BMJ Open

Artificial intelligence-based CT-free quantitative thyroid SPECT for thyrotoxicosis: study protocol of a multicentre, prospective, non-inferiority study

Dongkyu Oh ,[1,2] Hyun Gee Ryoo ,[1] Hyun Woo Chung ,[3] Sang-Geon Cho ,[4] Kyounghyoun Kwon ,[1,5] Ji Hye Kim ,[1] Young So ,[3] Jae Hoon Moon ,[6] Soyeon Ahn ,[7] Won Woo Lee [1,2,5,8]

To cite: Oh D, Ryoo HG, Chung HW, *et al*. Artificial intelligence-based CT-free quantitative thyroid SPECT for thyrotoxicosis: study protocol of a multicentre, prospective, non-inferiority study. *BMJ Open* 2024;**14**:e089552. doi:10.1136/bmjopen-2024-089552

► Prepublication history for this paper is available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2024-089552).

DO and HGR contributed equally.

Received 02 June 2024
Accepted 24 September 2024

For numbered affiliations see end of article.

Correspondence to
Dr Won Woo Lee;
wwlee@snubh.org

ABSTRACT

Introduction Technetium thyroid uptake (TcTU) measured by single-photon emission CT/CT (SPECT/CT) is an important diagnostic tool for the differential diagnosis of Graves' disease and destructive thyroiditis. Artificial intelligence (AI) may reduce CT-induced radiation exposure by substituting the role of CT in attenuation correction (AC) and thyroid segmentation, thus realising CT-free SPECT. This study aims to compare the diagnostic accuracy for the differential diagnosis of thyrotoxicosis between CT-free SPECT and SPECT/CT.

Methods and analysis The AI-based CT-free SPECT is a single-blind, multicentre, prospective, non-inferiority, clinical trial with a paired design conducted in the Republic of Korea. Eligible participants are adult (≥19 years old) thyrotoxicosis patients without a previous history of hyperthyroidism or hypothyroidism. Approximately 160 subjects will be screened for quantitative thyroid SPECT/CT using Tc-99m pertechnetate. CT-free thyroid SPECT will be realised using only SPECT data by the trained convolutional neural networks. TcTU will be calculated by SPECT/CT and CT-free SPECT in each subject. The primary endpoint is the accuracy of diagnosing Graves' disease using TcTU. The trial will continue until 152 completed datasets have been enrolled to assess whether the 95% (two-sided) lower confidence limit of the accuracy difference (CT-free SPECT accuracy—SPECT/CT accuracy) for Graves' disease is greater than −0.1. The secondary endpoints include the accuracy of diagnosing destructive thyroiditis and predicting the need for antithyroid drug prescription within 1 month of the SPECT/CT.

Ethics and dissemination The study protocol has been approved by the institutional review board of Seoul National University Bundang Hospital (IRB No. B-2304-824-301), Konkuk University Medical Center (IRB No. 2023-05-022-006) and Chonnam National University Hospital (IRB No. CNUH-2023-108). Findings will be disseminated as reports, presentations and peer-reviewed journal articles.

Trial registration number KCT0008387, Clinical Research Information Service of the Republic of Korea (CRIS).

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This is a multicentre prospective study focused on the clinical application of artificial intelligence (AI).
⇒ In this study, the AI application automates the measurement of thyroid uptake parameter.
⇒ Patient radiation exposure from CT is eliminated with the aid of AI.
⇒ As a limitation, the study includes only thyrotoxicosis patients who meet specific inclusion criteria, which limits its ability to represent the broad spectrum of thyroid diseases that require thyroid single-photon emission CT imaging.

INTRODUCTION

It is crucial to distinguish between Graves' disease and destructive thyroiditis in patients with thyrotoxicosis because their respective treatments are significantly different. Graves' disease is managed by antithyroid drugs, while destructive thyroiditis is often handled conservatively, often without specific medications. Not all cases of Graves' disease display the specific symptoms/signs of the disease,[1] and biochemical laboratory findings are frequently inconclusive.[2–4] Nuclear medicine tests that measure the technetium thyroid uptake (TcTU) have been widely used for the differential diagnosis of thyrotoxicosis.[5 6] The TcTU is increasingly replacing the traditional radioactive iodine uptake (RAIU) because TcTU is easier to measure and renders patients to less radiation exposure than RAIU. Moreover, in addition to the quantitative parameter of TcTU, a thyroid scan using the same radioisotope (Tc-99m pertechnetate) often provides valuable information about radioactivity distribution, facilitating the diagnosis of toxic adenoma or multinodular goitre.

Traditionally, TcTU has been measured using a one-dimensional thyroid uptake system or two-dimensional planar gamma camera imaging. Now, three-dimensional single-photon emission CT/CT (SPECT/CT) is emerging as a promising alternative to conventional methods. The SPECT/CT allows thyroid function assessment in both quantitative and qualitative ways in a single session.[7–10] Quantitative SPECT/CT has proven to be the most accurate for the differential diagnosis of thyrotoxicosis, compared with the thyroid uptake system or planar scan.[7 11] Hence, the utility of quantitative SPECT/CT is expanding to other hyperfunctioning or hypofunctioning thyroid situations, such as predicting intractable Graves' disease, meticulously differentiating between thyroiditis and euthyroid and planning the treatment of autonomously functioning thyroid nodules.[8–10 12]

The quantitative SPECT/CT requires triple corrections for radioactivity. Attenuation correction (AC) using an attenuation map (μ-map) is the first correction, and the μ-map is derived from CT. Scatter correction (SC) and collimator-detector response correction (resolution recovery (RR)) are also necessary corrections, fulfilled by vendor-specific software.[13 14] In addition, segmentation of the organ of interest on the CT canvas is a critical process for quantitative SPECT/CT.[15 16] Therefore, CT plays a pivotal role in AC and organ segmentation in quantitative SPECT/CT.

The use of CT has sparked debate in medical imaging due to potential radiation hazards. The as-low-as-reasonably-achievable principle necessitates the optimisation of radiation exposure and asks for possible alternatives to CT. In this respect, CT-free SPECT imaging techniques based on artificial intelligence (AI) hold significant promise.[17 18] AI applications have facilitated AC through direct or indirect generation of attenuation-corrected SPECT images without the aid of CT. In a further step, we conducted an AI study where not only was AC realised but organ segmentation was also achieved without CT, effectively replacing 'thyroid SPECT/CT' with 'CT-free thyroid SPECT'.[19]

This prospective, multicentre study was designed to promote CT-free SPECT technology. The key question was whether the AI-based CT-free thyroid SPECT is not inferior to the conventional thyroid SPECT/CT in terms of differential diagnosis of thyrotoxicosis. We investigated the accuracy of diagnosing Graves' disease and destructive thyroiditis using TcTU from both CT-free SPECT and SPECT/CT perspectives.

