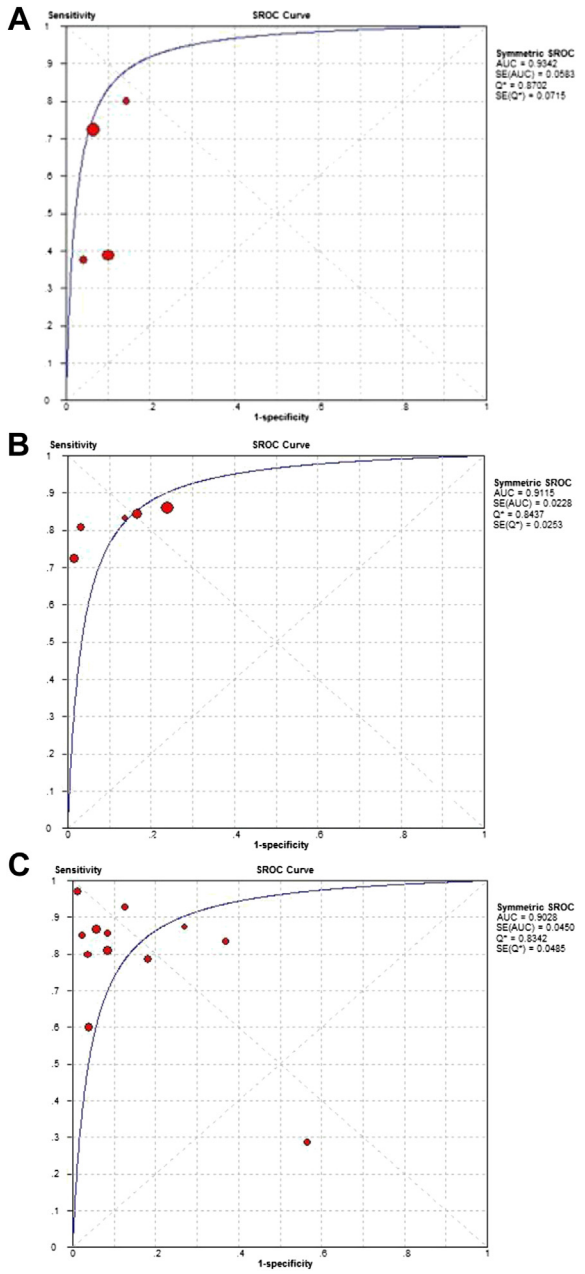


Fig. 3. Summary receiver-operating characteristic curve for differential diagnosis of between benign and malignant salivary gland tumors by ultrasonography (**A**), computed tomography (**B**), and magnetic resonance imaging (**C**). SROC, summary receiver operating characteristic; AUC, area under the curve; Q*, point at which sensitivity and specificity was equal.

DISCUSSION

In this study, we obtained the SROC for the diagnostic accuracy of the US, CT, and MRI modalities in patients with SGTs. AUC was considered the critical standard in judging diagnostic performance, and there was no statistical difference of AUC among the US, CT, and MRI modalities. From a forest map of sensitivity and specificity, there was high specificity but relatively poor sensitivity in the US modality; however, the combination of specificity and sensitivity in the MRI modality was the highest among the three modalities.

Imaging examination is very important for clinical diagnosis in patients with SGTs when FNAB is difficult to perform because of unusual location or patients' unwillingness to undergo FNAB. Early studies, in which the diagnostic criteria remained mostly consistent in each detection procedure, reported US to have high sensitivity.[17,18,27,30,33] With the new index applied in the detection procedure, diagnostic results varied greatly. For example, color Doppler flow imaging is an important tool in making a sufficiently definite diagnosis; however, the information on blood supply could not predict significant differences between benign and malignant SGTs.[34] Meanwhile, gray-scale sonographic images are effective features to calculate the properties of SGTs, and B-mode ultrasonography and sonoelastography could improve the diagnostic performance.[21,35] The specificity of US is generally good because the majority of SGTs are benign and only a small amount of SGTs are malignant (9.5%).[36] During the diagnostic procedure with US in patients with SGTs, some characteristics, such as lesion size, echogenicity, margin regularity, and vascularity, should be taken into consideration; furthermore, clinical data, such as medical history, speed of growth, pain, and facial palsy, should also be considered. For some cases, such as a large mass in a deep lobe of the salivary gland, differential diagnosis is difficult with US. In such situations, other imaging examinations, such as CT and MRI, might be helpful.

