Elastography reduces over-medical procedures in thyroid nodules: a systematic review and meta-analysis

Kun Yan (yankun@mail.xjtu.edu.cn)
the Second Affiliated Hospital of Medical College, Xi'an Jiaotong University
Wei Cheng
xi'an children hospital
jiancang Ma
Xian Jiao Tong University
donggang Han
xi'an jiaotong university
Xin Xu
Xi'an Jiaotong University
Yanni Chen
Xi'an Children Hospital
Jie Li
xi'an jiaotong university
Yiming Li
xi'an jiaotong university
Zhiliang Wang
xi'an jiaotong university
Pengyuan Jie
xi'an jiaotong university

Research

Keywords: thyroid nodule, elastography, systematic review, meta-analysis

Posted Date: June 19th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-32644/v1

Abstract

Background To evaluate the diagnostic performance of elastography in the differentiating the malignancy from thyroid nodule. Methods The literature published in English or Chinese prior to December 2019 which diagnosed by elastography in thyroid nodules were reviewed and summarized. According the inclusion and exclusion criteria, the elegant theses were reserved. Quality of each study was evaluated by the qu ality a ssessment of d iagnostic a ccuracy s tudies tool (QUADAS). Comprehensive Meta-Analysis Version 2.0 was used to describe primary results and explore homogeneity. Publication bias analysis was evaluated by Review Manager 5.3. Meta-Disc version 1.4 was applied to explore threshold effect, and pool sensitivity and specificity. A Summary ROC curve was constructed to calculate the area under the curve. Results As the heterogeneity was mild, we performance on random effect model. The publication bias was negligible. The elastography showed high accuracy in the diagnosis of malignant. The pooled sensitivity and specificity with corresponding 95% confidence interval were 0.80 (0.78, 0.82) and 0.80 (0.79, 0.81) respectively. AUC of sROC was 0.9087. Sensitivity analysis demonstrated that the pooled estimates were stable and reliable. Conclusions The elastography has obvious sensitivity and specificity in differentiate malignant from the thyroid nodule. Several different types of elastography could achieve the same aim, but the cut-off value is associate with the purpose of the screen program and the subtypes of the technique.

Introduction/background

Thyroid nodules are the most common clinical symptoms. They are the part of thyroid tissues which are abnormal in structural and/or functional，include cysts, inflammatory nodules, tumor nodules, hyperplasia nodules. Thyroid nodules could divide into non-cancerous (benign) and cancerous (malignant). Thyroid nodules are multifactorial diseases, which are effect by the complex of the contributing genes and environmental factors. It is very difficult to proceed etiological diagnosis.

In the recent decades, the morbidity rate of thyroid nodules is increasing rapidly due to both the explosive growth in the morbidity of thyroid neoplasm and the development of diagnostic technique. In clinic, only about 4% to 8% of all with thyroid nodules could detect by palpation, while 41% by ultrasonography, what is more, 56% to 67% by autopsy [1] . Physicians detect thyroid nodules by palpation occasionally. While palpation is subjective index, it may be involving in basis inevitably. Although thyroid nodules are extremely common, thyroid carcinomas are relatively rare. Only 5% to 15% of thyroid nodules could lead to regional and distant metastases and even death[2]. It is essential to identify the malignant nodules in order to establish an individual therapeutic strategy for each patient. The thyroid nodules could be detected by their imaging modalities (CT, MRI and US), but they could not accurately differentiate malignant and benign. FNA/FNB biopsy, an invasive and expensive method, are only applied to the suspected nodules[3].

Elastography, a subtype of ultrasound, is available to evaluate the stiffness of soft tissues and give diagnostic information about the presence or status of disease[4]. Elastography, as an objective index, is superior to palpation in sensitivity and specificity. Non-palpable nodules would barely detected by FNA/FNB，but it is more easily for elasticity. Furthermore, elastography can provide additional data, include the invasion and metastasis of thyroid cancer, which may influence the treatment procedures essentially.

The recent studies give conflicted result about the diagnostic performance of elastography in the malignant thyroid nodules [5-7]. The physician may be hesitated about whether the elastography should be the preferred strategy or not. We intend to perform a systematic review and meta-analysis, try to give an objective evaluation to the diagnostic performance of elastography in the differentiation performance of thyroid nodules

Methods

Types of studies

All types of diagnostic validation studies which aimed to evaluate the elastography as a diagnosis performance in screening the malignant lesion from thyroid nodules were considered. Reviews, letters, conference proceedings, cases report and personal opinions were also considered in order to compare their bibliographic references with the results of our literature search.

Search strategy

Abstracts of records were retrieved by the digital and manual searches. The papers were screened by two reviewers independently, which included one experiencer in thyroid diseases (Dr. Kun Yan) and a non-professional staff (Dr. Wei Cheng). Full text of potentially relevant articles was obtained and independently assessed by above reviewers. Final selection was always based on the full text of the publications. Disagreement between reviewers was resolved by discussion. The database of computer-aided literature search is PubMed, Embase, the Cochrane Library, and China biomedical literature service system (*SinoMed*) prior to December 2019. The search was performed by using terms "elastography" and "thyroid nodule" for clinical diagnosis trial.

Selection of articles

Inclusion criteria for the main analysis were follows: (1) Clinical diagnosis trials. (2) Adult patients with thyroid nodules. (3) Confirm pathological diagnosis as golden standardization. (4) Each study should cover several pathology types. (5) The case of each paper should not less than 30. (6) Sufficient information could be extracted. To avoid duplicate publication for data collection from the same group of patients, only the most details or recent series were included in our analysis.

Articles were selected according to the road map (Figure 1) of diagnostic systematic reviews and guidelines. Firsts, articles were excluded according to criteria. After the titles and abstracts of were searched and the duplications were removed, we recorded the final studies included in the meta-analysis.

Quality assessment and Date Extraction

The original data were exacted and filled into the same standardized from each article by two observers (Kun Yan and Wei Cheng) independently. To resolve disagreement between reviewers, a blind and neutral third reviewer (Yanni Chen) assessed all involved items. The quality assessment of diagnostic studies (QUADAS), an international quality assessment system, was performed to assess the quality of included studies [Dietrich, 2017 #50]. This method consists of 14 items (Figure 2). Each item was assessed by the scores as "low (no),""high (yes)" and"unclear". The phrased to the answer "high (yes)"indicates low risk of bias and vice versa. The general information was extracted from each study. Accurate data was extracted to construct 2 × 2 table[8].

