Postprint of Short-term Teaching Effectiveness Evaluation of the Standardized Thyroid Nodule Ultrasound Training and Assessment System

Authors: Gao Luying, Jiang Yuxin, Li Jianchu, Zhang Bo, Wang Ying, Liu Ruyu, Xi Xuehua, Gao Qiong, Yang Xiao, Lai Xingjian, Zhu Shenling, Zhang Xiaoyan, Zhao Ruina

Date: 2018-10-26T00:00:00+00:00

Abstract

Objective: This study aimed to investigate the short-term teaching effectiveness of the "Standardized Thyroid Nodule Ultrasound Training-Assessment System." Methods: Twelve trainees in the Department of Ultrasound at Peking Union Medical College Hospital in December 2017 were prospectively enrolled. First, the "assessment system" component of the "Standardized Thyroid Nodule Ultrasound Training-Assessment System" was used to evaluate the 12 trainees, with assessment content covering ultrasound feature identification and malignant risk stratification of thyroid nodules; the system automatically recorded scores. Subsequently, the training component of the system was utilized for corresponding training. At 12 hours and 1 month after training, all trainees were reassessed using the system, which automatically recorded scores. Differences in scores before and after training were compared. Results: In this study, trainees' thyroid nodule image interpretation scores significantly improved both at 12 hours and 1 month after receiving "thyroid nodule ultrasound training" (80.7$±7.3$vs$69.9±4.0, P < 0.01; 78.0±9.2vs69.9±$4.0, P<0.05), with diagnostic competence in both ultrasound feature identification of thyroid nodules and ultrasound-based risk stratification of thyroid nodules showing improvement (P<0.05). Compared with 12 hours after "thyroid nodule ultrasound training," the diagnostic level for thyroid nodules did not significantly decline at 1 month after training (P>0.05). Conclusion: The "Standardized Thyroid Nodule Ultrasound Training-Assessment System" can improve trainees' ultrasound diagnostic competence for thyroid nodules in the short term, providing a novel teaching approach for enhancing ultrasound diagnostic levels of thyroid nodules and promoting the standardization process of thyroid nodule ultrasound diagnosis.

Full Text

Short-term Teaching Effectiveness Evaluation of the Standardized Thyroid Nodule Ultrasound Training-Examination System

Luying Gao, MD, Yuxin Jiang, MD, Jianchu Li, MD, Bo Zhang, MD, Ying Wang, MD, Ruyu Liu, MB, Xuehua Xi, MB, Gao Qiong, MB, Xiao Yang, MD, Xingjian Lai, MD, Shenling Zhu, MD, Xiaoyan Zhang, MD, Ruina Zhao, MD

Department of Ultrasound, Chinese Academy of Medical Sciences & Peking Union Medical College Hospital, Beijing, China

Corresponding author: Bo Zhang
Phone: 69155493
E-mail: thyroidus@163.com

Abstract

Objective: This study aims to evaluate the short-term teaching effectiveness of the "Standardized Thyroid Nodule Ultrasound Training-Examination System."

Methods: We prospectively enrolled 12 trainees in the Department of Ultrasound at Peking Union Medical College Hospital in December 2017. All participants first completed an assessment using the examination component of the system, which evaluated their ability to identify ultrasound characteristics of thyroid nodules and classify malignant risk. The system automatically recorded scores. Participants then underwent training using the system's training module. Follow-up assessments were conducted at 12 hours and 1 month post-training using the same examination system, with scores automatically recorded. Pre- and post-training scores were compared.

Results: After completing the thyroid nodule ultrasound training, participants demonstrated significant improvement in image interpretation scores at both 12 hours and 1 month post-training $(80.7\$\pm7.3 vs 69.9\pm4.0, P < 0.01; 78.0\pm9.2 vs 69.9\pm\$4.0$, P<0.05). Diagnostic accuracy improved for all individual ultrasound feature identifications and nodule classification systems (P<0.05). Compared with scores at 12 hours post-training, performance did not significantly decline at 1 month (P>0.05).

Conclusion: The standardized thyroid nodule ultrasound training-examination system can rapidly improve trainees' diagnostic capabilities in thyroid nodule ultrasound within a short timeframe, providing a novel educational approach for advancing the standardization of thyroid nodule ultrasound diagnosis.