DIFFERENTIAL DIAGNOSIS OF THYROTOXICOSIS

Typical clinical symptoms and signs

Patients with excessive endogenous thyroid hormones can suffer serious complications from thyrotoxicosis, including cardiovascular or neuromuscular complications.[20] Excessive thyroid hormone can stem from either productive causes, such as Graves' disease, or destructive ones, such as painless thyroiditis, subacute thyroiditis and postpartum thyroiditis. Common signs of Graves' disease include goitre, orbitopathy and pretibial dermopathy. While there is no concrete study about the prevalence of these signs among Graves' disease patients, estimations are possible. In interviews with five renowned American endocrinologists, they reported a 70% prevalence (65%–75%) of these signs in cases of overt hyperthyroidism.[1] Additionally, approximately 50% of patients with Graves' disease were reported to exhibit orbitopathy.[21] With the widespread availability of thyroid hormone level tests, clinically severe Graves' disease, or its typical toxic appearance, is less frequently detected nowadays. The T3/T4 ratio may increase in Graves' disease[22] but decrease in destructive thyroiditis.[23] However, the performance of these ratios was found to be unsatisfactory in subsequent studies.[24 25] Thyroid-stimulating hormone (TSH) is often employed as a highly sensitive indicator of thyrotoxicosis, resulting in more frequent encounters with mild subclinical hyperthyroidism with suppressed TSH levels. As a result, relying solely on clinical signs or symptoms is not sufficient for differential diagnosis of thyrotoxicosis. Additional supportive measurements are required in the modern practice of managing thyrotoxicosis.

TSH receptor antibody (TSHRAb) measurement

TSHRAb, especially the TSH receptor stimulating immunoglobulin (TSI), is a pathognomonic autoantibody of Graves' disease. The TSI level is determined by a cell-based bioassay for intracellular cyclic adenosine monophosphate production, which has been proven to be highly sensitive, specific and cost-effective for Graves' disease.[1 26] However, the TSI bioassay is not widely used due to its complexity and the considerable laboratory expertise it requires, despite its superior diagnostic ability compared with other methods.[27 28] Currently, most TSHRAb assays are performed using a competitive-binding assay, which measures TSH receptor binding inhibitory immunoglobulin (TBII). The TBII assay has evolved to the third generation, characterised by competitive inhibition of monoclonal anti-human TSH receptor antibodies or labelled TSH binding to recombinant TSH receptors.[29] At present, this third-generation TBII assay is recommended for differential diagnosis of thyrotoxicosis.[20]

Despite TBII's reported high sensitivity and specificity for diagnosing Graves' disease,[30] several pitfalls exist. First, TBII does not differentiate between stimulatory, blocking and neutral immunoglobulins.[31] Second, population-based studies in various countries demonstrated that the TSHRAb positivity in Graves' disease is approximately 80%,[2 32–35] contrasting with the higher sensitivity and specificity (~99% each) seen in blood sample-based studies in highly selected patients.[36] Third, painless thyroiditis, a type of destructive thyroiditis most similar to Graves' disease, reportedly has a high incidence of TSHRAb positivity, up to 18.9%[37] or 22.4%.[38] Fourth, TSHRAb may be consistently negative in up to 10% of mild cases of Graves' disease.[39–41] Fifth, instead of a discrete cut-off level, a grey zone is suggested to exist between Graves'

disease and destructive thyroiditis, leading to interlaboratory variability.[42] Sixth, third-generation TSHRAb test kits are not available in all institutes/hospitals, even in well-developed countries.[43] In light of these factors, considering the sensitivity and specificity, the estimated clinical impact of the TSHRAb test for diagnosing Graves' disease in real-world settings was 85%.[1]

Colour Doppler ultrasonography

Blood flow is increased in hyperfunctioning thyroid diseases such as Graves' disease and autonomously functioning thyroid nodules. Colour Doppler ultrasonography was able to evaluate the increased peak systolic velocity in the intrathyroidal or inferior/superior thyroidal arteries[44 45] or the increased parenchymal colour flow echogenicity.[46] However, the sonographic findings were never pathognomonic for Graves' disease or any other thyroid disease.[47] The presence of experienced operators was also a prerequisite for the proper use of ultrasonography. In particular, for pregnant or lactating women, ultrasonography could be an alternative to nuclear imaging studies.[48]

Radioactive iodine uptake (RAIU)

RAIU is still the first-line test for the differential diagnosis of thyrotoxicosis.[20] The high RAIU in Graves' disease and the low RAIU in destructive thyroiditis are so clear-cut in most clinical or investigational conditions that RAIU is still advocated as the gold standard biomarker for thyrotoxicosis differentiation. Many laboratory tests[37] or imaging studies[49 50] have been conducted with RAIU as a reference. Iodine is fixed within the thyroid cells, while Tc-99m pertechnetate is not. The continuous accumulation of iodine increases the RAIU to more than 50%, which is substantially greater than the 1%–2% for TcTU in euthyroid cases.

However, there are some limitations of RAIU. RAIU usually takes a longer time (~24 hours) than TcTU (20–30 min). The radiation exposure by RAIU is greater (11 mSv/MBq) than that by TcTU (0.013 mSv/MBq).[6] Thyroid scintigraphy using iodine-131 is impractical due to high-energy (364 keV) gamma rays, but thyroid scintigraphy after TcTU is almost routine practice thanks to the ideal gamma energy (140 keV) of Tc-99m. Iodine-123 may be an alternative to iodine-131, but iodine-123 is not easily available in many hospitals/institutes.

Technetium thyroid uptake (TcTU)

The uptake of Tc-99m pertechnetate in the thyroid is dependent on TSH or TSHRAb. The use of Tc-99m pertechnetate for thyrotoxicosis evaluation has twofold applications: one is the quantitative parameter about the per cent of injected uptake to the thyroid (ie, TcTU), and the other is qualitative thyroid scintigraphy for visual assessment. Both TcTU and visual assessment have been employed as a gold standard for Graves' disease when TSHRAb,[38 42 51] colour Doppler ultrasonography[47 52 53] or MRI[54] were investigated for the differential diagnosis of thyrotoxicosis. Sometimes, not TcTU but thyroid scintigraphy was considered sufficient for thyroid function assessment. However, the interpretation of thyroid scintigraphy is too subjective to be standardised.[55 56] Many institutes/hospitals have adopted TcTU instead of RAIU for functional assessment of the thyroid due to technical simplicity and lower radiation exposure to patients.[57] In a direct head-to-head comparison, TcTU was superior to TSHRAb[41] or RAIU[58] for the differential diagnosis of thyrotoxicosis. Thus, TcTU is the objective biomarker of thyroid function.

The technique of TcTU measurement has evolved from a one-dimensional thyroid uptake system (in other words, a probe) to a two-dimensional planar scan method, and finally, to three-dimensional SPECT/CT. The planar scan and SPECT/CT can enable both TcTU measurement and visual assessment with a single injection of Tc-99m pertechnetate. However, TcTU by a one-dimensional thyroid uptake system is contaminated by the saliva or salivary gland Tc-99m uptake.[7] Moreover, the accuracy of TcTU from three-dimensional SPECT/CT was superior to that from the two-dimensional uptake ratio of the planar scan for the differential diagnosis of thyrotoxicosis.[8–11] In general, SPECT/CT is evolving into the highly advanced imaging technology of nuclear medicine. Therefore, thyroid SPECT/CT using Tc-99m pertechnetate may be the most accurate technology for the functional assessment of the thyroid.

