

Diagnostic Value of Dual-Energy CT in Differentiating Malignant and Benign Thyroid Nodules: A Systematic Review and Meta-Analysis

Kunlin Ye^{1,2}, Yongyao Kuang¹, Jianye Liang³, Xi Xu¹, Changzheng Shi¹, Qingqing Cheng¹, Zijie Jiang¹, Ziyun Xiang^{2*}, Liangping Luo^{1*}

¹Department of Medical Imaging Center, The First Affiliated Hospital of Jinan University, Guangzhou, China

²Department of Medical Imaging, Longgang District People's Hospital of Shenzhen, Shenzhen, China

³Department of Medical Imaging, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

Email: Kunlinye@foxmail.com, *szlgxzy@163.com, *tluolp@jnu.edu.cn

How to cite this paper: Ye, K.L., Kuang, Y.Y., Liang, J.Y., Xu, X., Shi, C.Z., Cheng, Q.Q., Jiang, Z.J., Xiang, Z.Y. and Luo, L.P. (2021) Diagnostic Value of Dual-Energy CT in Differentiating Malignant and Benign Thyroid Nodules: A Systematic Review and Meta-Analysis. *Journal of Biosciences and Medicines*, 9, 54-76.

<https://doi.org/10.4236/jbm.2021.94005>

Received: March 5, 2021

Accepted: April 12, 2021

Published: April 15, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Objectives: To evaluate the diagnostic performance of the quantitative iodine parameters, including Iodine Concentration (IC), Normalized Iodine Concentration (NIC), and λ_{HU} , in distinguishing malignant and benign thyroid nodules. **Methods:** Relevant studies were searched from Web of Science, PubMed, Embase, Cochrane Library, China National Knowledge Infrastructure database and other complementary sources from inception to May 20, 2020. Study selection, data extraction, quality assessment, and data analyses were performed following the Cochrane standards and the PRISMA-DTA guideline. **Results:** Eight studies were included (595 patients with 737 thyroid nodules). The pooled sensitivity, specificity and summary diagnostic odds ratio of IC were 79% (95% CI: 69% - 86%), 76% (95% CI: 65% - 84%) and 11 (95% CI: 5 - 27) respectively; those of NIC were 78% (95% CI: 70% - 85%), 80% (95% CI: 74% - 85%) and 15 (95% CI: 9 - 24) respectively; those of λ_{HU} were 80% (95% CI: 71% - 87%), 77% (95% CI: 70% - 83%) and 14 (95% CI: 8 - 24) respectively. **Conclusion:** DECT can be a potential evaluation tool for thyroid nodules. The NIC may be the most sensitive iodine parameter and could be comparable between different DECT machines in thyroid nodule assessment.

Keywords

Thyroid, Nodules, Differential Diagnosis, Dual-Energy CT, Iodine

1. Introduction

Thyroid nodules are common clinical findings in recent years, and the prevalence of thyroid nodules is 20% - 70% in the general population [1] [2] [3]. Autopsy results showed that thyroid nodules are present in about 50% of the population [1] [4] [5]. Recent statistics show that the incidence of thyroid diseases is rising rapidly in many countries, while mortality is relatively constant [2] [3] [6]. Currently, cytology is the gold standard for diagnosing thyroid nodules, but ultrasound-guided fine-needle aspiration (FNA) is invasive, with certain risks and limitations [7] [8]. Although most of the complications of FNA biopsy, including perithyroidal hemorrhage, and parenchyma edema, were reported to be mild and could be recovered within a few hours, the discomfort and the risk of tumor metastasis during aspiration should not be ignored [7] [8]. A non-invasive examination could reduce pain or discomfort or prevent the probability of cancer seeding along the track of the needle. Besides, it may reduce patients' mental pressure by painlessly recognizing the relatively low probability of having malignant diseases, which may be beneficial to patients' disease management and clinical results [9].

Conventional gray-scale ultrasonography (US) is widely used to detect thyroid nodules. It can estimate benign and malignant solid thyroid nodules based roughly on morphologic features, such as nodule shape, margin, size, and calcification [10]. However, we could not omit the relatively low interobserver reliability and the lack of convenient lexicon in US to characterize thyroid nodules. And US elastography, a new advance in US, might not be an ideal tool to differentiate the types of thyroid nodules either [11]. Single-energy computed tomography (SECT) can evaluate the size of nodules, but the correlation between the nodules size and the nature of nodules is unclear [12].

With the recent advance in dual-energy CT (DECT), dual-energy imaging is burgeoning in clinical practice. DECT provides information on material decomposition by simultaneously acquiring images at two different energy levels in one scan. It can quantify intrinsic attenuations related to different atomic numbers and tissue density in specific regions of interest (ROI) [13]-[18]. Evaluating thyroid nodules using DECT is a novel non-invasive technique. It can provide a quantitative measurement of iodine concentration, which may be related to the pathophysiological status of thyroid nodules [19] [20].

Several studies have attempted to evaluate this application [21] [22] [23]. However, due to the different techniques used, the value of relevant quantitative iodine parameters and the diagnostic accuracy were not determined, and it is worth exploring. Therefore, we conducted a meta-analysis to evaluate the diagnostic value of the quantitative DECT parameters in distinguishing malignant and benign thyroid nodules.

2. Material and Methods

2.1. Search Strategy

Based on the “PICO principle” of evidence-based medicine, we constructed a question that defines population, intervention, comparison, and outcome: for patients with thyroid nodules, whether quantitative DECT iodine parameters can be used to distinguish benign and malignant thyroid nodules as effectively as cytology. And we followed the preferred reporting items for systematic reviews and meta-analyses of diagnostic test accuracy studies (PRISMA-DTA) guidelines to conduct our meta-analysis [24] [25]. We used the following keywords to set up search formula: thyroid, dual energy computed tomography, Dual-Energy CT, and DECT. Chinese and English literatures published from January 2007 to October 2019 were searched through Web of Science, PubMed, Embase, Cochrane Library, and Chinese National Knowledge Infrastructure (CNKI), Wanfang database (<http://www.wanfangdata.com/>) and the details regarding the search strategies are reported in Supplementary materials. We review newly published research every two months and use manual retrieval if necessary.

2.2. Inclusion and Exclusion Criteria

Inclusion criteria: 1) Assess the use of Dual-energy CT in evaluating the thyroid nodules, and contain at least one quantitative parameter, including iodine concentration, normalized iodine concentration and Slope of spectrum curve; 2) a proper reference standard was used. Malignant thyroid nodules need to be pathologically confirmed, while benign thyroid nodules could be confirmed based on imaging follow-up of more than 12 months as an alternative; 3) sufficient data to construct a 2×2 contingency table to calculate diagnostic accuracy; 4) The imaging data of patients are preoperative or pre-treatment (such as chemotherapy) data.

Inclusion criteria: 1) Assess the use of Dual-energy CT in evaluating the thyroid nodules, and contain at least one quantitative parameter, including iodine concentration, normalized iodine concentration and Slope of spectrum curve; 2) a proper reference standard was used. Malignant thyroid nodules need to be pathologically confirmed, while benign thyroid nodules could be confirmed based on imaging follow-up of more than 12 months as an alternative; 3) sufficient data to construct a 2×2 contingency table to calculate diagnostic accuracy; 4) The imaging data of patients are preoperative or pre-treatment (such as chemotherapy) data.

2.3. Data Extraction

Two reviewers independently selected and extracted the data related to the study in the form of a data table. The data included first author, year of publication, country of origin, and research type, the patient’s data (the number of cases, the sex ratio, and the average age of patients), the nodules data (the pathological type of nodules, number, nodule size), the mean and standard deviation of re-

lated quantitative iodine parameters (IC, NIC, and λ_{HU}), and the value of true-positive, false-positive, false-negative and true-negative, which were used to calculate the diagnostic performance of each study. The imaging data included the type and manufactory of the CT scanner, the type of dual-energy parameter used to evaluate the lesion, and the protocol of contrast media injection (including contrast media type, flow rate, dose, and phase).

2.4. Assessment of Study Quality

The quality of the included literature was evaluated according to the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) [26]. Each literature will be evaluated on the basis of 14 criteria. Since the use of total scores may manifest bias itself, we do not calculate the scores but show the methodological quality evaluation of each study through a summary chart [27]. If there is a dispute over the quality assessment results, another senior clinician or statistician will be invited to discuss and reach a consensus [28]. We choose Begg's test as a quantitative indicator to evaluate publication bias, with $p < 0.05$ indicates the possibility of potential publication bias. Furthermore, we also try to minimize the publication bias through comprehensive search strategies.