Homogeneity test

The heterogeneity generally attributed to the actual difference among studies, which must be explained and removed in meta-analysis. The Cochrane Q and inconsistency index (I^2) were used to estimate the heterogeneity of individual studies. $P < 0.05$(Cochrane Q) suggests presence of heterogeneity could not explain by chance. I^2 describes the percentage of total variation across studies. I^2 (<50%) shows that homogeneity could be applicated to meta-analysis[9].

Threshold effect

Representation of accuracy estimates from each study in a receiver operating characteristic (ROC) space and computation of Spearman correlation coefficient between the log (Sen) and log(1-Spe) were assessed for threshold effect.

Different threshold effects may be result in false-positive due to lack of standardization[10] Differential threshold value maybe cause by the different in sensitivities and specificities of every studies. A typical pattern of "shoulder arm" plot in a receiver operating characteristic (ROC) space show a strong positive spearman correlation between the log(Sen) and log(1-Spe), which suggest the significant result[11, 12].

Publication bias

Studies with optimistic results are more likely to be accepted for publication than the unfavorable results. Since publication bias may cause an inaccurate threshold value, which would exaggerate clinical effects and result in potentially erroneous decision. So it is important to assess the extent of the bias and the potential impact on the conclusions[13]. In the literature on publication bias, the funnel plot is always regarded as a tool to detect the bias. In the absence of publication bias, the data points form a symmetric funnel-shaped distribution, whereas an asymmetric distribution indicates the presence of publication bias [14]. Publication bias also could detect by fail-safe N (NFS) quantitatively. NFS refers to the exact number of hypothetical studies with null results that would be required to nullify the statistical significance of combined effect [15]. A small NFS should be more concern. However, if NFS is large, it is confident to the combined effect.

Statistical pooled effect

The studies which are reasonably homogeneous could the apply to estimate the statistical pooled effect [16]. The estimates of the meta-analysis have two distinct types, the fixed effect model (FEM) and the random effect model (REM)[17]. The fixed effect model requires every theses come from the same distribution. While in the random effect model the heterogeneity could be involved. It means that all trials do not come from the same distribution. It always suggested that the outlier trials should be found and analyzed. Once it is detected the heterogeneity, the random effect model is recommended [18]. The reasons are as follow (a) The results of two models are same when heterogeneity is absent. (b) The results of the random models are more conservative and credible. (c) Measurement in biology has an intrinsic heterogeneity because of variability between the individuals. The variability can best be represented by a random effect model. For the reasons above, the results of systematic review are more credible when the random effects model is adopted.

The heterogeneity can be evaluated from the forest plots. If heterogeneity only due to threshold effect, the accuracy of statistical pooled effect could be described by a summary receiver operator characteristic (sROC) curve and it could be quantified by the area under the curve (AUC). A sROC curve summarizes and combines the true and false positive rates from different diagnostic studies. The overall performance of diagnostic studies can be visualized and reflected without affected by every cut-off value [19]. The best diagnostic procedure should yield a point in the upper left corner or coordinate (0, 1) of the sROC space, representing 100% sensitivity (no false negatives) and 100% specificity (no false positives) at the individual subject level. Similarly, AUC ranges from 1 for a perfect test which means all diagnoses are correct to 0 for a test which is invalid. If there is heterogeneity due to sources other than threshold effect, pooling outcome could only be accepted within homogeneous [20].

Sensitivity analysis

The pooled estimates were reappraised when suspicious studies were excluded, and the reappraised results were compared with the original results to assess stability and reliability of original meta-analysis.

Statistical analysis

The homogeneity test, threshold effect analysis, pooled weighted sensitivity and specificity, sROC curve and sensitivity analyses were performed by using Meta-Disc version 1.4. Publication bias analysis was performed by using Review Manager 5.3 and Comprehensive Meta-Analysis Version 2.0.

The sensitivity, specificity, and diagnostic odds ratio (OR) with corresponding 95% confidence intervals (CI) were calculated for each study. Meanwhile, the pooled sensitivity, specificity and diagnosis OR were also calculated for each group. The diagnostic OR expresses the ratio of the malignant to the negative nodules [19].

We further performed sensitivity analysis to explore the reasons of heterogeneity and examined characteristics of included studies. To summarize these results, we constructed a summary receiver operator characteristic (sROC) curve. Q-values were calculated from the sROC curves. Meanwhile, the area under sROC curve (AUC) was also calculated to show the probability of the correctly ranked diagnostic test values for a pair of diseased and negative subjects[21].

Results

Characteristics of included studies

The literature search was performed as described, and 2257 potentially relevant theses were identified. After reading the titles and abstracts, we kept 27 eligible articles. Then intensive reading the full texts, there are 16 articles met our inclusion criteria. The relevant literatures were excluded for the following reasons: The objective of studies was not explore the diagnostic performance of elastography on thyroid lesions (n=5); suspicious cases could not confirm or exclude by the pathological diagnosis (n=3); Researchers did not report the data which could be used to get TP, FP, TN, and FN (n=2). Elastography was not regarded an major method (n=1). Multiple subsets of data in one study were included because: The two methods of elastography were used to perform on thyroid nodules in Azizi G [22]. Elastography scores of Rago were classified on a scale of 1 to 5，the nodules with Rago scores of 4 to 5 were diagnosis as malignancy[23]. Asteria scores were range from1 to 4. Nodules with Asteria scores of 3 to 4 were diagnosis malignancy[24]. There were at least three different elastography score systems in clinic [25-27]. The Max values ranged from 3 to 7. The thresholds values of differentiating malignant lesions ranged from 3 to 5, while the sensitivity ranged from 15.7% to 100%; and specificity from 60.8% to 97.3%. Detailed characteristics and data of each included study are presented in Tables 1.

All the conditions and methods of the included studies, as shown in Figure 2, were used for different quality assessment of diagnostic accuracy. Each included study was strictly evaluated according to the QUADAS criteria [28], as shown in attached Table 1.