Nevertheless, the use of CT may be a source of concern due to radiation exposure. If the roles of CT in quantitative SPECT/CT can be fulfilled by other methods, it would be preferable. In this regard, the recent development of CT-free SPECT technology may be a solution to this issue if the diagnostic accuracy of the CT-free SPECT is not substantially lower compared with the conventional SPECT/CT. This potential has been suggested in studies with myocardial perfusion SPECT,[17 18] but it has not yet even explored in thyroid SPECT. In addition, CT-free SPECT is not labor intensive because thyroid segmentation is automatically performed by the trained networks, whereas human experts perform the same job for SPECT/CT.[19] The usefulness of AI applications has been tested in a variety of medical imaging fields, such as simple patient positioning during radiographic studies, artefact reduction, selection of optimal imaging protocols, etc.[59] CT dose reduction is one of those promising applications.[60] However, there are few studies that have proven the efficacy of AI applications in a prospective multi-institute setting.[61] In the current study, we aimed to prove the non-inferiority of CT-free thyroid SPECT in comparison with conventional SPECT/CT regarding the diagnostic accuracy for Graves' disease and destructive thyroiditis. The reduction in radiation exposure and savings in time and human resources would be benefits of CT-free SPECT if its non-inferiority to SPECT/CT is proven.

METHODS AND ANALYSIS

AI-based CT-free SPECT/CT (AISPECT) study objectives and design

The AISPECT study was registered on 10 April 2023 in the Clinical Research Information Service of the Republic of Korea (CRIS, https://cris.nih.go.kr as KCT0008387). This study is a self-controlled, single-blind, multicentre, prospective, non-inferiority, clinical trial. Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research. Adult patients (≥19 years old) with newly developed thyrotoxicosis will be the target population. The primary goal of the study is to investigate whether AI-based CT-free thyroid SPECT could provide TcTU values that are not inferior to those from conventional SPECT/CT in terms of the diagnosis of Graves' disease. For this purpose, approximately 160 subjects, including patients with Graves' disease or destructive thyroiditis, will be screened from three tertiary referral hospitals in South Korea. The subjects will undergo conventional thyroid SPECT/CT with Tc-99m pertechnetate (185 MBq) using SPECT/CT scanners from the same vendor (GE Healthcare). The quantitative SPECT/CT protocols are all the same, and the TcTU will be calculated after the SPECT/CT data are uploaded to the cloud platform (IRM, BEST Image; https://irmbestimage.com) for centralised data analysis. In a core lab at Seoul National University Bundang Hospital (SNUBH), TcTU will be calculated by a human for SPECT/CT and by AI for CT-free SPECT. The study will end when 152 complete paired datasets of SPECT/CT and CT-free SPECT are obtained, with an anticipated duration of 2 years. The secondary endpoints are the diagnostic accuracy for destructive thyroiditis and for the prediction of antithyroid drug prescription within 1 month of SPECT/CT. The study protocol was approved on 6 April 2023, which was also the planned start date, and the planned end date of the study is 8 July 2025.

Study population

Patients of this study were prospectively enrolled from 13 April 2023. To be included in the study, participants should be at least 19 years old and have thyrotoxicosis. A reduced level of TSH with elevated or normal range T3 (or free T3) and free T4 (or T4) is the main indicator of thyrotoxicosis in the current study. In a typical case scenario, subjects will be referred from primary or secondary hospitals for thyrotoxicosis evaluation to the three tertiary referral hospitals: SNUBH, Konkuk University Medical Center (KUMC) and Chonnam National University Hospital (CNUH). Initially, endocrinologists or family medicine doctors in the three hospitals will take care of the subjects, and thyroid SPECT/CT and other blood tests including TSHRAb are requested. When the subjects come to the department of nuclear medicine for thyroid SPECT/CT, they are screened for participation in the study. Table 1 shows the exclusion criteria which comprise previous hyperfunctioning or hypofunctioning thyroid diseases, concomitant other thyroid diseases, medication of thyroid-affecting agents and technical failure of AI operation. Approximately 2% of technical failure of AI operation was experienced in our previous tests (unpublished data). Therefore, a 5% dropout rate after screening is expected in the study. The schema of the case enrolment process is outlined in figure 1.

Table 1 Exclusion criteria

Pre-existing functional thyroid diseases	
Hyperthyroidism	History of antithyroid drug (propylthiouracil, methimazole, carbimazole) medication
Hypothyroidism	History of thyroid hormone (thyroxine, triiodothyronine) medication
Other thyroid diseases	History of thyroidectomy
	Coexisting thyroid cancer
	Intrathoracic goitre
Medications that affect thyroid	
	Iodine contrast-enhanced CT scan within 2 months
	History of lithium, amiodarone or immune-checkpoint inhibitors
Technically not feasible for AI operation	

AI, artificial intelligence.

SPECT/CT imaging

The subjects, who provide informed consent, will undergo quantitative thyroid SPECT/CT with Tc-99m pertechnetate (185 MBq intravenous injection). The general process of the quantitative SPECT/CT is available in the literature.[7] In brief, the system sensitivities between the dose calibrators and the SPECT/CT scanners are determined in advance. The protocols for SPECT/CT acquisition are harmonised to be exactly the same, but the acquisition protocols for planar scans are allowed to differ. In addition, the effective radiation dose of a thyroid CT scan in each subject will be calculated from the dose–length product value.

TcTU measurement by SPECT/CT and CT-free SPECT

TcTU will be measured in a core laboratory located at SNUBH. All the imaging data (sinograms for emission and scattering, NCRR (non-corrected for attenuation and scattering but with resolution recovery applied) SPECTs for emission and scattering, CT and planar scan) in digital imaging and communications in medicine (DICOM) format will be uploaded from the three participating hospitals to the cloud platform (IRM, BEST Image; https://irmbestimage.com) after anonymisation. Clinical information will also be recorded on the same platform.