2.5. Statistical Analysis

The primary purpose of this systematic review is to evaluate the DECT quantitative parameter's diagnostic capability by calculating pooled sensitivity, pooled specificity, and summary diagnostic odds ratio (SDOR). We use bivariate model to create summarized receiver operating characteristic (SROC) curves for each group. The diagnostic parameters of benign and malignant nodules will be estimated if data have a good homogeneity. Also, we attempt to explore the sources of heterogeneity.

2.5.1. Test of Heterogeneity

In the diagnostic meta-analysis, the sources of heterogeneity mainly include diagnostic threshold difference, random error, publication bias, different clinical data (gender and ages, pathological type of nodules), and methodological differences. Cochran's Q test and Higgins' statistics were used to evaluate the heterogeneity of each study. $p < 0.05$ in the Q test indicated that the study's heterogeneity was more significant than that caused by random error. The I^2 statistic was used to quantitatively describe the percentage of inter-study heterogeneity in the overall difference, and $I^2 > 50\%$ suggests heterogeneity and may require further evaluation of the source of heterogeneity [29].

2.5.2. Threshold Effect

The threshold effect refers to the difference in the diagnostic boundaries selected in different diagnostic studies, which may lead to a correlation between the log (sensitivity) and log (1-specificity) of each study result. The correlation was analyzed using Spearman's correlation test. If the Spearman correlation coefficient

shows a strong positive correlation ($p < 0.05$), it may indicate a threshold effect. The summarized receiver operating characteristic (SROC) curve, which derived from the bivariate model, can reflect the estimated diagnostic capacity [29] [30] [31].

2.5.3. Statistical Model

Typically, when calculating the effect size, the calculation was performed using random-effects or fixed-effects models according to the heterogeneity. However, the heterogeneity of diagnostic tests is usually greater than that of interventional tests due to patient diversity and environmental differences, so the random-effect model should be used to calculate the effect size [31]. The recommended bivariate mixed-effects regression model was used to calculate the combined sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, SDOR, area under the curve (AUC) and its 95% confidence interval (95% CI) [32].

The diagnostic effectiveness will be calculated using standardized mean difference (SMD) based on the consistency of parameters, measurement methods, and scanning phase.

Meta-regression was used to analyze the influence of SDOR from covariates, including study design type, country of origin, DECT manufacturer, contrast media concentration, contrast media flow rate, and scanning phase. $p < 0.05$ was selected as the criteria of statistical difference. Softwares used for data analysis and statistics include Stata14 (StataCorp LP, College Station, TX) and Review Manager 5.3 (Cochrane Collaboration, Oxford, UK).

3. Result

3.1. Study Selection

A total of 395 articles were preliminarily retrieved. After excluding duplicate items, there were 262 remaining studies. Screening of titles and abstracts yielded 23 articles for full-paper review. Of the remaining 23 full-text articles, 15 did not meet the predetermined inclusion criteria and were excluded. Finally, eight articles were included in this meta-analysis (**Figure 1**) [22] [23] [33]-[38].

3.2. Characteristics of Studies and Quality Assessment

The basic information with the diagnostic performance of the included literature is shown (**Table 1**, diagnostic performance data in **Supplementary Table S1**). Seven of the included studies were from China, and one was from South Korea. Three studies had a prospective design, and the rest were retrospective. Half of the studies were published after 2016 (range 2013-2019). Eight studies recorded the measurement of one or more DECT parameters, among which five listed the IC results, 6 listed the NIC results, and four recorded the λ_{HU} .

In this systematic review and meta-analysis, a total of 595 patients (32% males and 68% females, age 50.64 ± 11.46 years on average) were included, with an

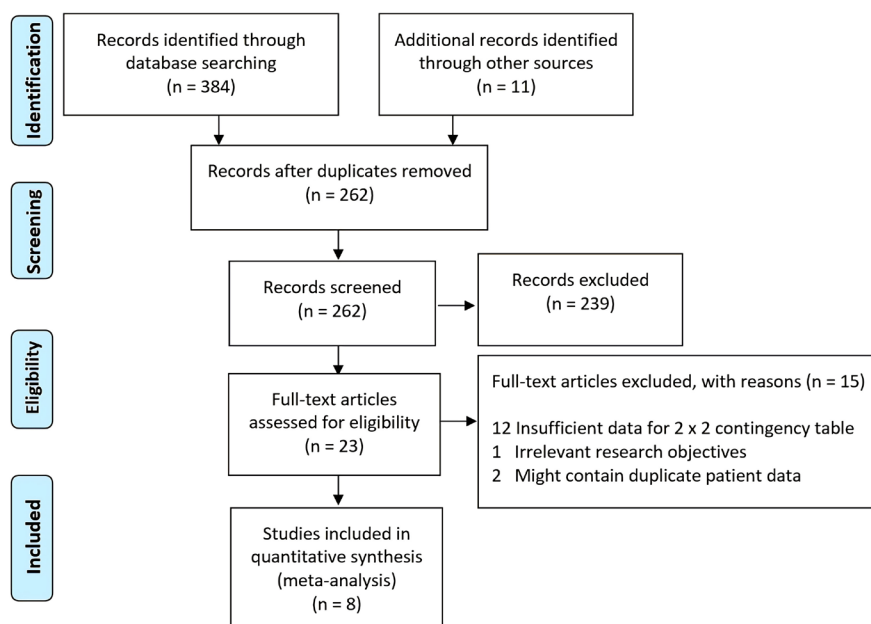


Figure 1. Flowchart of the selection and identification process based on inclusion and exclusion criteria.

Table 1. Summary data of the eight included studies.

First Author (Reference No.)	Year	Design	Country of origin	DECT Method*	Scanner	CM Volume	Injection Rate (mL/s)	CM Type	No. of Patients	Mean age ± SD (y)	No. of Nodules	Lesion Diameter (cm)
Li [36]	2015	R	China	IC _{non} NIC _{vp}	GE 750	1.3 - 1.5 ml/kg	3 - 4	Obitridol 300	123	M: 47.01 ± 2.37 F: 51.74 ± 12**	122	≥1
Gao [22]	2016	R	China	IC _{non}	GE 750	1.5 ml/kg	3 - 4	Omnipaque 300	53	54 ± 9	53	NR
Cheng [33]	2019	P	China	IC _{vp} Slope _{vp}	GE 750	1.3 ml/kg	3.5	Omnipaque 300	96	46.89 ± 12.67	101	NR
He [37]	2014	P	China	IC _{non} Slope _{non}	GE 750	1.0 - 1.5 ml/kg	3	Omnipaque 300	32	48.49 ± 11.85*	75	both >1 and ≤1
Xue [38]	2013	P	China	NIC _{vp} Slope _{vp}	GE 750	1.0 ml/kg	2.0 - 2.5	Optiray 320	55	52.4 ± 13.5	80	NR
Lee [23]	2018	R	Korea	IC _{vp} NIC _{vp}	Philips IQon	80 ml	2.5	Omnipaque 300/Pamiray 300/Optiray 350	63	52.8 ± 11.8	63	M: 1.1 (0.8 - 1.4) B: 1.6 (0.6 - 2.6)
Zhao [35]	2015	R	China	NIC _{vp}	Siemens Flash	1.0 ml/kg	3	Iopromide 370	82	56.2 ± 8.9	90	NR
Zhang [34]	2017	R	China	NIC _{vp} Slope _{vp}	Siemens Flash	50 ml	2.5 - 3	Omnipaque 350	91	45.7 ± 6.1	153	≥1

Abbreviations: DECT, dual-energy CT; CM, contrast medium; R, retrospective; P, prospective; IC, iodine quantification; IC, iodine concentration; NIC, normalized iodine concentration; Slope, Slope of spectrum curve; NR, not reported; M= male; F, female; GE 750, a dual-energy gemstone spectral CT (Discovery CT 750 HD, GE Healthcare, USA); Philips IQon, a dual-layer spectral CT (IQon spectral CT, Philips Healthcare, Tokyo, Japan); Siemens Flash, a second-generation dual-source dual-energy CT (SOMATOM Definition Flash, Siemens Healthcare, Forchheim, Germany); DECT Method*: the subscript indicating the phase when the parameters were measured, non = non-enhanced phase; vp = venous phase; Lee's phase was estimated according to the image acquisition time; **: Estimated value using statistical approach.

average of 74 patients (range of 32 to 123) per study. And 737 cases of thyroid nodules were analyzed (364 malignant nodules, 373 benign nodules). Malignant nodules included papillary carcinoma (n = 325, 89.29%), follicular carcinoma (n = 15, 4.12%), medullary carcinoma (n = 10, 2.75%), undifferentiated carcinoma (n = 5, 1.37%), poorly differentiated carcinoma (n = 4, 1.10%), lymphoma (n = 4, 1.10%), metastatic carcinoma (n = 1, 0.27%). Benign nodules included nodular goiter (n = 246, 65.95%) follicular adenoma (n = 80, 21.45%) nodular goiter with adenoma (n = 21, 6.17%) and granuloma (n = 3, 0.80%). 23 cases (6.17%) were follow-up cases or unknown pathological results (Table 2).