CT and MRI are commonly used as imaging diagnostic methods.[16,17,23,27,37] Unfortunately, neither of them has been shown to reach the ideal AUC achieved by US. The advantages of CT and MRI are significant, and they continue to play an important role in the management of patients with SGTs. CT, with its good anatomic resolution, soft tissue contrast, and detailed morphology, can provide meaningful information to surgeons during the procedure.[23,28,31] MRI, with its good spatial and contrast resolution and avoidance of radiation exposure and interfering factors, such as imaging parameters and iron accumulation, could also provide useful information.[16,17] The disadvantages of CT and MRI include time and monetary costs; for patients with an allergenic constitution and kidney dysfunction, use of the contrast agent is inappropriate.[38] Furthermore, some parents are uncomfortable about the radiation exposure to their children during CT scanning.

There are some limitations in this study when selecting studies, because a few studies come from

non-English articles; evaluation standard is varies in different clinical researches; radiologists from different hospitals may have different opinions on diagnosing periods. So, further clinical studies are encouraged to give more powerful results by improving inclusion and diagnostic criteria.

Although there is no significant difference with regard to diagnostic accuracy between benign and malignant SGTs in the present study, it is recommended that CT or MRI could be used as a recommended examination method in patients with SGTs with clinical symptoms, such as pain, rapid growth, and facial paralysis. Since CT and MRI have good sensitivity, they provide useful anatomic information to surgeons and for nonsurgical treatment planning.[15,21]

CONCLUSIONS

US, CT, and MRI are reliable methods in diagnosing SGTs clinically. There is no statistical difference between CT and MRI; however, MRI is more expensive than CT. CT is recommended as an effective imaging tool in patients with primary SGTs; MRI is also recommended for its highest sensitivity and specificity for differential diagnosis between benign and malignant SGTs.

We thank Dr. Jian-feng Luo for providing statistical support and editing the manuscript.