For TcTU by CT-free SPECT (figure 2), the sinograms for emission and scattering will be reconstructed at each hospital (Q.Volumetrix MI, GE Healthcare) into corresponding NCRR SPECTs, which means that not corrections

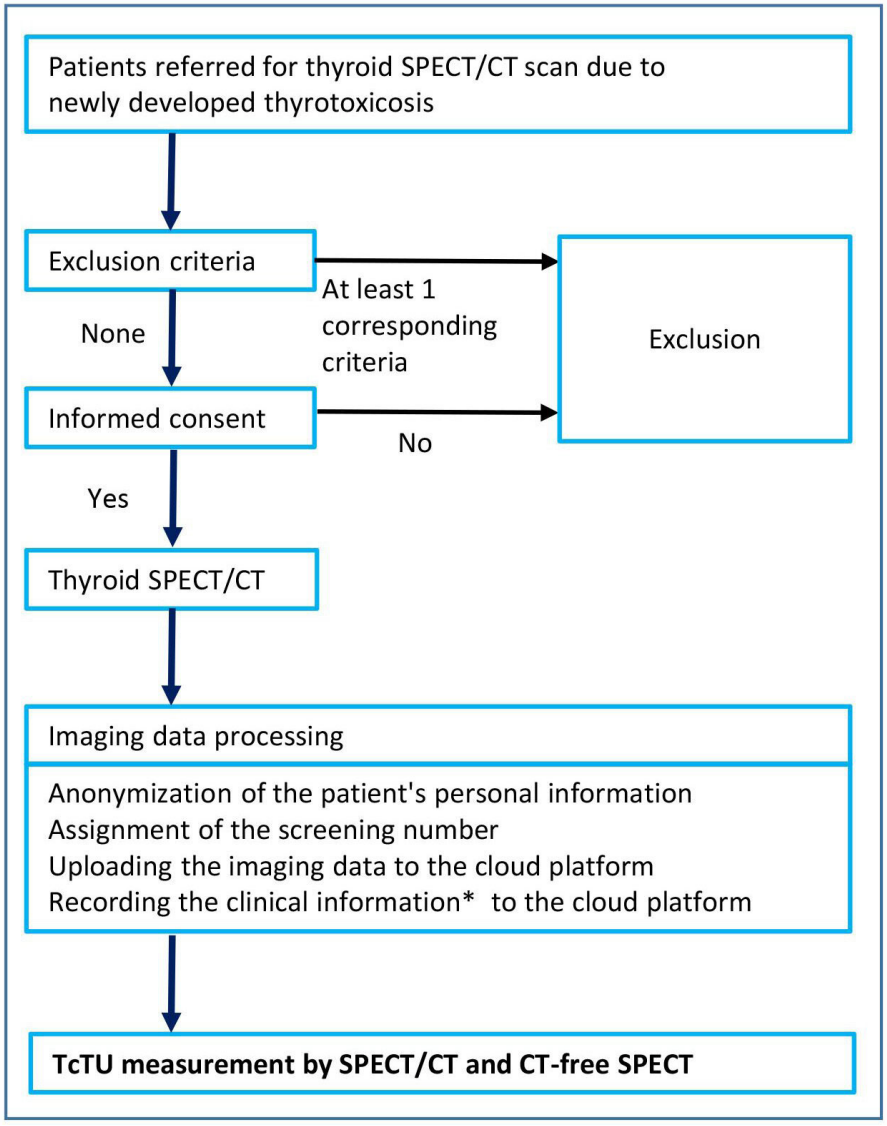

Figure 1 The schema of the case enrolment process. *This includes the following information: the patient's gender, body weight, height and blood tests (serum T3, T4, TSH and TSHRAb). SPECT/CT, single-photon emission CT/CT; TcTU, technetium thyroid uptake; TSH, thyroid-stimulating hormone; TSHRAb, thyroid-stimulating hormone receptor antibody.

for attenuation and scattering but RR application is used. The NCRR SPECTs for emission and scattering are uploaded to the cloud platform (IRM, BEST Image; https://irmbestimage.com), where an AI server is connected via Ethernet. The previously established sequential deep-learning algorithms (the first for synthetic μ-map generation from SPECT images and the second for automatic thyroid segmentation from the generated μ-map), which demonstrated successful performance in both Graves' disease and thyroiditis cases in our previous study, will be applied. Detailed architectures and training methods of the deep-learning model are described in the previous study.[19] The NCRR SPECTs are input for the first convolutional neural network (CNN), and the output is a synthetic attenuation map (μ-map). The process of μ-map generation will take less than 1 min. The generated μ-map and emission SPECT are used as the input of the second CNN after being cropped and down sampled. The output of the second CNN is a thyroid segmentation map. The cropping/downsampling and the second CNN workout will take less than 1 min. In another process, the synthetic μ-map is downloaded to each hospital and used for AC in conjunction with sinograms for emission and scattering to reconstruct ACSCRR (AC, SC and RR applied) SPECT, which is a quantitative SPECT with AC, SC and RR applications (Q.Volumetrix MI, GE Healthcare). The reconstructed ACSCRR SPECT is then uploaded to the cloud platform, where the thyroid segmentation map (the output of the second CNN) is applied to the ACSCRR SPECT to produce TcTU by CT-free SPECT. An investigator (DO), who is not aware of clinical information, is in full charge of the AI operation.

For TcTU by conventional SPECT/CT at the core laboratory, a human expert (JHK), who participated in the development of the AI,[19] will download the sinograms (for emission and scattering) and CT from the cloud platform. She will reconstruct the conventional quantitative ACSCRR SPECT (using Q.Volumetrix MI, GE