The method used in the literature quality assessment is based on the QUASAS-2 evaluation tool (Figure 2, Figure 3). Although the included studies had made a meticulous choice in the section of patient selection, there was still an unknown risk of bias due to the variety of the patients with thyroid nodules (e.g., some studies excluding pregnant patients or patients aged under 20). In terms of the reference standard, one study using long-term follow-up results as one of the gold standards may pose a hidden risk of bias in determining thyroid benignity and malignancy [22]. While most studies use cytological results as the gold standard, we believe that the risk of bias is low. Besides, two studies did not mention the interpretation of the index test result without knowledge of the reference standard [37] [38]. The PRISMA-DTA checklist (Table S2, Table S3) are provided in the Supplementary Materials.

3.3. Diagnostic Effectiveness

The pooled sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, DOR, and AUC corresponding to the three parameters are shown in Table 3. NIC has a specific threshold effect in the estimation of diagnostic accuracy (p

Table 2. The pathologic types of thyroid nodules in eight studies.

Author (Year)	Malignant Nodules (n = 364, 100%)						Benign Nodules (n = 373, 100%)			
	Papillary carcinoma	Medullary carcinoma	Poor-differentiated carcinoma	Anaplastic carcinoma	Follicular carcinomas	Others*	Nodular goiters	Follicular adenoma	Granuloma	Others**
Li (2015)	27	1	4	2	-	1	71	15	1	-
Gao (2016)	25	-	-	-	1	-	5	5	-	17
Cheng (2019)	55	3	-	2	3	-	32	4	2	-
He (2014)	45	-	-	-	-	-	30	-	-	-
Xue (2013)	30	-	-	1	2	-	31	16	-	-
Lee (2018)	46	-	-	-	-	-	5	6	-	6
Zhao (2015)	60	-	-	-	-	-	25	5	-	-
Zhang (2017)	37	6	-	-	9	4	47	29	-	21
Total	325 (89.29%)	10 (2.75%)	4 (1.10%)	5 (1.37%)	15 (4.12%)	5 (1.37%)	246 (65.95%)	80 (21.45%)	3 (0.80%)	44 (11.80%)

NOTE: Others*: Li: Metastatic carcinoma (n = 1), Zhang: thyroid lymphoma (n = 4); Others**: Gao: no deterioration on long-term follow-up (n = 17), Lee: not reported (n = 6), Zhang: nodular goiter with adenoma (n = 21).

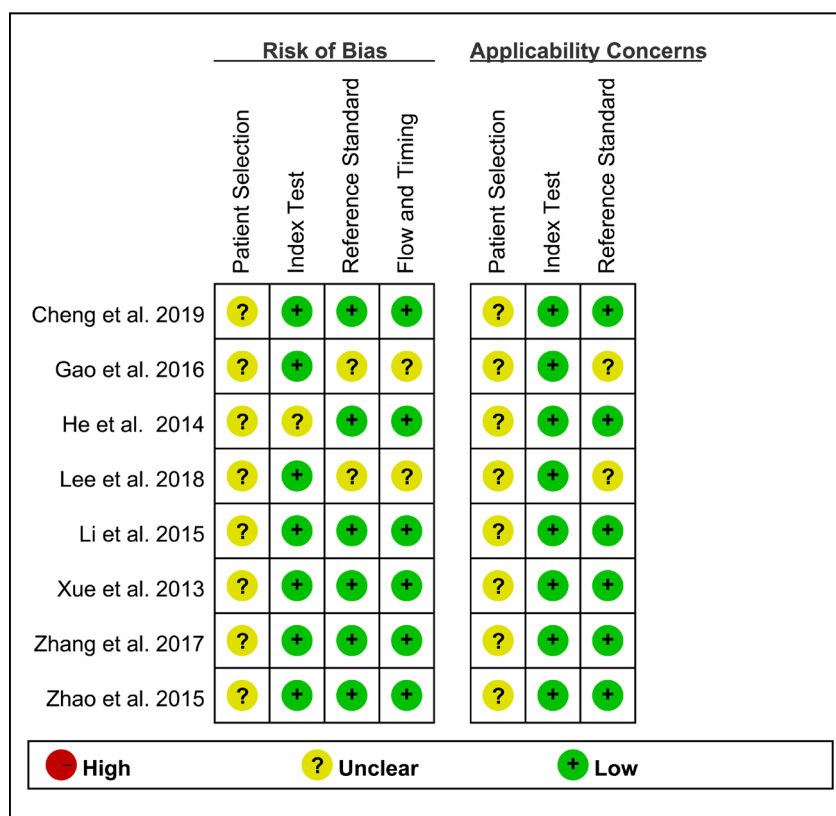


Figure 2. Risk of bias and applicability concerns summary.

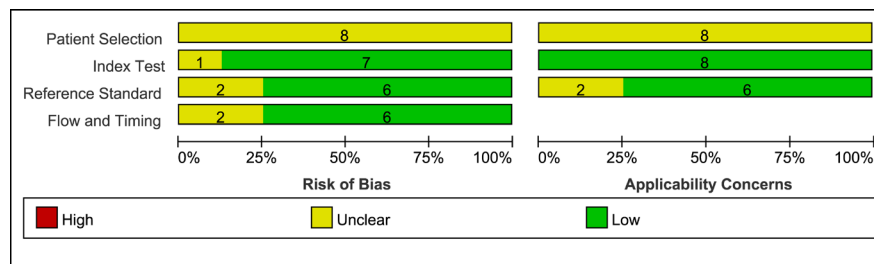


Figure 3. Risk of bias and applicability concerns graph.

Table 3. Diagnosis values of IC, NIC and λ_{HU} .

Index	Combined Studies	No. of Nodules	Sensitivity (95% CI)	Specificity (95% CI)	PLR (95% CI)	NLR (95% CI)	DOR (95% CI)	AUC (95% CI)	I ² Sensitivity Specificity	
IC	5	420	0.79 (0.69, 0.86)	0.76 (0.65, 0.84)	3.2 (2.1, 5.1)	0.28 (0.18, 0.45)	11 (5, 27)	0.84 (0.81, 0.87)	47.85%	50.69%
NIC	5	427	0.78 (0.69, 0.85)	0.80 (0.73, 0.86)	4.0 (2.8, 5.6)	0.27 (0.19, 0.38)	15 (8, 26)	0.85 (0.82, 0.88)	64.71%	0%
λ_{HU}	4	409	0.80 (0.71, 0.87)	0.77 (0.70, 0.83)	3.5 (2.6, 4.6)	0.26 (0.17, 0.39)	14 (8, 24)	0.81 (0.77, 0.84)	72.44%	0%

Abbreviations: AUC, area under the curve; CI, confidence interval; DOR, diagnostic odds ratio; IC, iodine concentration; NLR, negative likelihood ratio; PLR, positive likelihood ratio; NIC, normalized iodine concentration; Slope, slope of spectral curve; I², inconsistency index.

< 0.05), IC and λ_{HU} have no noticeable threshold effect in the evaluation of diagnostic accuracy ($p = 0.188$, $p = 0.6$). In terms of sensitivity, the heterogeneity of NIC and λ_{HU} was moderate (I² = 64.73%, 72.44%), while the heterogeneity of

IC was not significant ($I^2 = 47.85\%$). In terms of specificity, IC might show moderate ($I^2 = 50.69\%$), while NIC and λ_{HU} did not show obvious heterogeneity ($I^2 = 0\%$). The results showed that the three parameters' diagnostic capabilities were comparable, and the NIC was relatively better, as its DOR was 15 (95% CI: 8 - 26), and the AUC was 0.85 (95% CI: 0.81 - 0.88) (Table 3).

3.3.1. Iodine Concentration (IC)

IC was used to differentiate between malignant and benign thyroid nodules in five of the included literature. There were 215 cases of malignant thyroid nodules and 205 cases of benign thyroid nodules. The heterogeneity test ($I^2 = 71\%$, $p < 0.05$) suggested the existence of moderate heterogeneity without an obvious threshold effect ($p = 0.188$). In the random-effects model, the SMD of the IC was 1.22 (95% CI: 0.81 - 1.64) $p < 0.01$. The pooled sensitivity and specificity of IC were 79% (95% CI: 69% - 86%; $I^2 = 47.9\%$) and 76% (95% CI: 65%-84%; $I^2 = 0\%$) respectively. SDOR is 11 (95% CI: 5 - 27; $I^2 = 50.7\%$) (Figure 4). No significant publication bias was identified by Begg's tests ($p = 0.086$).