Meta-analysis

The pooled outcome is shown in Figure 3. All 16 studies involving 4582 patients were identified. The total incidence of malignant thyroid nodule(s) was 33.22% (1706 of 5135 nodules) in the suspicious thyroid(s) group. The incidence of general group was from 3.6 to 30.8%[29-31].

The forest plots of sensitivities and specificities with a 95% CI from 16 studies are shown in Figure 3a. A homogeneity test of the Diagnostic Odds Ratio (DOR) shows Q= 158.21 (P=0.000), I^2= 90.5%. None of significant heterogeneities is detected. And Sensitivity and Specificity shows Q = 787.81 (P= 0.0000), I^2= 98.1%, and Q = 321.00 (P=0.0000), I^2= 95.3%, respectively. The heterogeneities are also rejected [32] as be show in Figure 3b & 3c. We adopt the random effect model, and the meta-regression is followed to analyses the possible of heterogeneities. Next a sROC space was draw to explore the threshold effect of sensitivity against 1-specificity. The pattern of the points in this plot is not a"shoulder-arm" shape (Figure 4). A Spearman correlation is performed as a further test about threshold effect. The Spearman correlation coefficient was equal to 0.009(P= 0.974) and indicates that there is no threshold effect resulting in variations in accuracy estimates among individual studies. The expressed as AUC of sROC curve is 0.9087 (Figure 4).

Analysis of Heterogeneity and Sensitivity

In a meta-analysis of diagnostic test, heterogeneity is an important index of the accuracy diagnosis. And it also estimates the appropriateness of statistical pooling of results from various studies [33]. The heterogeneities are brought by several factors, such as the cut-off value, assay method, and the distribution of patients and so on.

In fact, heterogeneity may not be entirely avoided in meta-analysis, so it is necessary for us to explore the reason and extent of heterogeneity. Generally, one of the most important sources of heterogeneity is the threshold effect in a diagnostic study. So, we explored the threshold effect firstly, which was evaluated the Spearman correlation coefficient by inverse variance. We found that there was no statistically significant difference (Spearman's correlation coefficient = 0.009, P= 0.974). The source of heterogeneity was explored by Meta regression analysis. The results show that cut-off value (score≥4) is the significant factor RDOR (relative diagnostic odds ratio) =13.21, Q-value =0.0863, while the effects of the age (RDOR =0.32, Q-value =0.493), assay method (RDOR = 0.09, Q-value =0.1505), and QUADAS (RDOR = 1.32, Q-value =0.7052) are negligible. The sensitivity analysis was also performed to identify further analysis of these sources. It could convince the elastography as a popular screen method in thyroid nodule from our results.

The analysis of the heterogeneity, sensitivity, specificity, and AUC are unchanged, when suspicious literature Ugur Unluturk [34] and Franco Uliaque C [35] were excluded respectively due to the seemingly differential sensitivity and specificity from others. The pooled sensitivity and specificity is similar, and the corresponding 95% CIs are predominantly overlapped with each other. The homogeneity is not reversed. The results of sensitivity analysis demonstrate that the pooled estimates are stable and reliable.

Publication Bias

The funnel plot in Figure 5 shows that the studies are distributed symmetrically about the combined effect size. There is none of publication bias. The NFS is 260. This means that it would at least include 260 'null' studies to exceed the above result.

Discussion

Thyroid nodules are the most common symptoms of malignant thyroid lesion with certain mortality. Recently, the incidence of the diseases increases about 19 times [37], but the mortality does not change significantly. So, it is urgent to explore an efficient, repeatable screening method to reduce the over medical surgery.

Thus, efficient screening procedure is vital to improve detection rate and reduce the unfavorable prognosis. Due to the clinical signs are subtle and nonspecific, and the serum marker are always unreliable [36]. Patients are always recommended to take imaging procedures in screening thyroids nodules [37], such as ultrasonography (US), computerized tomography (CT), magnetic resonance imaging (MRI).

It is a major challenge for clinicians, pathologists and radiologists how to screen the thyroid nodules efficiently. However, any single procedure looks like insufficiently to cover all kind of thyroid nodules [38, 39]. Therefore, it is necessary to evaluate the reliable screening procedures. Here, we give a review about the related procedures to screen thyroid nodules.

There are several diagnosis techniques once been used to screen thyroid nodules and detected the malignant lesion. The programs of screening malignant lesion from the thyroid nodule include physical examination, biochemical test, imaging methods, cytological examination and histological examination. Physician usually inspects nodules and regional lymph nodes by palpation. The biochemical tests could detected biomarker of cancer within serum[40]. They include Calcitonin, neuron-specific enclose, and thyroglobulin. Imaging methods includes radio nuclide scanning, CT and MRI, but it is difficult to distinguish by a mere anatomical image, morphologic and their functional characterization.

It well know that the histopathological diagnosis is the gold standard [41], but the accurate diagnosis of malignant thyroid nodules requires the surgery operation and sufficient specimens. Cytological diagnosis is mini-invasive procedure, which is inferior to the histopathology in the accuracy. Recently, guidelines for fine-needle aspiration (FNA) and fine-needle biopsy (FNB) have been published [42], but the invasive and the rates of false-positive or false-negative results remain the headache challenges. Unfortunately, up to two-thirds cytology suspected cases may be misdiagnosed and result in unnecessary the operation, especially for the follicular cancer [43]. In clinic, only 5% of thyroid nodule cytological biopsies detected by FNA/FNB were confirmed malignant finally, while a large multi-center Cohort study conducted that about more than 10% benign nodules confirmed by FNA/FNB are proved malignant after operation [44]. It is suggested that the half of the malignant and the suspected diagnosed by cytology have to be following up. That will be a long-term heavy burden on the patient. The practitioner would face a dilemma to the suspected cases whether surgically resection should be taken [45]. The majority postoperative date suggested most of lesions were in the early stages. The conservative strategy to the suspected nodules is the enormous specimens and invasive repeatedly. Furthermore, many factors can affect adequacy about FNA/FNB such as the experience of the aspirator, procedural variations, and lesion characteristics, specimen preparation, and needles gauge size and so on. Therefore, the effects of FNA/FNB are inferior to ultrasonography, especially to elastography.