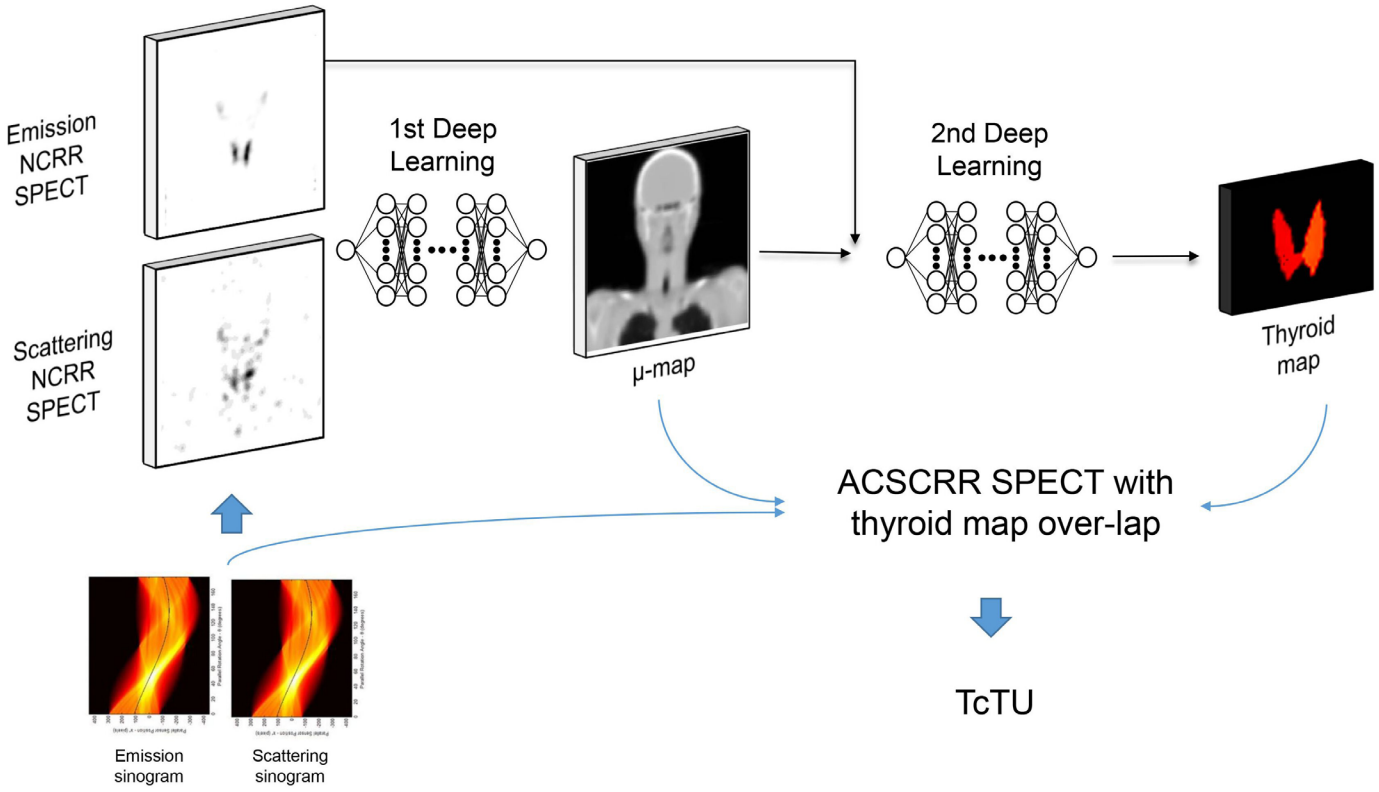

Figure 2 The process of artificial intelligence-based CT-free SPECT for TcTU calculation. SPECT, single-photon emission CT; TcTU, technetium thyroid uptake; NCRR, non-corrected for attenuation and scattering but with resolution recovery applied; ACSCRR, attenuation correction, scatter correction and resolution recovery applied.

Healthcare) and manually segment the thyroid using the CT, which will lead to TcTU by SPECT/CT. The segmentation process usually takes at least 40 min. The clinical information of the subjects and the TcTU results generated by the AI will be kept blind to her.

Study objectives

Primary objective

The primary objective of the current study is to determine whether CT-free thyroid SPECT is not significantly inferior to conventional thyroid SPECT/CT for diagnosing Graves' disease. The diagnosis of Graves' disease will be based on the optimal TcTU value that is determined using the Youden index test, which measures the maximum distance between the ROC curve and the diagonal line. The gold standard for diagnosing Graves' disease is the fulfilment of all of the following four criteria: (1) elevated or within normal range levels of T4 (either free T4 or total T4) and T3 (either free T3 or total T3), (2) reduced TSH, (3) increased TSHRAb and (4) hyperthyroid pattern of increased thyroid uptake as shown by planar scan findings.[55] An experienced endocrinologist (JHM), with more than 15 years of clinical practice, will determine the final diagnosis based on the above four findings, which will be available on the cloud platform. If any of these criteria are not met, the subject will be categorised as 'indeterminate'. The endocrinologist will not have access to the TcTU results.

Secondary objective

The study's secondary objectives are twofold. The first is to evaluate whether CT-free thyroid SPECT is not significantly inferior to conventional thyroid SPECT/CT for diagnosing destructive thyroiditis. The second is to ascertain if CT-free thyroid SPECT can accurately predict the need for antithyroid medication prescriptions within 1 month of the SPECT/CT. The optimal TcTU values for these objectives will be determined by the Youden index test and ROC analysis. The gold standard for diagnosing destructive thyroiditis comprises the following four criteria: (1) T4 (free T4 or total T4) and T3 (free T3 or total T3) are elevated or within normal range, (2) TSH is reduced, (3) TSHRAb is within normal range and (4) planar scan demonstrates a hypothyroid pattern of decreased thyroid uptake.[55] The endocrinologist (JHM), who is blinded to the TcTU results, will also determine the final diagnosis of destructive thyroiditis. Similar to Graves' disease, if any of these criteria are not met, the subject will be labelled as 'indeterminate'. The decision to prescribe antithyroid drugs (propylthiouracil, methimazole, carbimazole) will be made by the physicians who ordered the thyroid SPECT/CT. These physicians are not involved in the clinical trial and will prescribe the medications without considering the TcTU results.

Statistical analyses

The enrolled datasets of CT-free SPECT and SPECT/CT will be used for the statistical analyses. The diagnostic accuracy of TcTU from both CT-free SPECT and SPECT/CT will be determined for Graves' disease. The two-sided 95% CI for the accuracy difference (accuracy of CT-free SPECT—accuracy of SPECT/CT) will be calculated by an asymptotic restricted maximum-likelihood estimation (RMLE) test statistic which has been suggested for paired binary data.[62]

The null hypothesis (H_0) and the alternate hypothesis (H_1) of the current non-inferiority trial are:

$$H_0 : P_{CT-free\ SPECT} - P_{SPECT/CT} \leq -\Delta$$
$$H_1 : P_{CT-free\ SPECT} - P_{SPECT/CT} > -\Delta$$

where, $P_{CT\text{-free SPECT}}$ is the diagnostic accuracy of TcTU by CT-free SPECT, $P_{SPECT/CT}$ is the diagnostic accuracy of TcTU by SPECT/CT and Δ is the non-inferiority margin (=0.1). Non-inferiority of CT-free SPECT to SPECT/CT will be defined as a lower limit of the 95% CI of the accuracy difference for >−0.1. Additionally, paired t-test, Pearson's correlation analysis, McNemar test and ROC analysis will be conducted. A two-sided p value of <0.05 is considered statistically significant.

Non-inferiority margin determination

The non-inferiority margin (Δ) is set as 0.1 (=10% points) for the accuracy difference between $P_{CT\text{-free SPECT}}$−$P_{SPECT/CT}$. As an active control, the $P_{SPECT/CT}$ is assumed to be 95% because (1) the planar TcTU study using the newest gamma camera scanner reported 94.1% accuracy for Graves' disease[63] and (2) the TcTU by quantitative SPECT/CT was superior to the TcTU by planar scan.[11] Meanwhile, using as a placebo study (ie, not imaging study but clinical judgement alone), 85% accuracy is suggested for Graves' disease because (1) typical sign and symptoms were reported in about 50%–70% of Graves' disease,[1 21] (2) TSHRAb positivity in Graves' disease was found to be around 80% of tested populations[2 32–35] and (3) the painless thyroiditis, the most similar thyroiditis type to Graves' disease had 20% positivity of TSHRAb.[37 38]

There should exist a reference study directly comparing the SPECT/CT (active control) and the clinical judgement (placebo) for the non-inferiority margin determination. However, there is no such study. Furthermore, it is hard to expect such a study design because TcTU is already considered as one of the gold standards for Graves' disease and there is no need for reconfirmation of TcTU competency in the near future. In this regard, 10% points of non-inferiority margin by 95% of SPECT/CT accuracy minus 85% of clinical judgement accuracy is reasonably acceptable.