3.3.2. Normalized Iodine Concentration (NIC)

NIC was an indicator to identify between malignant and benign thyroid nodules

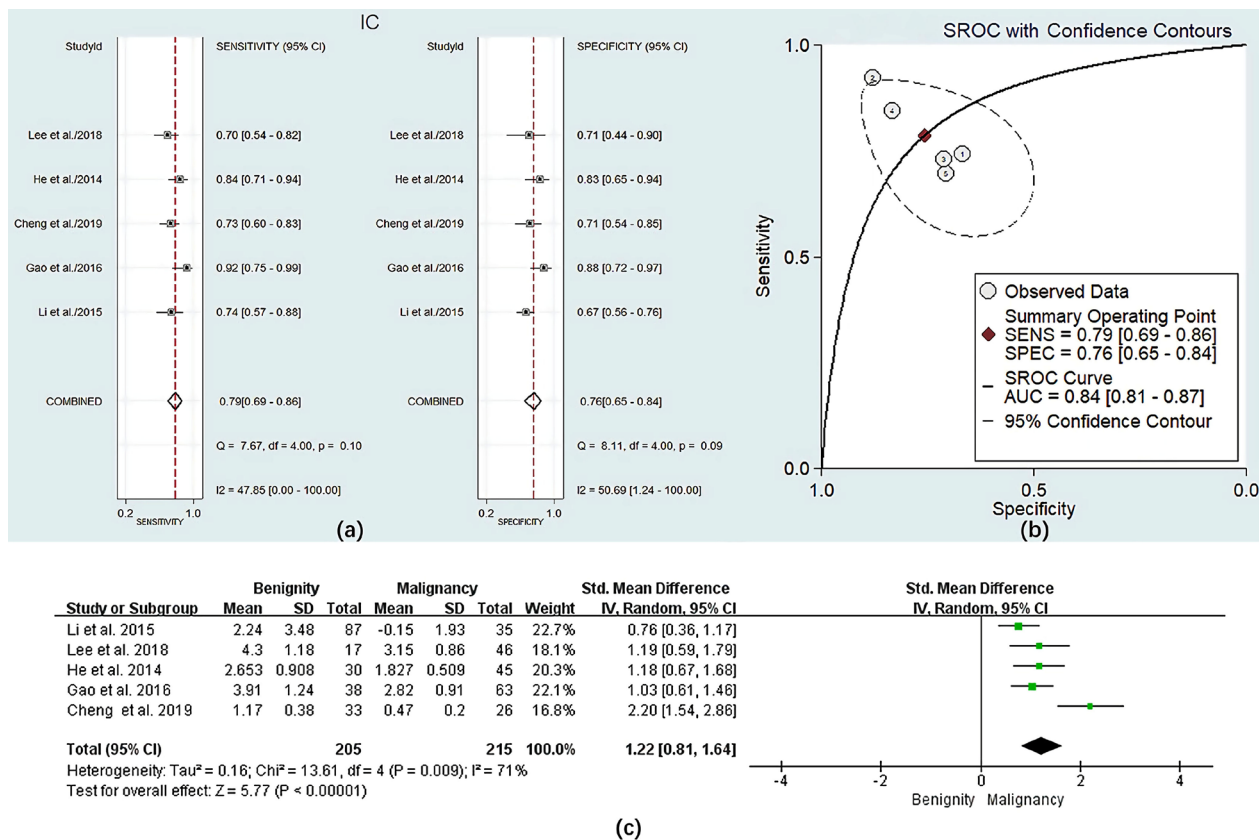


Figure 4. Diagnostic accuracy of dual-energy CT for thyroid nodules evaluation using IC technique. Forest plots of (a) diagnostic sensitivity (top left) and specificity (top right) and (b) SROC curve show results in discriminating between benign and malignant nodules. (c) Forest plot of the mean value of IC between benign and malignant nodules (Left chart listed by mean IC \pm SD mg·mL⁻¹), indicating obvious difference between benign and malignant nodules.

in five of the included literatures. There were 252 cases of malignant thyroid nodules and 175 cases of benign thyroid nodules. The heterogeneity test ($I^2 = 47\%$, $p = 0.076$) suggested insignificant heterogeneity with a potential threshold effect ($p < 0.05$). In the random-effects model, the SMD of the NIC was 0.92 (95% CI: 0.71 - 1.13), $p < 0.01$. The pooled sensitivity and specificity of NIC were 78% (95% CI: 69% - 85%; $I^2 = 64.7\%$) and 80% (95% CI: 73% - 86%; $I^2 = 0\%$) respectively. SDOR is 15 (95% CI: 8 - 26; $I^2 = 0\%$) (Figure 5). No significant publication bias was identified by Begg's tests ($p = 0.806$).

3.3.3. Slope of Spectrum Curve (λ_{HU})

λ_{HU} was used to differentiate between malignant and benign thyroid nodules in four of the included literature. There were 238 cases of malignant thyroid nodules and 171 cases of benign thyroid nodules. The heterogeneity test ($I^2 = 70\%$, $p < 0.05$) suggested moderate heterogeneity without an obvious threshold effect ($p = 0.6$). In the random-effects model, the SMD of the λ_{HU} was 1.08 (95% CI: 0.69 - 1.48) $p < 0.01$. The pooled sensitivity and specificity of λ_{HU} were 80% (95% CI: 71% - 87%; $I^2 = 72.4\%$) and 77% (95% CI: 70% - 83%; $I^2 = 0\%$) respectively. SDOR is 14 (95% CI: 8 - 24; $I^2 = 0\%$) (Figure 6). No significant publication bias was identified by Begg's tests ($p = 0.734$).

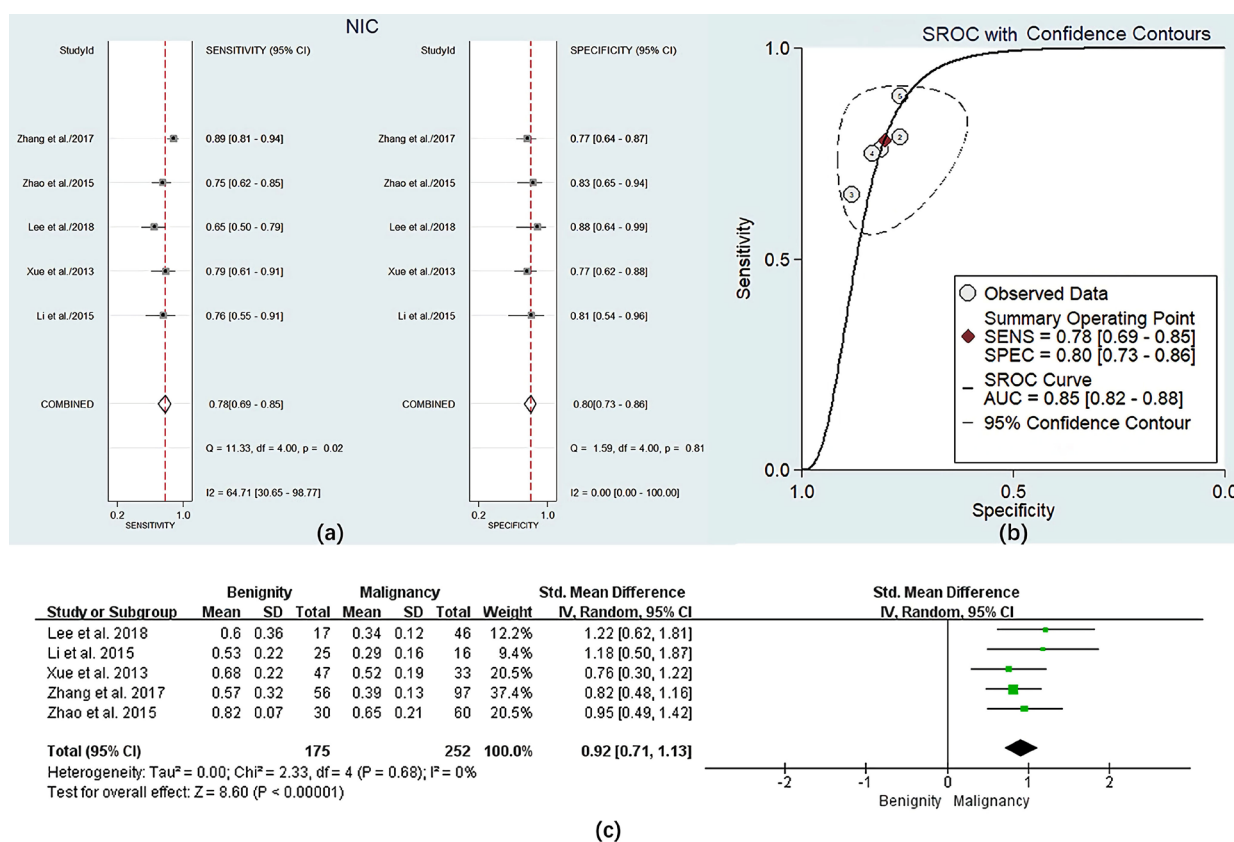


Figure 5. Diagnostic accuracy of dual-energy CT for thyroid nodules evaluation using NIC technique. Forest plots of (a) diagnostic sensitivity (top left) and specificity (top right) and (b) SROC curve show results in discriminating between benign and malignant nodules. (c) Forest plot of the mean value of NIC between benign and malignant nodules (Left chart listed by mean NIC \pm SD), indicating a positive result between benign and malignant nodules.