Ultrasonography is a noninvasive, painless, repeatable, without a risk of hemorrhage or infection technique which could cover the regional lymph node and thyroid lesion simultaneously[46]. Conventional ultrasonography is always regard as basic checkup, which consider hypoechogenicity, blurred margins, microcalcifications, neovascularization as malignancy. But neither single one nor combination of the above criteria has sufficient sensitivity and specificity to detected malignant lesions. Elastography is a subtype of ultrasonography, which evaluates the stiffness-deformability index. It is more objectively than the palpation [47]. Recent studies show that the malignant lesions tend to be more vascularized and 3 times stiffer than the surrounding tissues, while benign nodules are less than 1.7 times stiffer than normal tissue[48]. There are several prominent elastography technologies involve in thyroid nodules evaluation [49], include Quasistatic Elastography (Strain Imaging), Acoustic Radiation Force Impulse imaging (ARFI), Shear Wave Elasticity Imaging (SWEI), Supersonic Shear Imaging (SSI), Transient Elastography, and so on. The elastography scores which display the qualitative or quantitative outcome compared with background tissues. Our review reveals that elastography is an efficient preoperative screening technology which could reduce unnecessary invasive biopsies on benign thyroid nodules [50]. Elastography could also be as a triage procedure and navigation of surgery including the FNA/FNB. It could also follow up _suspect malignant nodules. Furthermore, elastogarphy could give more information to individualizing strategy for the thyroid nodules that will require surgery [Deeks, #16].

Elastography is a medical imaging technique which is used for the investigation of many diseases, especially in subcutaneous lesions. It can get not only a mere anatomical image, but also additional diagnostic information superimposed on the image. It can also be used to guide biopsies and even replace them entirely. With development and widespread of elsstography, more and more suspicious cases can benefit from it. There are numerous ultrasound elastographic techniques, since they applied in clinical. But only certain types of elastography are suitable for thyroid nodules screening. The current meta-analysis shows that several elastographic techniques are valuable and stable in screening the thyroid nodules. Elastography is prominent for completely noninvasive, painless, presenting no risk of hemorrhage and infection.

The ideal treatment procedure should have more efficient, more economic and less side effect[51], whereas it is should be based on the accurate diagnosis. The normal thyroid nodules does not need special treatment, except following-up regularly [52]. While the therapeutic strategy to malignant lesion is significant differences [53]. Traditional thyroidectomy was recommended for patients with malignant or highly risk thyroid nodules as basic treatment [54]. But such strategy would perhaps lead to trauma for the patients without malignant lesions. Although the thyroidectomy is the most common surgical procedures to thyroid nodules, the complications incidence varies widely in the literature, include recurrent laryngeal nerve injury, hypoparathyroidism, aggressive and metastasis. Increased complications associated with the dimensions of thyroidectomy, extensive malignancy, and repeat surgery. Among all methods which devote to reduce the complications, the early accurate diagnosis is more effectively.

The technique of elastography was developing, the cut-off values and equipment are different in different institutions. So, there are numbers of differences in the procedure and facility involved into the study. Despite these, the quality and reliability of our meta-analysis was faithful according to the QUADAS.

Although all of include studies showed a high prevalence of malignant lesions deemed to be stiff, there was some crossover value between the elastography mean values of malignant and benign lesions, resulting in it difficult to determine the recommended cut-off values to screen thyroid nodules. The variance of cut-off values might be influenced by different screen procedure, patient selection, pathological characteristic of lesions and measurement of elastography values[55].

In conclusion, on the base of our systematic review, elastography on thyroid nodules is a noninvasive technique with a high sensitivity and a good specificity to detected malignant thyroid lesions. As the anatomical characteristics of the thyroid and the frequency of nodular pathology make it is perfect technique. Because the cut-off value should be determined according to the purpose of the screen program, the popular screening cut-off value for malignant would be based on many factors.

Conclusions

The elastography has obvious specificity and sensitivity in differentiate malignancy from the thyroid nodules. Several different types of elastography could achieve the same aim, but the cut-off value is associate with the purpose of the screen program and the subtypes of the technique.

Declarations

Ethics approval and consent to participate: The study does not use any animal or human data or tissue.

Consent for publication: The authors consent for publication.

Availability of data and material: The study does not involve any individual human tissue material.

Competing interests：The authors declare that they have no competing interests.

Funding: This project was supported by the key project of Xi'an Children's Hospital (No. 2019B04) and Xi’an Science and Technology Research Project (No.2020ms11), and the State Scholarship Fund of China (NO. 201806285146).

Authors' contributions: YK, CW, and Prof. CYn reviewed and summarized literatures. They peform the meta-analysis. Prof HDg gave the professional advice about elastography, **XX, LJ, LYm, WZl and JPy** gave help in writing manuscript.

Acknowledgements: We sincerely appreciate the Professor Libin Ye, who is from Department of Cell Biology, Microbiology and Molecular Biology, University of South Florida, for his forthright comments on this paper.

References

[1] Girardi FM, Silva LMD, Flores CD.A predictive model to distinguish malignant and benign thyroid nodules based on age, gender and ultrasonographic features[J].Brazilian journal of otorhinolaryngology.2017.

[2] James BC, Aschebrook-Kilfoy B, Cipriani N, Kaplan EL, Angelos P, Grogan RH.The Incidence and Survival of Rare Cancers of the Thyroid, Parathyroid, Adrenal, and Pancreas[J].Annals of surgical oncology.2016,23(2):424-33.

[3] Paajanen I, Metso S, Jaatinen P, Kholova I.Thyroid FNA diagnostics in a real-life setting: Experiences of the implementation of the Bethesda system in Finland[J].Cytopathology : official journal of the British Society for Clinical Cytology.2017.

[4] Ahn HS, Lee JB, Seo M, Park SH, Choi BI.Distinguishing benign from malignant thyroid nodules using thyroid ultrasonography: utility of adding superb microvascular imaging and elastography[J].La Radiologia medica.2017.

[5] Tian W, Hao S, Gao B, Jiang Y, Zhang S, Guo L, et al.Comparison of Diagnostic Accuracy of Real-Time Elastography and Shear Wave Elastography in Differentiation Malignant From Benign Thyroid Nodules[J].Medicine.2015,94(52):e2312.