Determination of sample size

The study is planned to observe 152 self-controlled SPECT/CT cases to rule out an accuracy difference of 0.1 at the 95% (2-sided) lower confidence limit ($Z_{\alpha/2}$=1.96) with 90% power (Z_{β}=1.28). The sample size was calculated using the following equation, which was derived from the RMLE-based test statistics for non-inferiority clinical trial of paired binary data[62]:

$$n = 2p_{01}\left\{\frac{Z_{\alpha/2}/\bar{w}+Z_{\beta}}{\Delta}\right\}^2$$

where

$$\bar{w} = (2p_{01})^{1/2}/(\bar{p}_{01} - \Delta - \Delta^2)^{1/2} \text{ and } \bar{p}_{01} = \left\{-a + \left(a^2 - 8b\right)^{1/2}\right\}/4.$$

Here,

$$a = -\theta(1-\Delta) - 2(p_{01}+\Delta) \text{ and } b = \Delta(1+\Delta)p_{01}.$$

In a preliminary test of 34 thyrotoxicosis patients, the CT-free SPECT and the SPECT/CT showed the same accuracy for Graves' disease with the same discordant accuracy of SPECT/CT (p_{01}) and CT-free SPECT (p_{10}) as 6%. With $\theta = P_{CT-free\ SPECT} - P_{SPECT/CT} = p_{10} - p_{01} = 6\% - 6\% = 0\%$, a and b are calculated as −0.32 and 0.0066, respectively. Accordingly, $\bar{p}_{01}$ and $\bar{w}$ are calculated as 0.1357 and 0.8623, respectively.

Thus, the required sample size is calculated as $2 \times 0.06 \times \left(\frac{1.96/0.8623+1.28}{0.1}\right)^2 \approx 152$.

Therefore, the study will continue until 152 complete sets of SPECT/CT and CT-free SPECT are obtained. Assuming a 5% dropout rate, 160 is determined as the final sample size for screening.

Study organisation

The leadership of AISPECT comprises three nuclear medicine physicians (WWL, HWC and S-GC) representing the three investigating hospitals. They are the key members of the Korean Research Council of Hybrid SPECT (K-SPECT) group, which is a branch society of Korean Society of Nuclear Medicine. The leadership designed the study and holds full responsibility for its integrity. Statistical analysis, including the determination of sample size, was conducted in consultation with the Medical Research Collaborating Center at SNUBH. The quality assurance team of the institutional review board (IRB) at SNUBH provided monitoring services for the clinical research. The IRBs of the three investigating hospitals approved the study protocol, and all subjects will provide their informed consent. All the authors read and approved the final draft for submission.

Perspective

The AISPECT study is a single-blind, multicentre, prospective, non-inferiority, clinical trial in a paired design regarding AI application in medical imaging. Despite many promising retrospective AI studies, there are few prospective studies validating the clinical impact in real-world settings.[64] To our knowledge, a multicentre prospective trial of AI application has never been reported before. Among various potential AI applications, CT-free SPECT technology has not been thoroughly investigated, except in a few proof-of-concept studies of myocardial

SPECT for just AC.[18] In the current study, both AC and thyroid segmentation will be automatically performed by the trained AI, and the performance of the AI will be tested in comparison to conventional SPECT/CT. The TcTU is still an important practical tool for the differential diagnosis of thyrotoxicosis. The TcTU will be calculated by the CT-free SPECT technology with the aid of AI. If the non-inferiority of CT-free SPECT to SPECT/CT is proven, this could lead to a reduction in patients' radiation exposure and save human resources for analysis.

ETHICS AND DISSEMINATION

The study protocol has been approved by the IRB of SNUBH (IRB No. B-2304-824-301), KUMC (IRB No. 2023-05-022-006) and CNUH (IRB No. CNUH-2023-108). Findings will be disseminated as reports, presentations and peer-reviewed journal articles.

Author affiliations
[1]Department of Nuclear Medicine, Seoul National University Bundang Hospital, Seongnam-si, Gyeonggi-do, Republic of Korea
[2]Department of Nuclear Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea
[3]Department of Nuclear Medicine, Konkuk University Medical Center, Konkuk University School of Medicine, Seoul, Republic of Korea
[4]Department of Nuclear Medicine, Chonnam National University Hospital and Medical School, Gwangju, Republic of Korea
[5]Department of Health Science and Technology, The Graduate School of Convergence Science and Technology, Seoul National University, Suwon-si, Gyeonggi-do, Republic of Korea
[6]Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam-si, Gyeonggi-do, Republic of Korea
[7]Medical Research Collaborating Center, Seoul National University Bundang Hospital, Seongnam-si, Gyeonggi-do, Republic of Korea
[8]Institute of Radiation Medicine, Medical Research Center, Seoul National University, Seoul, Republic of Korea

Acknowledgements We sincerely thank Heather Lee, Business Manager at IRM, for her technical support of data collection.

Contributors Study concept and design: HWC, S-GC, YS, SA and WWL. Data acquisition: DO, HWC, S-GC, JHK and YS. Data analysis: DO, HGR, KK and JHK. Data interpretation: KK, JHK and JHM. Manuscript drafting: WWL. Manuscript revision: DO, HGR and S-GC. Manuscript editing: HGR and HWC. Clinical studies; HWC, S-GC, KK, YS and JHM. Experimental studies: KK and JHK. Statistical analysis: SA and WWL. Literature research: YS. Guarantors of integrity of entire study: HWC, S-GC and WWL. Agrees to ensure any question related to the work: HWC, S-GC and KK. Approval of final version for submission: DO, HGR, HWC, S-GC, KK, JHK, YS, JHM, SA and WWL.

Funding This study was supported by the Korean Society of Nuclear Medicine Clinical Trial Network Research Award (KSNM-CTN-2022-01-1), funded by the Korean Society of Nuclear Medicine, and by the National Research Foundation of Korea grant, funded by the Korea government (MSIT) (RS-2023-00241409). The funders do not have a role in the study design, data management or site management.

Competing interests WWL and KK hold a patent for the CT-free thyroid SPECT technology. The title of the invention is 'Artificial intelligence algorithm enabling CT-free quantitative thyroid single-photon emission computed tomography (SPECT)' (PCT/KR2023/007541). Other than this, there are no conflicts of interest to declare.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Dongkyu Oh http://orcid.org/0000-0002-7935-3479
Hyun Gee Ryoo http://orcid.org/0000-0001-5020-5846
Hyun Woo Chung http://orcid.org/0000-0003-4167-3060
Sang-Geon Cho http://orcid.org/0000-0002-1373-1887
Kyounghyoun Kwon http://orcid.org/0000-0002-4969-9655
Ji Hye Kim http://orcid.org/0009-0009-2393-0861
Young So http://orcid.org/0009-0009-3825-8422
Jae Hoon Moon http://orcid.org/0000-0001-6327-0575
Soyeon Ahn http://orcid.org/0000-0003-3440-2027
Won Woo Lee http://orcid.org/0000-0002-9188-6921