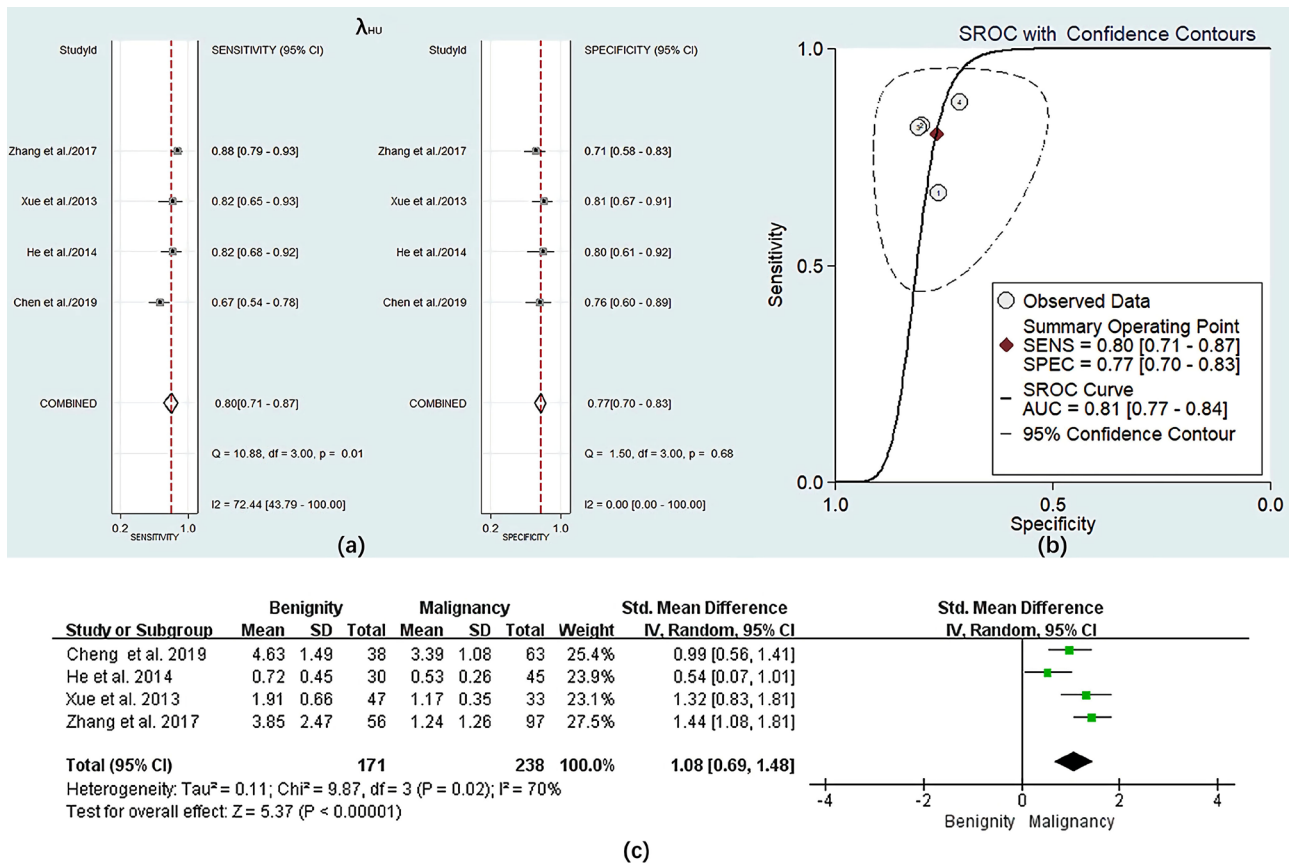


Figure 6. Diagnostic accuracy of dual-energy CT for thyroid nodules evaluation using λ_{HU} . Forest plots of (a) diagnostic sensitivity (top left) and specificity (top right) and (b) SROC curve show results in discriminating between benign and malignant nodules. (c) Forest plot of the mean value of λ_{HU} between benign and malignant nodules (Left chart listed by mean $\lambda_{HU} \pm SD$) shows a statistical difference.

3.3.4. Prediction of Post-Test Probability

Based on our results, the post-test probability, calculated by using the likelihood ratio of the NIC groups, was shown in the Fagan nomogram [32]. Referring to the currently recorded incidence of malignant thyroid tumors, we assume that the pre-test probability of malignant thyroid nodules is about 10%. It seems that if the test result is positive after the non-invasive DECT examination, the probability of malignancy of thyroid nodules will be increased by about three times from 10% to 31%. If the result is negative, the probability of having malignant thyroid nodules decreases from 10% to 3% (Figure 7).

3.3.5. Heterogeneity Exploration

Meta-regression analyses were performed separately for each group (IC, NIC, λ_{HU}). It was found that all the covariates (publication year, study design, country of origin, CT manufacturers, CM Volume, CM Type, Injection Rate) did not have a significant effect on the SDOR ($p > 0.05$). As shown in the inconsistency index (I^2), there was little heterogeneity among the specificities of IC, NIC, and λ_{HU} , while moderate heterogeneity was observed in the sensitivities of IC, NIC, and λ_{HU} .

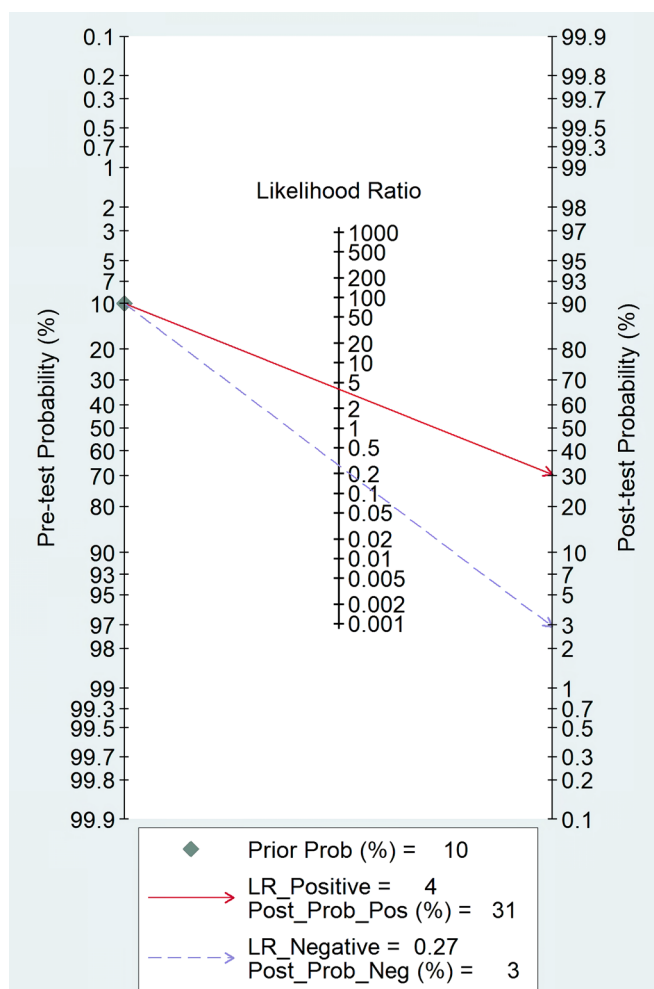


Figure 7. Fagan nomogram using the NIC technique to predict the post-test probability of malignant thyroid nodules in patients with thyroid nodules. As shown, a positive result increases the probability from 10% to 31%, while a negative result decreases the probability from 10% to 3%.