[6] Trimboli P, Paone G, Zatelli MC, Ceriani L, Giovanella L.Real-time elastography in autonomously functioning thyroid nodules: relationship with TSH levels, scintigraphy, and ultrasound patterns[J].Endocrine.2017,58(3):488-94.

[7] Negro R, Greco G, Colosimo E.Ultrasound Risk Categories for Thyroid Nodules and Cytology Results: A Single Institution's Experience after the Adoption of the 2016 Update of Medical Guidelines by the American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi[J].Journal of thyroid research.2017,2017:8135415.

[8] Hurley J.Meta-analysis of clinical studies of diagnostic tests: developments in how the receiver operating characteristic "works"[J].Archives of pathology & laboratory medicine.2011,135(12):1585-90.

[9] Pawlowska M.Evaluation of three proposed markers for language impairment in English: a meta-analysis of diagnostic accuracy studies[J].Journal of speech, language, and hearing research : JSLHR.2014,57(6):2261-73.

[10] Glasmacher SA, Stones W.A systematic review and diagnostic test accuracy meta-analysis of the validity of anion gap as a screening tool for hyperlactatemia[J].BMC research notes.2017,10(1):556.

[11] Cronin P, Kelly AM, Altaee D, Foerster B, Petrou M, Dwamena BA.How to Perform a Systematic Review and Meta-analysis of Diagnostic Imaging Studies[J].Academic radiology.2018.

[12] Takwoingi Y, Riley RD, Deeks JJ.Meta-analysis of diagnostic accuracy studies in mental health[J].Evidence-based mental health.2015,18(4):103-9.

[13] Ding Y, Yan JL, Fang AN, Zhou WF, Huang L.Circulating miRNAs as novel diagnostic biomarkers in hepatocellular carcinoma detection: a meta-analysis based on 24 articles[J].Oncotarget.2017,8(39):66402-13.

[14] Papageorgiou SN, Dimitraki D, Coolidge T, Kotsanos N.Publication bias & small-study effects in pediatric dentistry meta-analyses[J].The journal of evidence-based dental practice.2015,15(1):8-24.

[15] Deeks JJ, Macaskill P, Irwig L.The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed[J].Journal of clinical epidemiology.2005,58(9):882-93.

[16] Liao LJ, Chen HW, Hsu WL, Chen YS.Comparison of Strain Elastography, Shear Wave Elastography, and Conventional Ultrasound in Diagnosing Thyroid Nodules[J]. *J Med Ultrasound* 2019, **27**(1):26-32

[17] Coronado-Zarco R, Cruz-Medina E, Arellano-Hernandez A, Chavez-Arias D, Leon-Hernandez SR.Effectiveness of calcitonin in intermittent claudication treatment of patients with lumbar spinal stenosis: a systematic review[J].Spine.2009,34(22):E818-22.

[18] Nikoloulopoulos AK.A mixed effect model for bivariate meta-analysis of diagnostic test accuracy studies using a copula representation of the random effects distribution[J].Statistics in medicine.2015,34(29):3842-65.

[19] Sengul D, Sengul I: **Association between Tsukuba elasticity scores 4 and 5 on elastography and Bethesda undetermined cytology on US-guided FNA with 27-G needle, verified by histopathology: a cut-off point of 20 mm of diameter designated for thyroid nodules[J]**. *J BUON* 2019, **24**(1):382-390.

[20] Lee J, Kim KW, Choi SH, Huh J, Park SH.Systematic Review and Meta-Analysis of Studies Evaluating Diagnostic Test Accuracy: A Practical Review for Clinical Researchers-Part II. Statistical Methods of Meta-Analysis[J].Korean journal of radiology.2015,16(6):1188-96.

[21] Doebler P, Holling H.Meta-analysis of Diagnostic Accuracy and ROC Curves with Covariate Adjusted Semiparametric Mixtures[J].Psychometrika.2015,80(4):1084-104.

[22] Moon HJ, Sung JM, Kim EK, Yoon JH, Youk JH, Kwak JY.Diagnostic performance of gray-scale US and elastography in solid thyroid nodules[J].Radiology.2012,262(3):1002-13.

[23] Rago T, Santini F, Scutari M, Pinchera A, Vitti P.Elastography: new developments in ultrasound for predicting malignancy in thyroid nodules[J].The Journal of clinical endocrinology and metabolism.2007,92(8):2917-22.

[24] Asteria C, Giovanardi A, Pizzocaro A, Cozzaglio L, Morabito A, Somalvico F, et al.US-elastography in the differential diagnosis of benign and malignant thyroid nodules[J].Thyroid : official journal of the American Thyroid Association.2008,18(5):523-31.

[25] Yan CX, Luo ZY, Liu XM, Huang PT, Mo GQ, Hong YR, et al.[Ultrasonic scores of conventional ultrasound and ultrasound elastography in the diagnosis of thyroid nodular lesions][J].Zhonghua yi xue za zhi.2013,93(21):1630-3.

[26] Azizi G, Keller J, Lewis M, Puett D, Rivenbark K, Malchoff C.Performance of elastography for the evaluation of thyroid nodules: a prospective study[J].Thyroid : official journal of the American Thyroid Association.2013,23(6):734-40.

[27] Pandey NN, Pradhan GS, Manchanda A, Garg A.Diagnostic Value of Acoustic Radiation Force Impulse Quantification in the Differentiation of Benign and Malignant Thyroid Nodules[J].Ultrasonic imaging.2017,39(5):326-36.

[28] Downs SH, More SJ, Goodchild AV, Whelan AO, Abernethy DA, Broughan JM, et al.Evaluation of the methodological quality of studies of the performance of diagnostic tests for bovine tuberculosis using QUADAS[J].Preventive veterinary medicine.2017.

[29] Yang L, Zheng RS, Wang N, Zeng HM, Yuan YN, Zhang SW, et al.[Analysis of Incidence and Mortality of Thyroid Cancer in China, 2013][J].Zhonghua zhong liu za zhi [Chinese journal of oncology].2017,39(11):862-7.

[30] Topstad D, Dickinson JA.Thyroid cancer incidence in Canada: a national cancer registry analysis[J].CMAJ open.2017,5(3):E612-E6.

[31] Davies L, Morris L, Hankey B.Increases in Thyroid Cancer Incidence and Mortality[J].Jama.2017,318(4):389-90.