REFERENCES

1 McKee A, Peyerl F. TSI assay utilization: impact on costs of Graves' hyperthyroidism diagnosis. *Am J Manag Care* 2012;18:e1–14.
2 Khoo DH, Fok AC, Tan CE, *et al*. Thyroid stimulating hormone receptor antibody levels in Singaporean patients with autoimmune thyroid disease. *Ann Acad Med Singap* 1997;26:435–8.
3 Costagliola S, Morgenthaler NG, Hoermann R, *et al*. Second generation assay for thyrotropin receptor antibodies has superior diagnostic sensitivity for Graves' disease. *J Clin Endocrinol Metab* 1999;84:90–7.
4 Kamijo K, Ishikawa K, Tanaka M. Clinical evaluation of 3rd generation assay for thyrotropin receptor antibodies: the M22-biotin-based ELISA initiated by Smith. *Endocr J* 2005;52:525–9.
5 Atkins HL. Technetium-99m pertechnetate uptake and scanning in the evaluation of thyroid function. *Semin Nucl Med* 1971;1:345–55.
6 Giovanella L, Avram AM, Iakovou I, *et al*. EANM practice guideline/SNMMI procedure standard for RAIU and thyroid scintigraphy. *Eur J Nucl Med Mol Imaging* 2019;46:2514–25.
7 Lee H, Kim JH, Kang Y, *et al*. Quantitative single-photon emission computed tomography/computed tomography for technetium pertechnetate thyroid uptake measurement. *Medicine (Abingdon)* 2016;95:e4170.
8 Kim HJ, Bang J-I, Kim J-Y, *et al*. Novel Application of Quantitative Single-Photon Emission Computed Tomography/Computed Tomography to Predict Early Response to Methimazole in Graves' Disease. *Korean J Radiol* 2017;18:543.
9 Kim J-Y, Kim JH, Moon JH, *et al*. Utility of Quantitative Parameters from Single-Photon Emission Computed Tomography/Computed Tomography in Patients with Destructive Thyroiditis. *Korean J Radiol* 2018;19:470–80.
10 Lee R, So Y, Song YS, *et al*. Evaluation of Hot Nodules of Thyroid Gland Using Tc-99m Pertechnetate: a Novel Approach Using Quantitative Single-Photon Emission Computed Tomography/Computed Tomography. *Nucl Med Mol Imaging* 2018;52:468–72.
11 Oh D, Kim JH, Yoo J, *et al*. Absolute %uptake of quantitative SPECT/CT is more accurate than relative uptake ratio of planar scan for identification of Graves' disease and destructive thyroiditis. *J Nucl Med* 2023;64:1501.
12 Lee WW, Song YS, So Y. Quantitative Iodine-123 single-photon emission computed tomography/computed tomography for Iodine-131 therapy of an autonomously functioning thyroid nodule. *Eur J Hybrid Imaging* 2023;7:4.
13 Suh MS, Lee WW, Kim Y-K, *et al*. Maximum Standardized Uptake Value of (99m)Tc Hydroxymethylene Diphosphonate SPECT/CT for the Evaluation of Temporomandibular Joint Disorder. *Radiology* 2016;280:890–6.
14 Lee WW, K-SPECT Group. Clinical Applications of Technetium-99m Quantitative Single-Photon Emission Computed Tomography/Computed Tomography. *Nucl Med Mol Imaging* 2019;53:172–81.
15 Park J, Bae S, Seo S, *et al*. Measurement of Glomerular Filtration Rate using Quantitative SPECT/CT and Deep-learning-based Kidney Segmentation. *Sci Rep* 2019;9:4223.
16 Park J, Lee JS, Oh D, *et al*. Quantitative salivary gland SPECT/CT using deep convolutional neural networks. *Sci Rep* 2021;11:7842.
17 Shi L, Onofrey JA, Liu H, *et al*. Deep learning-based attenuation map generation for myocardial perfusion SPECT. *Eur J Nucl Med Mol Imaging* 2020;47:2383–95.