4. Discussion

As far as we know, a few studies have attempted to explore the relationship between thyroid nodules and diverse quantitative parameters [12] [22] [23] [36] [39]. One of the previous effective methods using conventional single-energy CT is to evaluate the size of nodules to distinguish benign from malignant thyroid nodules quantitatively. But, the correlation between the size of nodules and the nature of nodules is not ideal [12]. Also, for non-contrast DECT images, Tomita *et al.* found that it is unfeasible to differentiate between benign and malignant thyroid nodules using material decomposition and attenuation slopes [39]. Interestingly, our meta-analysis found that quantitative DECT parameters have good diagnostic accuracy in evaluating benign and malignant thyroid nodules, especially using contrast-enhanced images.

The collection of IC, NIC, and λ_{HU} are relatively sufficient data at present. The diagnostic ability of these three parameters seems similar. The diagnostic effi-

ciency of NIC (AUC = 0.85) is almost the same as that of IC (AUC = 0.84), and both of them are slightly higher than that of the λ_{HU} (AUC = 0.81). One possible reason might be that all these measurements could reflect specific element concentration in thyroid mass, and the injection of the contrast media increases the difference of iodine concentration between benign and malignant nodules. In addition, the NIC value means the normalizing IC value, which is obtained by taking both the IC of thyroid mass and the IC of ipsilateral common carotid artery into account. This process might reduce the effect of inconsistent scan delay between studies and improve the homogeneity.

Benign and malignant thyroid nodules differ in iodine uptake ability and tissue composition for the cytological and metabolic changes [21] [40] [41] [42]. Li *et al.* found follicular epithelial cells with abnormal formation in papillary carcinoma, while the follicular structure and iodine-uptake ability was normal in nodular goiters and adenomas [21]. Studies had also demonstrated that the expression of thyrocyte markers like sodium-iodide symporter (NIS) and thyroglobulin (TG) in thyroid cancer were less than those in benign lesions [41] [42] [43]. Among these, the decrease of follicular cells and iodine transporter on the cell membrane might closely correlate with the disfunction of uptake iodine. Based on our estimation, the NIC values of malignant nodules were significantly lower on contrast-enhanced CT images than that of benign nodules ($p < 0.01$), which may support the disfunction of iodine uptake in malignant thyroid nodules.

When x-ray passing through a material, its photon energy is absorbed proportionally. And the remaining energy is collected by CT detectors and will be converted to a specific signal, which finally represents the X-ray energy attenuation values on this material, the so-called Hounsfield units (HU) [17] [44]. The spectrum curve is, essentially, the plot of HU against various X-ray energy levels (generally within the range of 40 to 140 keV). In some instances, two materials (e.g., calcium and iodine) can exhibit the same HU due to the similar attenuation coefficients at a given energy level. However, the material's attenuation coefficient depends primarily on the mass density, composition, and the photon energies interacting with the material. After measuring an additional attenuation at a second energy level, the HU's change helps to differentiate between two materials [17]. The slope of spectrum curve (λ_{HU}) was used to depict the spectrum curve. Previous *ex vivo* study reported that the λ_{HU} measured in benign nodules (nodular goiters and follicular adenomas) was positive, whereas in malignant nodules (papillary carcinoma) λ_{HU} was negative [21]. But our included studies showed that all of the λ_{HU} were positive in the venous phase, which was inconsistent with the previous *ex vivo* findings. And the absolute value of λ_{HU} was larger in benign nodules than in malignant nodules according to our analysis.

As a previous study stated, different DECT data obtained by different DECT devices might have some heterogeneity, which requires caution when comparing various studies [45]. Our comprehensive study contains three kinds of DECT

devices: five dual-energy gemstone spectral CTs, one dual-layer spectral CT, and two second-generation dual-source dual-energy CTs. However, our research found that the heterogeneity of NIC's specificity was not significant, and the heterogeneity of NIC's sensitivity was also acceptable, for it could be explained, to some extent, by the threshold effect. This evidence indicates that the measurement of NIC in the venous phase might be comparable between different CT machines.

Moreover, we did not perform the same estimation in the groups of IC and λ_{HU} due to the uninterpretable heterogeneity. The heterogeneity of IC and λ_{HU} groups may not mainly originate from the threshold effect but from other factors. For example, different imaging phases may cause considerable heterogeneity shown in the mean difference of diagnostic efficacy ($I^2 = 73\%$, $I^2 = 95\%$ respectively). By meta-regression analysis of three parameters, however, the known covariates did not show a significant influence on the SDOR. One possibility might be that the heterogeneity is low. But it is more likely that the literature's insufficiency makes it difficult to use the known variables to explain the heterogeneity in existing studies. Nevertheless, we should not ignore that the present analysis does not consider some important variables, such as ages, type of nodules, size of nodules, the timing of the contrast injection, and the trigger threshold of CT scan.

5. Conclusion

DECT has a good diagnostic accuracy in thyroid nodule evaluation. The normalized iodine quantification technique is comparable between different dual-energy CT machines in thyroid nodule assessment. Quantitative DECT parameter may be a potential tool to estimate thyroid nodules noninvasively, reducing the need for unnecessary examinations of nodules, the associated patient anxiety, and inconvenience. More extensive prospective design studies are necessary to further optimize the diagnostic efficacy of quantitative DECT parameters.

Limitations

This study has some limitations. First, it has been pointed out that the incidence of thyroid cancer is increasing while thyroid nodules may be overdiagnosed, which could lead to overtreatment [6] [45]. We did not focus on whether we should add extra diagnostic methods like DECT to the thyroid examination, though it is still a matter of concern. Second, the thyroid is radiation sensitive. Presently, there are still some challenges in terms of radiation safety, though the radiation dose could be reduced by up to 50% in DECT [46]. Third, because the application of DECT is still developing and spreading, available literature data are limited, which leads to a relatively small number of studies meeting our inclusion criteria. And we did not estimate the combined diagnostic thresholds due to the heterogeneity among studies. Fourth, the category of thyroid nodules included in the study might need further subdivisions. Although various patho-

logical types of nodules were collected in the paper, no study proposed to crosslink parameters and the specific types of thyroid nodules (e.g., suspicious malignant nodules in FNA) that might occur progressive events, such as nodules enlargement or tendency for lymph node metastasis. And it is worth noting that it may be challenging to accurately implement iodine value measurement for micro thyroid lesions (maximum diameter < 10 mm), due to the partial volume effect in CT imaging. Of the studies we included, only two studies contained nodules less than 10 mm [36] [37], which implies that the quantitative diagnosis of sub-centimeter nodules still requires further study. Fifth, in terms of data statistics, some data are provided in the form of median and quartile but not mean and stand deviation [36]. So we estimated some data by combining median and quartile with original data size [47] [48], which may bring into errors when estimating the mean and standard deviation.

Acknowledgements

This research was supported by the National Natural Science Foundation of China: 21317241, 81971672, Key Program of Natural Science Foundation of Guangdong Province of China: 2018B0303110011, Guangzhou Key Laboratory of Molecular and Functional Imaging for Clinical Translation: 201905010003, the Fundamental Research Funds for the Central Universities: 21620308, 21620101, and the Engineering Research Center of Medical Imaging Artificial Intelligence for Precision Diagnosis and Treatment, Guangdong Province.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Parsa, A.A. and Gharib, H. (2018) Chapter 1 Epidemiology of Thyroid Nodules. In: *Thyroid Nodules*, Springer International Publishing, Cham, 1-11. https://doi.org/10.1007/978-3-319-59474-3_1
- [2] Pereira, M., Williams, V.L., Hallanger Johnson, J. and Valderrabano, P. (2020) Thyroid Cancer Incidence Trends in the United States: Association with Changes in Professional Guideline Recommendations. *Thyroid*, **30**, 1132-1140. <https://doi.org/10.1089/thy.2019.0415>
- [3] Feng, R.M., Zong, Y.N., Cao, S.M. and Xu, R.H. (2019) Current Cancer Situation in China: Good or Bad News from the 2018 Global Cancer Statistics? *Cancer Communications*, **39**, 22. <https://doi.org/10.1186/s40880-019-0368-6>
- [4] Burman, K.D. and Wartofsky, L. (2015) Clinical Practice. Thyroid Nodules. *The New England Journal of Medicine*, **373**, 2347-2356. <https://doi.org/10.1056/NEJMcp1415786>
- [5] Bomeli, S.R., LeBeau, S.O. and Ferris, R.L. (2010) Evaluation of a Thyroid Nodule. *Otolaryngologic Clinics of North America*, **43**, 229-238. <https://doi.org/10.1016/j.otc.2010.01.002>
- [6] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A. and Jemal, A. (2018)