[32] Gupta S, Fischer JK, Lunkenheimer P, Loidl A, Novak E, Jalarvo N, et al.Effect of adding nanometre-sized heterogeneities on the structural dynamics and the excess wing of a molecular glass former[J].Scientific reports.2016,6:35034.

[33] Sridharan S, Jones B, Caudill B, Nakaima A.Steps towards incorporating heterogeneities into program theory: A case study of a data-driven approach[J].Evaluation and program planning.2016,58:88-97.

[34] Unluturk U, Erdogan MF, Demir O, Gullu S, Baskal N.Ultrasound elastography is not superior to grayscale ultrasound in predicting malignancy in thyroid nodules[J].Thyroid : official journal of the American Thyroid Association.2012,22(10):1031-8.

[35] Franco Uliaque C, Pardo Berdun FJ, Laborda Herrero R, Perez Lorenz C.Usefulness of semiquantitative elastography in predicting malignancy in thyroid nodules[J].Radiologia.2016,58(5):366-72.

[36] Martin Martin G, Ismail A, Sancho Moya C, Precioso Estiguin J, Roig Vila JV.[Brain metastases as the first clinical sign of papillary thyroid cancer][J].Cirugia espanola.2011,89(7):472-4.

[37] Enomoto Y, Enomoto K, Uchino S, Shibuya H, Watanabe S, Noguchi S.Clinical features, treatment, and long-term outcome of papillary thyroid cancer in children and adolescents without radiation exposure[J].World journal of surgery.2012,36(6):1241-6.

[38] Trimboli P, Nigri G, Guidobaldi L, Romanelli F, Aurello P, Crescenzi A, et al.Early diagnosis by FNA cytology should not influence the outcome of differentiated thyroid cancer[J].Endocrine journal.2016,63(4):367-73.

[39] Farra JC, Picado O, Liu S, Ouyang W, Teo R, Franco AM, et al.Clinically significant cancer rates in incidentally discovered thyroid nodules by routine imaging[J].The Journal of surgical research.2017,219:341-6.

[40] Patel HH, Goyal N, Goldenberg D.Imaging, genetic testing, and biomarker assessment of follicular cell-derived thyroid cancer[J].Annals of medicine.2014,46(6):409-16.

[41] Dos Reis LL, Tuttle RM, Alon E, Bergman DA, Bernet V, Brett EM, et al.What is the gold standard for comprehensive interinstitutional communication of perioperative information for thyroid cancer patients? A comparison of existing electronic health records with the current American Thyroid Association recommendations[J].Thyroid : official journal of the American Thyroid Association.2014,24(10):1466-72.

[42] Layfield LJ, Abrams J, Cochand-Priollet B, Evans D, Gharib H, Greenspan F, et al.Post-thyroid FNA testing and treatment options: a synopsis of the National Cancer Institute Thyroid Fine Needle Aspiration State of the Science Conference[J].Diagnostic cytopathology.2008,36(6):442-8.

[43] Saglietti C, Piana S, La Rosa S, Bongiovanni M.Hyalinizing trabecular tumour of the thyroid: fine-needle aspiration cytological diagnosis and correlation with histology[J].Journal of clinical pathology.2017,70(8):641-7.

[44] Mehrotra P, McQueen A, Kolla S, Johnson SJ, Richardson DL.Does elastography reduce the need for thyroid FNAs?[J].Clinical endocrinology.2013,78(6):942-9.

[45] Abboud B, Sleilaty G, Braidy C, Zeineddine S, Ghorra C, Abadjian G, et al.Careful examination of thyroid specimen intraoperatively to reduce incidence of inadvertent parathyroidectomy during thyroid surgery[J].Archives of otolaryngology--head & neck surgery.2007,133(11):1105-10.

[46] Fujiwara K, Fukuhara T, Koyama S, Donishi R, Kataoka H, Kitano H, et al.Ultrasound-Guided Transoral Videolaryngoscopic Surgery for Retropharyngeal Lymph Node Metastasis of Papillary Thyroid Cancer[J].Case reports in oncology.2017,10(2):649-55.

[47] Gazdag A, Nagy EV, Erdei A, Bodor M, Berta E, Szabo Z, et al.Aortic stiffness and left ventricular function in patients with differentiated thyroid cancer[J].Journal of endocrinological investigation.2015,38(2):133-42.

[48] Rianna C, Radmacher M.Comparison of viscoelastic properties of cancer and normal thyroid cells on different stiffness substrates[J].European biophysics journal : EBJ.2017,46(4):309-24.

[49] Dietrich CF, Barr RG, Farrokh A, Dighe M, Hocke M, Jenssen C, et al.Strain Elastography - How To Do It?[J].Ultrasound international open.2017,3(4):E137-E49.

[50] Wojtaszek-Nowicka M, Slowinska-Klencka D, Sporny S, Popowicz B, Kuzdak K, Pomorski L, et al.The efficiency of elastography in the diagnostics of follicular lesions and nodules with an unequivocal FNA result[J].Endokrynologia Polska.2017.

[51] Ibrahim EY, Busaidy NL.Treatment and surveillance of advanced, metastatic iodine-resistant differentiated thyroid cancer[J].Current opinion in oncology.2017,29(2):151-8.

[52] Chung YS, Yoo C, Jung JH, Choi HJ, Suh YJ.Review of atypical cytology of thyroid nodule according to the Bethesda system and its beneficial effect in the surgical treatment of papillary carcinoma[J].Journal of the Korean Surgical Society.2011,81(2):75-84.

[53] El-Zohairy M, Zaher A.Re-operation for the treatment of well differentiated thyroid cancer: necessity, safety and impaction on further management[J].Journal of the Egyptian National Cancer Institute.2004,16(3):130-6.

[54] Yu WB, Zhang NS.Protection and Dissection of Recurrent Laryngeal Nerve in Salvage Thyroid Cancer Surgery to Patients with Insufficient Primary Operation Extent and Suspicious Residual Tumor[J].Asian Pacific journal of cancer prevention : APJCP.2015,16(17):7457-61.

[55] Frederix GW, Hovels AM, Severens JL, Raaijmakers JA, Schellens JH.[Threshold value for reimbursement of costs of new drugs: cost-effectiveness research and modelling are essential links][J].Nederlands tijdschrift voor geneeskunde.2015,159:A7728.