REFERENCES

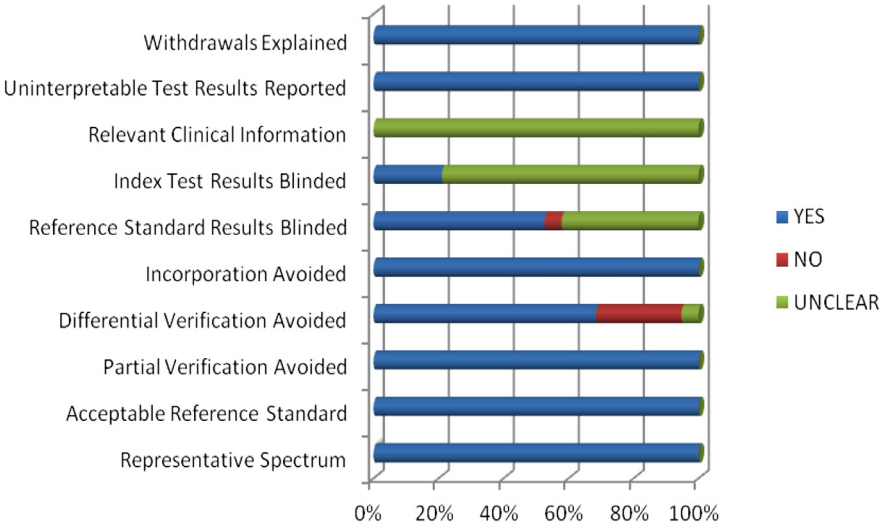
1. Lee YY, Wong KT, King AD, Ahuja AT. Imaging of salivary gland tumours. *Eur J Radiol*. 2008;66:419-436.
2. Rudack C, Jorg S, Kloska S, Stoll W, Thiede O. Neither MRI, CT nor US is superior to diagnose tumors in the salivary glands—an extended case study. *Head Face Med*. 2007;3:19.
3. Sherman PM, Yousem DM, Loevner LA. CT-guided aspirations in the head and neck: assessment of the first 216 cases. *AJNR Am J Neuroradiol*. 2004;25:1603-1607.
4. Maier H, Fruhwald S, Sommer S, Tisch M. Can pre-operative fine-needle aspiration of parotid tumors pose problems for a definitive histologic diagnosis? *Head Neck Oncol*. 2006;54: 166-170.
5. Mukunyadzi P, Bardales RH, Palmer HE, Stanley MW. Tissue effects of salivary gland fine-needle aspiration. Does this procedure preclude accurate histologic diagnosis? *Am J Clin Pathol*. 2000;114:741-745.
6. Balakrishnan K, Castling B, McMahon J, et al. Fine needle aspiration cytology in the management of a parotid mass: a two-centre retrospective study. *Surgeon*. 2005;3:67-72.
7. Ochicha O, Malami S, Mohammed A, Atanda A. A histopathologic study of salivary gland tumors in Kano, northern Nigeria. *Indian J Pathol Microbiol*. 2009;52:473-476.
8. Shishegar M, Ashraf MJ, Azarpira N, Khademi B, Hashemi B, Ashrafi A. Salivary gland tumors in maxillofacial region: a retrospective study of 130 cases in a southern Iranian population. *Pathol Res Int*. 2011;2011:9343-9350.
9. Casselman JW, Mancuso AA. Major salivary gland masses: comparison of MR imaging and CT. *Radiology*. 1987;165:183-189.
10. Weber AL. Imaging of the salivary glands. *Curr Opin Radiol*. 1992;4:117-122.
11. Whiting P, Rutjes AW, Reitsma JB, Bossuty PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003;3:25.
12. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med*. 1993;12:1293-1316.
13. Walter SD. Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. *Stat Med*. 2002;21:1237-1256.
14. Eida S, Sumi M, Sakihama N, Takahashi H, Nakamura T. Apparent diffusion coefficient mapping of salivary gland tumors: prediction of the benignancy and malignancy. *AJNR Am J Neuroradiol*. 2007;28:116-121.
15. Motoori K, Ueda T, Uchida Y, Chazono H, Suzuki H, Ito H. Identification of Warthin tumor: magnetic resonance imaging versus salivary scintigraphy with technetium-99 m pertechnetate. *J Comput Assist Tomogr*. 2005;29:506-512.
16. Kurabayashi T, Ida M, Tetsumura A, Ohbayashi N, Yasumoto M, Sasaki T. MR imaging of benign and malignant lesions in the buccal space. *Dentomaxillofac Radiol*. 2002;31:344-349.
17. Takashima S, Wang J, Takayama F, et al. Parotid masses: prediction of malignancy using magnetization transfer and MR imaging findings. *AJR Am J Roentgenol*. 2001;176:1577-1584.
18. Takashima S, Sone S, Takayama F, et al. Assessment of parotid masses: which MR pulse sequences are optimal? *Eur J Radiol*. 1997;24:206-215.
19. Inohara H, Akahani S, Yamamoto Y, et al. The role of fine-needle aspiration cytology and magnetic resonance imaging in the management of parotid mass lesions. *Acta Otolaryngol*. 2008;128:1152-1158.
20. Arbab AS, Koizumi K, Toyama K, et al. Various imaging modalities for the detection of salivary gland lesions: the advantages of 201 Tl SPET. *Nucl Med Commun*. 2000;21:277-284.
21. Klintworth N, Mantsopoulos K, Zenk J, Psychogios G, Iro H, Bozzato A. Sonoelastography of parotid gland tumours: initial experience and identification of characteristic patterns. *Eur Radiol*. 2012;22:947-956.
22. Wu S, Liu G, Chen R, Guan Y. Role of ultrasound in the assessment of benignity and malignancy of parotid masses. *Dentomaxillofac Radiol*. 2012;41:131-135.
23. Jin GQ, Su DK, Xie D, Zhao W, Liu LD, Zhu XN. Distinguishing benign from malignant parotid gland tumours: low-dose multi-phasic CT protocol with 5-minute delay. *Eur Radiol*. 2011;21:1692-1698.
24. Lechner Goyault J, Riehm S, Neuville A, Gentine A, Veillon F. Interest of diffusion-weighted and gadolinium-enhanced dynamic MR sequences for the diagnosis of parotid gland tumors. *J Neuroradiol*. 2011;38:77-89.
25. Paris J, Facon F, Pascal T, Chrestian MA, Moulin G, Zanaret M. Preoperative diagnostic values of fine-needle cytology and MRI in parotid gland tumors. *Eur Arch Otorhinolaryngol*. 2005;262:27-31.
26. Takashima S, Takayama F, Wang Q, Kurozumi M, Sekiyama Y, Sone S. Parotid gland lesions: diagnosis of malignancy with MRI and flow cytometric DNA analysis and cytology in fine-needle aspiration biopsy. *Head Neck*. 1999;21:43-51.
27. Corr P, Cheng P, Metreweli C. The role of ultrasound and computed tomography in the evaluation of parotid masses. *Australas Radiol*. 1993;37:195-197.
28. Kim KH, Sung MW, Yun JB, et al. The significance of CT scan or MRI in the evaluation of salivary gland tumors. *Auris Nasus Larynx*. 1998;25:397-402.
29. Yabuuchi H, Fukuya T, Tajima T, Hachitanda Y, Tomita K, Koga M. Salivary gland tumors: diagnostic value of gadolinium-enhanced dynamic MR imaging with histopathologic correlation. *Radiology*. 2003;226:345-354.