- Global Cancer Statistics 2018: Globocan Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, **68**, 394-424. <https://doi.org/10.3322/caac.21492>
- [7] Egset, A.V., Holm, C., Larsen, S.R., Nielsen, S.H., Bach, J., Helweg-Larsen, J.P., *et al.* (2017) Risk of Malignancy in Fine-Needle Aspiration Biopsy in Patients with Thyroid Nodules. *Danish Medical Journal*, **64**, A5320.
- [8] Cappelli, C., Pirola, I., Agosti, B., Tironi, A., Gandossi, E., Incardona, P., *et al.* (2017) Complications after Fine-Needle Aspiration Cytology: A Retrospective Study of 7449 Consecutive Thyroid Nodules. *British Journal of Oral and Maxillofacial Surgery*, **55**, 266-269. <https://doi.org/10.1016/j.bjoms.2016.11.321>
- [9] Antoni, M.H. and Dhabhar, F.S. (2019) The Impact of Psychosocial Stress and Stress Management on Immune Responses in Patients with Cancer. *Cancer*, **125**, 1417-1431. <https://doi.org/10.1002/cncr.31943>
- [10] Moon, W.J., Jung, S.L., Lee, J.H., Na, D.G., Baek, J.H., Lee, Y.H., *et al.* (2008) Benign and Malignant Thyroid Nodules: US Differentiation-Multicenter Retrospective Study. *Radiology*, **247**, 762-770. <https://doi.org/10.1148/radiol.2473070944>
- [11] Moon, H.J., Sung, J.M., Kim, E.K., Yoon, J.H., Youk, J.H. and Kwak, J.Y. (2012) Diagnostic Performance of Gray-Scale US and Elastography in Solid Thyroid Nodules. *Radiology*, **262**, 1002-1013. <https://doi.org/10.1148/radiol.11110839>
- [12] Hoang, J.K., Langer, J.E., Middleton, W.D., Wu, C.C., Hammers, L.W., Cronan, J.J., *et al.* (2015) Managing Incidental Thyroid Nodules Detected on Imaging: White Paper of the ACR Incidental Thyroid Findings Committee. *Journal of the American College of Radiology*, **12**, 143-150. <https://doi.org/10.1016/j.jacr.2014.09.038>
- [13] Baer, M., Maaß, C., Kalender, W.A. and Kachelrieß, M. (2009) Image-Based Dual Energy CT Using Optimized Precorrection Functions: A Practical New Approach to Material Decomposition in the Image Domain. In: *Proceedings of World Congress on Medical Physics and Biomedical Engineering*, Springer, Berlin, 205-208. https://doi.org/10.1007/978-3-642-03879-2_58
- [14] Forghani, R., Srinivasan, A. and Forghani, B. (2017) Advanced Tissue Characterization and Texture Analysis Using Dual-Energy Computed Tomography: Horizons and Emerging Applications. *Neuroimaging Clinics of North America*, **27**, 533-546. <https://doi.org/10.1016/j.nic.2017.04.007>
- [15] Goo, H.W. and Goo, J.M. (2017) Dual-Energy CT: New Horizon in Medical Imaging. *Korean Journal of Radiology*, **18**, 555-569. <https://doi.org/10.3348/kjr.2017.18.4.555>
- [16] Liu, X., Yu, L., Primak, A.N. and McCollough, C.H. (2009) Quantitative Imaging of Element Composition and Mass Fraction Using Dual-Energy CT: Three-Material Decomposition. *Medical Physics*, **36**, 1602-1609. <https://doi.org/10.1118/1.3097632>
- [17] McCollough, C.H., Leng, S., Yu, L. and Fletcher, J.G. (2015) Dual- and Multi-Energy CT: Principles, Technical Approaches, and Clinical Applications. *Radiology*, **276**, 637-653. <https://doi.org/10.1148/radiol.2015142631>
- [18] Mendonca, P.R., Lamb, P. and Sahani, D.V. (2014) A Flexible Method for Multi-Material Decomposition of Dual-Energy CT Images. *IEEE Transactions on Medical Imaging*, **33**, 99-116. <https://doi.org/10.1109/TMI.2013.2281719>
- [19] Marin, D., Davis, D., Roy Choudhury, K., Patel, B., Gupta, R.T., Mileto, A., *et al.* (2017) Characterization of Small Focal Renal Lesions: Diagnostic Accuracy with Single-Phase Contrast-Enhanced Dual-Energy CT with Material Attenuation Analysis Compared with Conventional Attenuation Measurements. *Radiology*, **284**, 737-747. <https://doi.org/10.1148/radiol.2017161872>

- [20] Liang, D.Z., Zhu, D. and Liu, Y.J. (2014) Clinical Study on Thyroid Iodine Concentration with Gemstone Spectral CT. *Chinese Journal of Anatomy and Clinics*, **19**, 33-36.
- [21] Li, M., Zheng, X., Li, J., Yang, Y., Lu, C., Xu, H., *et al.* (2012) Dual-Energy Computed Tomography Imaging of Thyroid Nodule Specimens: Comparison with Pathologic Findings. *Investigative Radiology*, **47**, 58-64. <https://doi.org/10.1097/RLI.0b013e318229fef3>
- [22] Gao, S.Y., Zhang, X.Y., Wei, W., Li, X.T., Li, Y.L., Xu, M., *et al.* (2016) Identification of Benign and Malignant Thyroid Nodules by *in Vivo* Iodine Concentration Measurement Using Single-Source Dual Energy CT: A Retrospective Diagnostic Accuracy Study. *Medicine (Baltimore)*, **95**, e4816. <https://doi.org/10.1097/MD.0000000000004816>
- [23] Lee, D.H., Lee, Y.H., Seo, H.S., Lee, K.Y., Suh, S.I., Ryoo, I., *et al.* (2019) Dual-Energy CT Iodine Quantification for Characterizing Focal Thyroid Lesions. *Head Neck*, **41**, 1024-1031. <https://doi.org/10.1002/hed.25524>
- [24] Frank, R.A., Bossuyt, P.M. and McInnes, M.D.F. (2018) Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy: The PRISMA-DTA Statement. *Radiology*, **289**, 313-314. <https://doi.org/10.1148/radiol.2018180850>
- [25] McInnes, M.D.F., Moher, D., Thombs, B.D., McGrath, T.A., Bossuyt, P.M., *et al.* (2018) Preferred Reporting Items for a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *JAMA*, **319**, 388-396. <https://doi.org/10.1001/jama.2017.19163>
- [26] Whiting, P.F., Rutjes, A.W., Westwood, M.E., Mallett, S., Deeks, J.J., Reitsma, J.B., *et al.* (2011) QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Annals of Internal Medicine*, **155**, 529-536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
- [27] McInnes, M.D. and Bossuyt, P.M. (2015) Pitfalls of Systematic Reviews and Meta-Analyses in Imaging Research. *Radiology*, **277**, 13-21. <https://doi.org/10.1148/radiol.2015142779>
- [28] Leeflang, M.M., Deeks, J.J., Gatsonis, C., Bossuyt, P.M. and Cochrane Diagnostic Test Accuracy Working Group (2008) Systematic Reviews of Diagnostic Test Accuracy. *Annals of Internal Medicine*, **149**, 889-897. <https://doi.org/10.7326/0003-4819-149-12-200812160-00008>
- [29] Higgins, J.P., Thompson, S.G., Deeks, J.J. and Altman, D.G. (2003) Measuring Inconsistency in Meta-Analyses. *BMJ*, **327**, 557-560. <https://doi.org/10.1136/bmj.327.7414.557>
- [30] Zamora, J., Abraira, V., Muriel, A., Khan, K. and Coomarasamy, A. (2006) Meta-DiSc: A Software for Meta-Analysis of Test Accuracy Data. *BMC Medical Research Methodology*, **6**, 31. <https://doi.org/10.1186/1471-2288-6-31>
- [31] Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T.J., Page, M.J., *et al.* (2019) *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons, Hoboken. <https://doi.org/10.1002/9781119536604>
- [32] Dwamena, B. (2009) MIDAS: Stata Module for Meta-Analytical Integration of Diagnostic Test Accuracy Studies. <https://EconPapers.repec.org/RePEc:boc:bocode:s456880>
- [33] Cheng, Y. (2019) Application of Energy Spectrum CT Parameters in Differential Diagnosis of Benign and Malignant Thyroid Lesions. *Chinese Journal of CT and MRI*, **17**, 26-28.