Table

Table 1 The characteristics of included studies.

	author publication	years(ref)	country	number of patients	number of nodules	number of mailgnant nodules	% mailgnant nodules	reference surgery	reference FNA	Measurement quantitative=1 qualitative=2 both=3	mean age(years)	male	threshold value	full score
1	Zhan J	2017	China	30	58	21	36.20%	30	0	2	51	7	≥4	5
2	Liao Li-Jen	2019	Taiwan/China	185	185	21	11.35%	21	185	3	51.3	56	≥38.3kpa	95.0kpa
3	Yang BR	2017	Korea	80	80	6	7.50%	6	74	2	46	14	≥4	5
4	Wang Yi-jun	2012	China	168	208	41	19.70%	41	206	2	37	32	≥3	5
5	Ugur Unluturk	2012	Turkey	194	237	58	24.47%	58	237	2	78	44	≥2	3
6	T.Rago	2007	Italy	92	92	31	33.69%	0	92	2	43	29	≥3	5
7	Stoian.D	2012	Romania	69	107	11	10.28%	107	0	2	50.08	2	≥3	6
8	Shao ning-ning	2010	China	115	68	40	58.80%	40	115	2	48.7	24	≥5	7
9	Park AY	2014	Korea	453	476	379	79.62%	398	123	1	45.7	89	≥85.2kpa	94.0kpa
10	Pandey NN	2017	India	40	40	14	35.00%	21	19	1	40	7	≥2.53m/s	5.52m/s
11	Demet Sengul	2019	Turkey	547	655	609	92.98%	609	655	2	47.1	121	≥4	5
12	Hee Jung Moon	2012	Korea	676	703	217	30.87%	489	221	2	49.7	120	≥3	5
13	Ghobad Azizi	2011	USA	706	912	86	9.43%	0	912	2	43.8	34	≥2	4
14	Franco Uliaque	2016	Spanish	295	314	19	6.05%	19	267	2	N	N	≥2	3
15	Colakoglu B	2016	Turkey	256	293	71	24.23%	71	185	2	34	23	≥3	4
16	Azizi G	2015	usa	676	707	82	11.61%	132	575	3	52.29	96	≥3.53m/s	5
				4582	5135	1706		2042	3866			698		

Figures

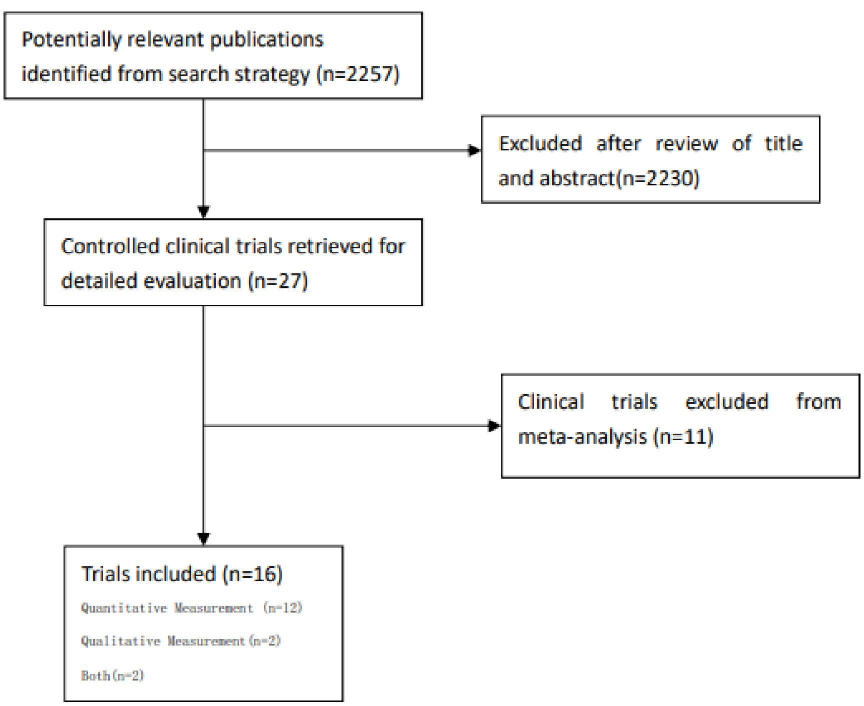
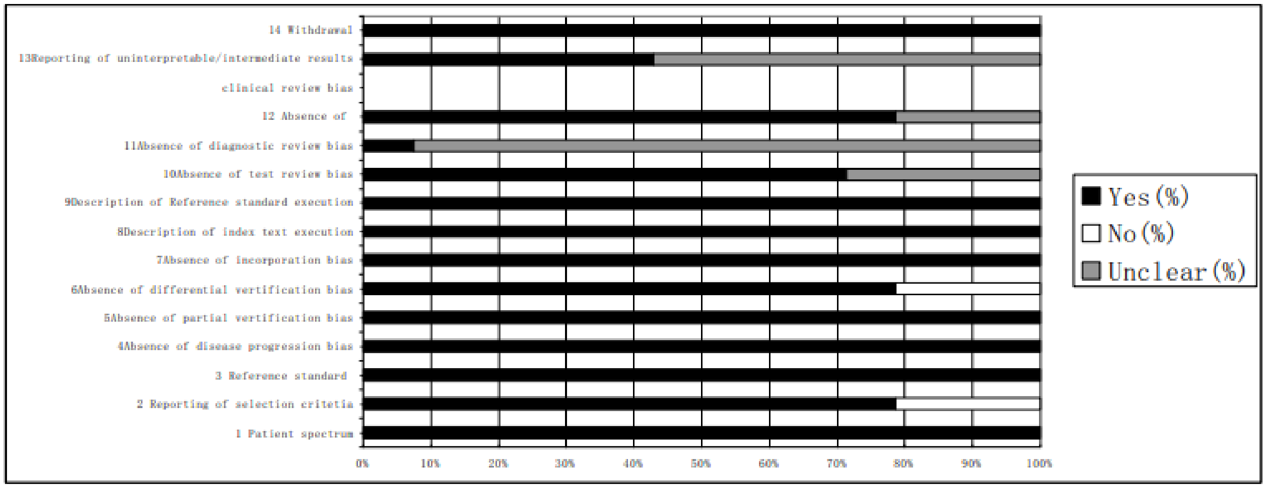


Figure 1

Follow chart illustrating the selection process



QUADAS graphic. Analysis of study quality considering the Quality assessment of studies of diagnostic performance included in systematic reviews (QUADAS) checklist.