18 Shanbhag AD, Miller RJH, Pieszko K, *et al*. Deep Learning-Based Attenuation Correction Improves Diagnostic Accuracy of Cardiac SPECT. *J Nucl Med* 2023;64:472–8.
19 Kwon K, Hwang D, Oh D, *et al*. CT-free quantitative SPECT for automatic evaluation of %thyroid uptake based on deep-learning. *EJNMMI Phys* 2023;10:20.
20 Ross DS, Burch HB, Cooper DS, *et al*. 2016 American Thyroid Association Guidelines for Diagnosis and Management of Hyperthyroidism and Other Causes of Thyrotoxicosis. *Thyroid* 2016;26:1343–421.
21 Smith TJ, Hegedüs L. Graves' Disease. *N Engl J Med* 2016;375:1552–65.
22 Carlé A, Knudsen N, Pedersen IB, *et al*. Determinants of serum T4 and T3 at the time of diagnosis in nosological types of thyrotoxicosis: a population-based study. *Eur J Endocrinol* 2013;169:537–45.
23 Shigemasa C, Abe K, Taniguchi S, *et al*. Lower serum free thyroxine (T4) levels in painless thyroiditis compared with Graves' disease despite similar serum total T4 levels. *J Clin Endocrinol Metab* 1987;65:359–63.
24 Sriphrapradang C, Bhasipol A. Differentiating Graves' disease from subacute thyroiditis using ratio of serum free triiodothyronine to free thyroxine. *Ann Med Surg* 2016;10:69–72.
25 Tura Bahadır Ç, Yılmaz M, Kılıçkan E. Free triiodothyronine to free thyroxine ratio in the differential diagnosis of thyrotoxicosis and hyperthyroidism: A retrospective study. *Int J Clin Pract* 2021;75:e14003.
26 Lytton SD, Schluter A, Banga PJ. Functional diagnostics for thyrotropin hormone receptor autoantibodies bioassays prevail over binding assays. *Front Biosci* 2018;23:2028–43.
27 Kahaly GJ. Bioassays for TSH Receptor Antibodies: Quo Vadis? *Eur Thyroid J* 2015;4:3–5.
28 Dwivedi SN, Kalaria T, Buch H. Thyroid autoantibodies. *J Clin Pathol* 2023;76:19–28.
29 Tozzoli R, Kodermaz G, Villalta D, *et al*. Accuracy of receptor-based methods for detection of thyrotropin-receptor autoantibodies: a new automated third-generation immunoassay shows higher analytical and clinical sensitivity for the differential diagnosis of hyperthyroidism. *Auto Immun Highlights* 2010;1:95–100.
30 Tozzoli R, Bagnasco M, Giavarina D, *et al*. TSH receptor autoantibody immunoassay in patients with Graves' disease: improvement of diagnostic accuracy over different generations of methods. Systematic review and meta-analysis. *Autoimmun Rev* 2012;12:107–13.
31 Kahaly GJ, Diana T. TSH Receptor Antibody Functionality and Nomenclature. *Front Endocrinol (Lausanne)* 2017;8:28.
32 Zouvanis M, Panz VR, Kalk WJ, *et al*. Thyrotropin receptor antibodies in black South African patients with Graves' disease and their response to medical therapy. *J Endocrinol Invest* 1998;21:771–4.
33 Mo SI, Ryu AJ, Kim YJ, *et al*. The High Proportion of Painless Thyroiditis as a Cause of Thyrotoxicosis in Korea. *J Korean Thyroid Assoc* 2015;8:61.
34 Mondal RK, Bhattacharjee R, Mukherjee K, *et al*. Clinical Association of TSH Receptor Antibody and Related Autoimmune Markers with Subclinical Graves' Disease: A Study from Eastern Indian Population. *JCDR* 2022;16:BC01–4.
35 Tee SA, Bishop DR, Al-Sharefi A, *et al*. Use of TSH-receptor antibodies (TRAb) in the assessment of new onset thyrotoxicosis within the North East of England. *Thyroid Res* 2017;10.
36 Schott M, Hermsen D, Broecker-Preuss M, *et al*. n.d. Clinical value of the first automated TSH receptor autoantibody assay for the diagnosis of Graves. 71.
37 Morita T, Tamai H, Oshima A, *et al*. The occurrence of thyrotropin binding-inhibiting immunoglobulins and thyroid-stimulating antibodies in patients with silent thyroiditis. *J Clin Endocrinol Metab* 1990;71:1051–5.
38 Sanyal D, Chatterjee S. Thyrotropin receptor antibody immunoassays may not be reliable in confirming diagnosis of Painless Thyroiditis. *Acta Endocrinol (Buchar)* 2020;16:530–4.
39 Ilicki A, Gamstedt A, Karlsson FA. Hyperthyroid Graves' disease without detectable thyrotropin receptor antibodies. *J Clin Endocrinol Metab* 1992;74:1090–4.
40 Mukuta T, Tamai H, Oshima A, *et al*. Immunological findings and thyroid function of untreated Graves' disease patients with undetectable TSH-binding inhibitor immunoglobulin. *Clin Endocrinol (Oxf)* 1994;40:215–9.
41 Scappaticcio L, Trimboli P, Keller F, *et al*. Diagnostic testing for Graves' or non-Graves' hyperthyroidism: A comparison of two thyrotropin receptor antibody immunoassays with thyroid scintigraphy and ultrasonography. *Clin Endocrinol (Oxf)* 2020;92:169–78.
42 Kamijo K. Study on cutoff value setting for differential diagnosis between Graves' disease and painless thyroiditis using the TRAb (Elecsys TRAb) measurement via the fully automated electrochemiluminescence immunoassay system. *Endocr J* 2010;57:895–902.
43 Perdomo CM, García-Goñi M, Sancho L, *et al*. Evaluation of the role of thyroid scintigraphy in the differential diagnosis of thyrotoxicosis. *Clin Endocrinol (Oxf)* 2021;94:466–72.
44 Erdoğan MF, Anil C, Cesur M, *et al*. Color flow Doppler sonography for the etiologic diagnosis of hyperthyroidism. *Thyroid* 2007;17:223–8.
45 Hiraiwa T, Tsujimoto N, Tanimoto K, *et al*. Use of color Doppler ultrasonography to measure thyroid blood flow and differentiate graves' disease from painless thyroiditis. *Eur Thyroid J* 2013;2:120–6.
46 Morosini PP, Simonella G, Mancini V, *et al*. Color Doppler sonography patterns related to histological findings in Graves' disease. *Thyroid* 1998;8:577–82.
47 Zuhur SS, Ozel A, Kuzu I, *et al*. The Diagnostic Utility of Color Doppler Ultrasonography, Tc-99m Pertechnetate Uptake, and TSH-Receptor Antibody for Differential Diagnosis of Graves' Disease and Silent Thyroiditis: A Comparative Study. *Endocr Pract* 2014;20:310–9.
48 Peng X, Wu S, Bie C, *et al*. Mean peak systolic velocity of superior thyroid artery for the differential diagnosis of thyrotoxicosis: a diagnostic meta-analysis. *BMC Endocr Disord* 2019;19:56.
49 Kurita S, Sakurai M, Kita Y, *et al*. Measurement of thyroid blood flow area is useful for diagnosing the cause of thyrotoxicosis. *Thyroid* 2005;15:1249–52.
50 Meng Z, Zhang G, Sun H, *et al*. Differentiation between Graves' disease and painless thyroiditis by diffusion-weighted imaging, thyroid iodine uptake, thyroid scintigraphy and serum parameters. *Exp Ther Med* 2015;9:2165–72.
51 Vos XG, Smit N, Endert E, *et al*. Frequency and characteristics of TBII-seronegative patients in a population with untreated Graves' hyperthyroidism: a prospective study. *Clin Endocrinol (Oxf)* 2008;69:311–7.
52 Malik SA, Choh NA, Misgar RA, *et al*. Comparison between peak systolic velocity of the inferior thyroid artery and technetium-99m pertechnetate thyroid uptake in differentiating Graves' disease from thyroiditis. *Arch Endocrinol Metab* 2019;63:495–500.
53 Narkar RR, Mishra I, Baliarsinha AK, *et al*. Rapid Differential Diagnosis of Thyrotoxicosis Using T3/T4 Ratio, FT3/FT4 Ratio and Color Doppler of Thyroid Gland. *Indian J Endocrinol Metab* 2021;25:193–7.
54 Tezuka M, Murata Y, Ishida R, *et al*. MR imaging of the thyroid: correlation between apparent diffusion coefficient and thyroid gland scintigraphy. *J Magn Reson Imaging* 2003;17:163–9.
55 Intenzo CM, dePapp AE, Jabbour S, *et al*. Scintigraphic Manifestations of Thyrotoxicosis. *Radiographics* 2003;23:857–69.
56 Qiao T, Liu S, Cui Z, *et al*. Deep learning for intelligent diagnosis in thyroid scintigraphy. *J Int Med Res* 2021;49:0300060520982842.
57 Tzen KY, Chen HY, Kao PF, *et al*. Using a computerized method to measure 99mTc pertechnetate uptake for the assessment of thyroid function: a clinical validation. *J Formos Med Assoc* 1990;89:356–62.
58 Siraj QH, Renilson S, Masoomi M, *et al*. Comparison of different techniques for assessment of thyroid functional status in patients with hyperthyroidism. *Iran J Nucl Med (Suppl)* 2012;20:137–8.
59 Potočnik J, Foley S, Thomas E. Current and potential applications of artificial intelligence in medical imaging practice: A narrative review. *J Med Imaging Radiat Sci* 2023;54:376–85.
60 Nam JG, Hong JH, Kim DS, *et al*. Deep learning reconstruction for contrast-enhanced CT of the upper abdomen: similar image quality with lower radiation dose in direct comparison with iterative reconstruction. *Eur Radiol* 2021;31:5533–43.
61 Varoquaux G, Cheplygina V. Machine learning for medical imaging: methodological failures and recommendations for the future. *NPJ Digit Med* 2022;5:48.
62 Liu J, Hsueh H, Hsieh E, *et al*. Tests for equivalence or non-inferiority for paired binary data. *Stat Med* 2002;21:231–45.
63 Jin M, Ahn J, Jo S-G, *et al*. Comparison of ^{99m}Tc Pertechnetate Thyroid Uptake Rates by Gamma Probe and Gamma Camera Methods for Differentiating Graves' Disease and Thyroiditis. *Nucl Med Mol Imaging* 2022;56:42–51.
64 Nam JG, Hwang EJ, Kim J, *et al*. AI Improves Nodule Detection on Chest Radiographs in a Health Screening Population: A Randomized Controlled Trial. *Radiology* 2023;307:e221894.