- [34] Zhang, Y., Chen, J. and Chen, Z.G. (2017) Clinical Application of Dual Energy Scanning Technology with Dual-Source CT in Differentiating Benign from Malignant Thyroid Nodules. *Journal of Practical Radiology*, **33**, 455-459.
- [35] Zhao, W., Zhang, Z.H., Han, D., Yang, J., Liu, C.R. and Zhao, W. (2015) Iodine Concentration on Dual-Energy Contrast Enhanced CT in Differential Diagnosis of Thyroid Nodules. *Chinese Journal of Medical Imaging*, **11**, 815-818.
- [36] Li, H.W., Wu, X.W., Liu, B., Liu, W.D. and Gao, N. (2015) Clinical Values of Gemstone Spectral CT in Diagnosing Thyroid Disease. *Journal of X-Ray Science and Technology*, **23**, 45-56. <https://doi.org/10.3233/XST-140469>
- [37] He, L., Jia, X.C., Liu, R.H., Sun, J.L. and Zhang, S.Q. (2014) Diagnostic Value of Spectral CT Imaging in Papillary Thyroid Carcinoma. *Diagnostic Imaging & Interventional Radiology*, **6**, 506-510.
- [38] Xue, Y.J., Duan, Q., Sun, B., Chen, L.H. and Ge, H.T. (2013) Clinical Application of CT Gemstone Spectral Imaging in Distinguishing Malignant from Benign Primary Thyroid Lesions. *Chinese Journal of Medical Imaging Technology*, **1**, 30-33.
- [39] Tomita, H., Kuno, H., Sekiya, K., Otani, K., Sakai, O., Li, B., *et al.* (2020) Quantitative Assessment of Thyroid Nodules Using Dual-Energy Computed Tomography: Iodine Concentration Measurement and Multiparametric Texture Analysis for Differentiating between Malignant and Benign Lesions. *International Journal of Endocrinology*, **2020**, Article ID: 5484671. <https://doi.org/10.1155/2020/5484671>
- [40] Lazar, V., Bidart, J.M., Caillou, B., Mahe, C., Lacroix, L., Filetti, S., *et al.* (1999) Expression of the Na⁺/I-Symporter Gene in Human Thyroid Tumors: A Comparison Study with Other Thyroid-Specific Genes. *The Journal of Clinical Endocrinology & Metabolism*, **84**, 3228-3234. <https://doi.org/10.1210/jc.84.9.3228>
- [41] Ringel, M.D., Anderson, J., Souza, S.L., Burch, H.B., Tambascia, M., Shriver, C.D., *et al.* (2001) Expression of the Sodium Iodide Symporter and Thyroglobulin Genes Are Reduced in Papillary Thyroid Cancer. *Modern Pathology*, **14**, 289-296. <https://doi.org/10.1038/modpathol.3880305>
- [42] Sodre, A.K., Rubio, I.G., Galrao, A.L., Knobel, M., Tomimori, E.K., Alves, V.A., *et al.* (2008) Association of Low Sodium-Iodide Symporter Messenger Ribonucleic Acid Expression in Malignant Thyroid Nodules with Increased Intracellular Protein Staining. *The Journal of Clinical Endocrinology & Metabolism*, **93**, 4141-4145. <https://doi.org/10.1210/jc.2007-0353>
- [43] Lakshmanan, A., Scarberry, D., Shen, D.H. and Jhiang, S.M. (2014) Modulation of Sodium Iodide Symporter in Thyroid Cancer. *Hormones and Cancer*, **5**, 363-373. <https://doi.org/10.1007/s12672-014-0203-0>
- [44] Yu, L., Leng, S. and McCollough, C.H. (2012) Dual-Energy CT-Based Monochromatic Imaging. *American Journal of Roentgenology*, **199**, S9-S15. <https://doi.org/10.2214/AJR.12.9121>
- [45] Jacobsen, M.C., Schellingerhout, D., Wood, C.A., Tamm, E.P., Godoy, M.C., Sun, J., *et al.* (2018) Intermanufacturer Comparison of Dual-Energy CT Iodine Quantification and Monochromatic Attenuation: A Phantom Study. *Radiology*, **287**, 224-234. <https://doi.org/10.1148/radiol.2017170896>
- [46] Chai, X., Zhang, L.J., Lu, G.M. and Zhou, C.S. (2009) Dual Energy CT Intracranial Angiography: Image Quality, Radiation Dose and Initial Application Results. *Chinese Journal of Radiology*, **43**, 725-729.
- [47] Wan, X., Wang, W.Q., Liu, J.M. and Tong, T.J. (2014) Estimating the Sample Mean and Standard Deviation from the Sample Size, Median, Range and/or Interquartile Range. *BMC Medical Research Methodology*, **14**, 135.

<https://doi.org/10.1186/1471-2288-14-135>

- [48] Luo, D.H., Wan, X., Liu, J.M. and Tong, T.J. (2018) Optimally Estimating the Sample Mean from the Sample Size, Median, Mid-Range, and/or Mid-Quartile Range. *Statistical Methods in Medical Research*, **27**, 1785-1805.

<https://doi.org/10.1177/0962280216669183>

Supplementary Materials

The following are available online, Full search strategy, **Table S1**: diagnostic performance of DECT parameters, **Table S2**: PRISMA-DTA checklist, **Table S3**: PRISMA-DTA for Abstracts Checklist.

Full search strategy in PubMed

1# DECT

2# dual-energy CT

3# dual-energy computed tomography

4# thyroid

5# 1# OR 2# OR 3#

6# 4# AND 5#

OR:

((DECT) OR (dual-energy CT)) OR (dual-energy computed tomography)

AND (thyroid)

Full search strategy in Web of Science

TS = (DECTOR dual-energy CT OR dual-energy computed tomography)

AND TS = thyroid

Full search strategy in Embase

(DECT OR (“dual energy” AND CT) OR (“dual energy” AND computed AND tomography)) AND thyroid

Table S1. Diagnostic performance of DECT parameters.

Method	Author	Year	No. of Nodules					
			TP	FP	FN	TN	Benign	Malignant
λ_{HU}	Xue <i>et al.</i>	2013	27	9	6	38	47	33
	Cheng <i>et al.</i>	2019	42	9	21	29	38	63
	He <i>et al.</i>	2014	37	6	8	24	30	45
	Zhang <i>et al.</i>	2017	85	16	12	40	56	97
	Total						171	238
NIC	Li <i>et al.</i>	2015	13	6	3	19	25	16
	Xue <i>et al.</i>	2013	26	11	7	36	47	33
	Lee <i>et al.</i>	2018	30	2	16	15	17	46
	Zhao <i>et al.</i>	2015	45	5	15	25	30	60
	Zhang <i>et al.</i>	2017	86	13	11	43	56	97
Total						175	252	
IC	Li <i>et al.</i>	2015	26	29	9	58	87	35
	Gao <i>et al.</i>	2016	24	4	2	29	33	26
	Cheng <i>et al.</i>	2019	46	11	17	27	38	63
	He <i>et al.</i>	2014	38	5	7	25	30	45
	Lee <i>et al.</i>	2018	32	5	14	12	17	46
Total						205	215	

Abbreviations: λ_{HU} , Slope of spectrum curve; NIC, normalized iodine concentration; IC, iodine concentration; TP, true positive; FP, false positive; FN, false negative; TN, true negative.

Table S2. PRISMA-DTA checklist.

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
TITLE/ABSTRACT			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	1
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	2
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	2
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	2
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Suppl 3-7
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	2-3
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	2-3
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	2-3
Study selection	9	State the process for selecting studies (<i>i.e.</i> , screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	3
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	3
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	3
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	3
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	3
Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity, c) handling multiple index test readers, d) handling of indeterminate test results, e) grouping and comparing tests, f) handling of different reference standards	4
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	4
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	4
RESULTS			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	5, Figure 1
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	8-9, Table 1, Table 2

Continued

Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	5-7, Figure 1, Figure 2
Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2 × 2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	10-12, Figures 4-6, Table S1
Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	9-12, Table 3,
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	11-13, Figure 7
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	13-14
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	15-16
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	15
FUNDING			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	16

Table S3. PRISMA-DTA for abstracts checklist.

Section/topic	#	PRISMA-DTA for Abstracts Checklist Item	Reported on page #
TITLE and PURPOSE			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	1
Objectives	2	Indicate the research question, including components such as participants, index test, and target conditions.	1
METHODS			
Eligibility criteria	3	Include study characteristics used as criteria for eligibility.	1
Information sources	4	List the key databases searched and the search dates.	1
Risk of bias & applicability	5	Indicate the methods of assessing risk of bias and applicability.	1
Synthesis of results	A1	Indicate the methods for the data synthesis.	1
RESULTS			
Included studies	6	Indicate the number and type of included studies and the participants and relevant characteristics of the studies (including the reference standard).	1
Synthesis of results	7	Include the results for the analysis of diagnostic accuracy, preferably indicating the number of studies and participants. Describe test accuracy including variability; if meta-analysis was done, include summary results and confidence intervals.	1
DISCUSSION			

Continued

Strengths and limitations	9	Provide a brief summary of the strengths and limitations of the evidence	1
Interpretation	10	Provide a general interpretation of the results and the important implications.	1
OTHER			
Funding	11	Indicate the primary source of funding for the review.	-
Registration	12	Provide the registration number and the registry name.	-