Figure 2

QUADAS graphic. Analysis of study quality considering the Quality assessment of studies of diagnostic performance included in systematic reviews (QUADAS) checklist.

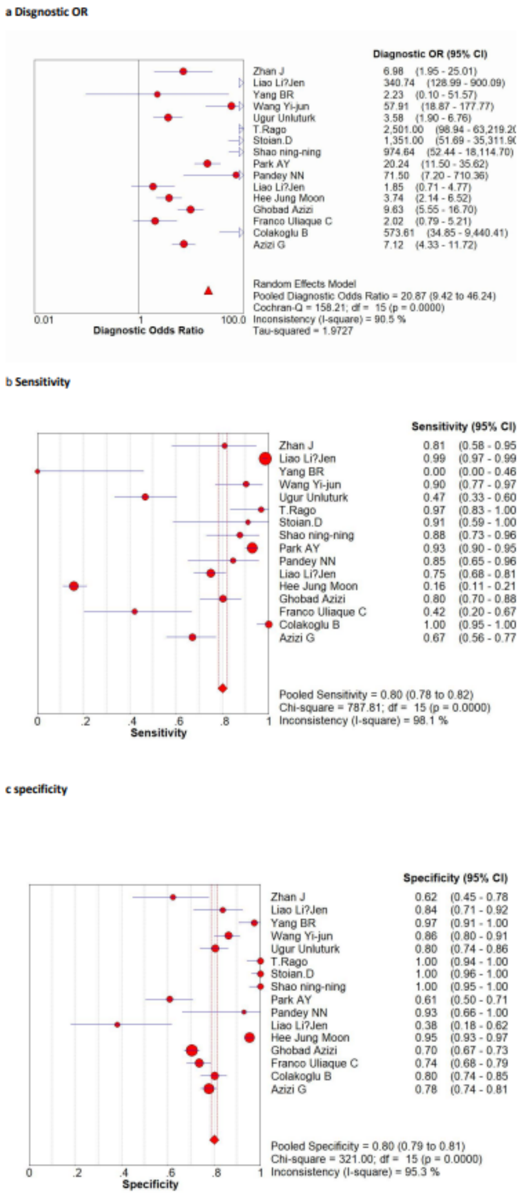


Figure 3

Forest polts of Disgnostic OR, sensitivity and specificity, with corresponding 95% CIs from all eligible studies. a Disgnostic OR b Sensitivity c specificity

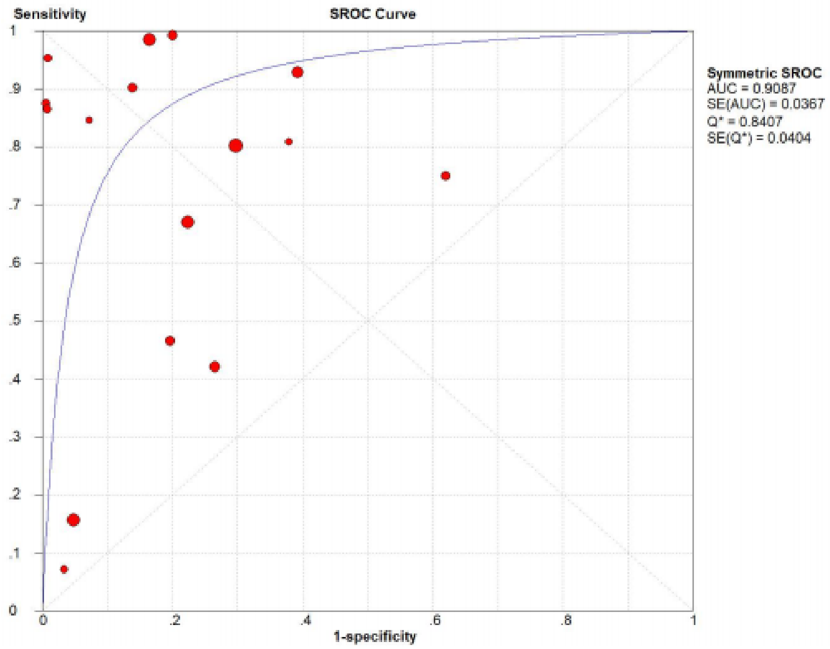
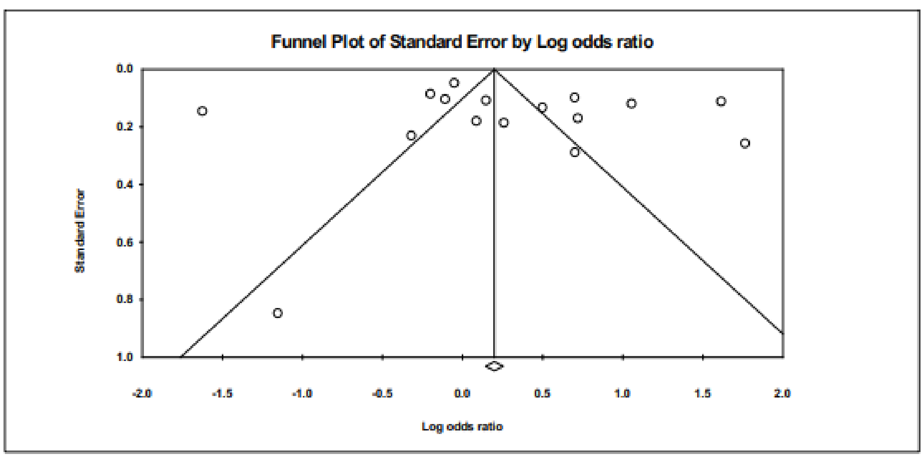


Figure 4

SROC curve for eligible studies.



Note: Flakes represent published articles. Rhombus represent actual and theoretical combined effect size.

Figure 5

Funnel plot of eligible studies.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AttachedTable1.doc
- AttachedTable1.doc
- AttachedTable1.doc