30. Gritzmann N. Sonography of the salivary glands. *AJR Am J Roentgenol*. 1989;153:161-166.
31. Bryan RN, Miller RH, Ferreyro RI, Sessions RB. Computed tomography of the major salivary glands. *AJR Am J Roentgenol*. 1982;139:547-554.
32. Park SB, Choi JY, Lee EJ, et al. Diagnostic criteria on 18 F-FDG PET/CT for differentiating benign from malignant focal hypermetabolic lesions of parotid gland. *Nucl Med Mol Imaging*. 2012;46:95-101.
33. Wittich GR, Scheible WF, Hajek PC. Ultrasonography of the salivary glands. *Radiol Clin North Am*. 1985;23:29-37.
34. Bradley MJ, Durham LH, Lancer JM. The role of colour flow Doppler in the investigation of the salivary gland tumour. *Clin Radiol*. 2000;55:759-762.
35. Yonetsu K, Ohki M, Kumazawa S, Eida S, Sumi M, Nakamura T. Parotid tumors: differentiation of benign and malignant tumors with quantitative sonographic analyses. *Ultrasound Med Biol*. 2004;30:567-574.
36. Li LJ, Li Y, Wen YM, Liu H, Zhao HW. Clinical analysis of salivary gland tumor cases in West China in past 50 years. *Oral Oncol*. 2008;44:187-192.
37. Burke CJ, Thomas RH, Howlett D. Imaging the major salivary glands. *Br J Oral Maxillofac Surg*. 2011;49:261-269.
38. Hasebroock KM, Serkova NJ. Toxicity of MRI and CT contrast agents. *Expert Opin Drug Metab Toxicol*. 2009;5:403-416.

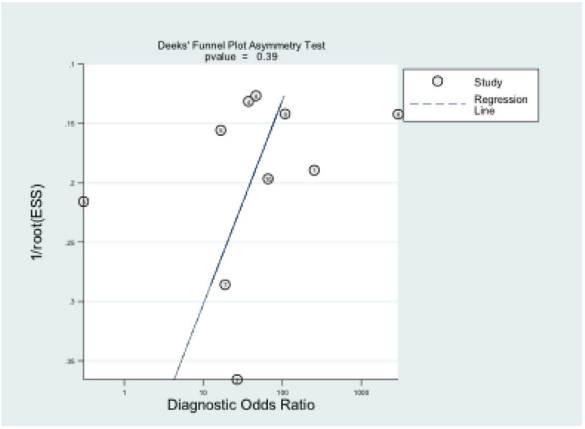
Reprint requests:

Lai-ping Zhong, MD, PhD
Department of Oral and Maxillofacial−Head and Neck Oncology
Zhonglp@hotmail.com

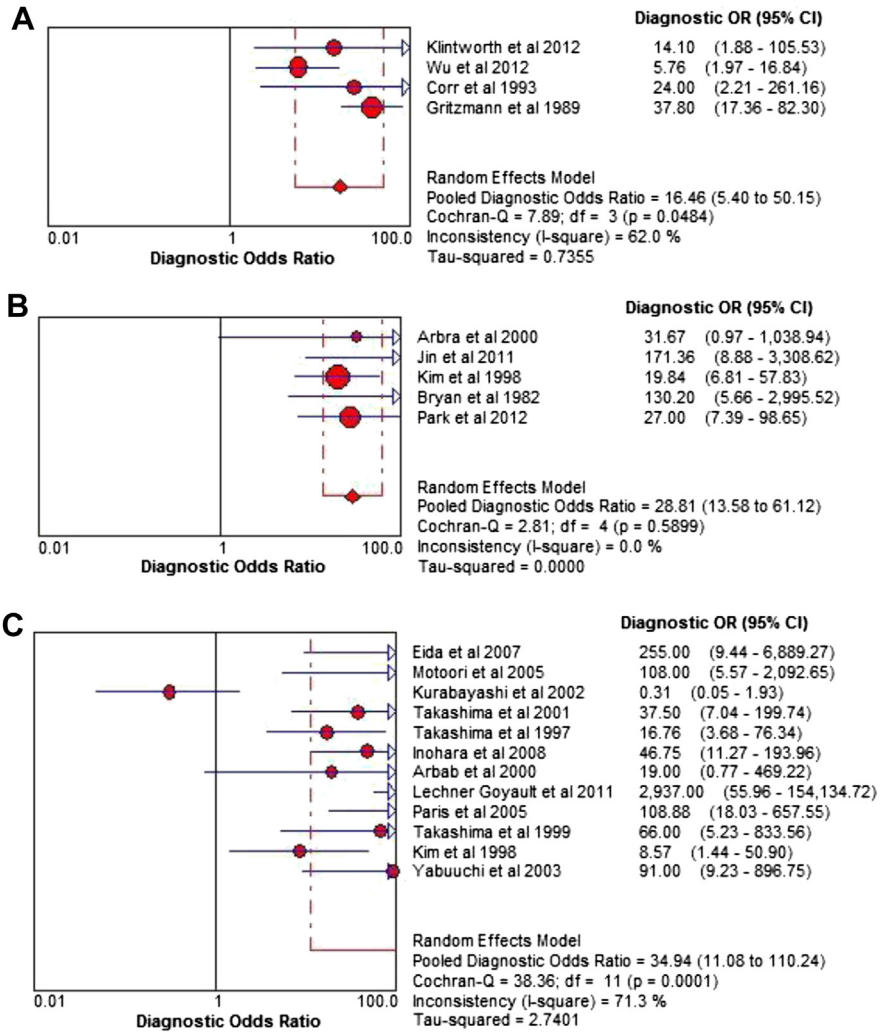
Ping Xiong, MD
Department of Ultrasound
Shanghai Ninth People's Hospital
Shanghai Jiao Tong University School of Medicine
No.639 Zhizaoju Road
Shanghai 200011
China
Xiongpxp@163.com



Supplementary Fig. S1. Chart showing the study design characteristics based on the QUADAS tool.



Supplementary Fig. S2. The Deek funnel plot for testing publication bias.



Supplementary Fig. S3. The Forest plot (random effects model) of pooled diagnostic odds ratio for differential diagnosis of between benign and malignant salivary gland tumors by ultrasonography (**A**), computed tomography (**B**), and magnetic resonance imaging (**C**).