Keywords: thyroid nodule; ultrasound; standardized training-examination system

Introduction

Thyroid disease represents the most common endocrine disorder, with reported prevalence rates of thyroid nodules ranging from 19% to 68% in the general population, of which 7% to 15% are malignant. With increasing health awareness and widespread use of imaging examinations, the detection rates of thyroid nodules and thyroid cancer have risen annually. Ultrasound serves as the primary diagnostic modality for evaluating thyroid nodule malignancy risk [1-5]. In 2017, the American College of Radiology (ACR) published the “ACR Thyroid Imaging-Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee” [6], while the American Thyroid Association (ATA) released its “Guidelines for the Management of Adults with Differentiated Thyroid Cancer” in 2015 [7]. Both guidelines propose risk stratification of thyroid nodules based on ultrasound characteristics to guide decisions regarding fine-needle aspiration biopsy, follow-up, and treatment. Therefore, accurate ultrasound-based malignant risk stratification is critically important.

However, ultrasound is an imaging-based diagnostic technique with inherent observer dependency and subjectivity, leading to high misdiagnosis rates among less experienced physicians. This represents a significant limitation in the accurate application of thyroid ultrasound guidelines. Traditional thyroid ultrasound teaching typically follows textbook sequences covering ultrasound principles, organ anatomy, standard scanning techniques, and normal and abnormal images, which proves difficult to memorize, comprehend, and translate into standardized diagnoses. Based on the professional characteristics of thyroid ultrasound and extensive teaching experience, the Department of Ultrasound at Peking Union Medical College Hospital has developed a “Standardized Thyroid Nodule Ultrasound Training-Examination System” grounded in the ACR TI-RADS and ATA guidelines. This study prospectively investigates the value of this system in short-term thyroid ultrasound teaching.

Materials and Methods

The “Standardized Thyroid Nodule Ultrasound Training-Examination System” comprises two components: the “Thyroid Nodule Ultrasound Training System” and a WeChat-based online “Thyroid Nodule Ultrasound Examination System.”

Thyroid Nodule Ultrasound Training System Based on the ultrasound lexicon clearly defined by the ACR TI-RADS committee, this system delivers instruction through online or offline lectures using typical images to explain ultrasound features including composition, echogenicity, shape, margin, and echogenic foci (5-10 images for each feature or sub-feature).

1. Composition: Describes the internal structure of nodules, representing the proportion of soft tissue to fluid components.
1.1 Solid nodule: Composed entirely or almost entirely of soft tissue with minimal cystic components.

2. Echogenicity: Refers to the echogenicity level of the solid portion of the nodule (non-calcified), using surrounding thyroid tissue as reference.
2.4 Markedly hypoechoic: Echogenicity lower than that of adjacent neck strap muscles.

3. Shape:
3.1 Taller-than-wide: Ratio of anteroposterior to transverse diameter >1 on transverse view.

4. Size: Measurements include maximum longitudinal diameter on longitudinal view and anteroposterior and transverse diameters on transverse view. Three dimensions are measured: maximum diameter on transverse view, maximum diameter perpendicular to the first measurement on the same image, and maximum longitudinal diameter on sagittal view. Note: Taller-than-wide ratio differs, calculated as anteroposterior diameter (parallel to ultrasound beam) to transverse diameter (perpendicular to beam) on transverse view. If a halo is present, include it in measurements.

5. Margin: The boundary or interface between the nodule and thyroid parenchyma or adjacent extrathyroidal structures.
5.5 Extrathyroidal extension: Nodule extending beyond the thyroid capsule.

6. Echogenic foci: Focal areas with significantly increased echogenicity relative to surrounding tissue, varying in size and shape, single or clustered, may demonstrate posterior acoustic shadowing.
6.1 Punctate echogenic foci: No posterior shadowing, diameter <1 mm, including small comet-tail artifacts within solid components.
6.2 Macrocalcification: Posterior shadowing present, may have irregular morphology.
6.3 Peripheral calcification: Calcification completely or partially surrounding or occupying most of the nodule margin, often obscuring internal architecture.
6.4 Comet-tail artifacts: Reverberation artifacts with decreasing echogenicity and gradually narrowing width with increasing depth, forming a triangular shape. Large comet-tail: >1 mm (0 points); small comet-tail: $ $1 mm.

Thyroid Nodule Ultrasound Examination System The WeChat-based online examination consists of 20 thyroid nodule image interpretation cases [Figure 20: see original paper]. Each case includes seven multiple-choice questions covering five ultrasound features and two classification systems, totaling 140 questions [Figure 1: see original paper].

Ultrasound features include: Margin/boundary (clear, spiculated, lobulated, infiltrative), shape (regular, irregular, taller-than-wide), echogenicity (hyperechoic, isoechoic, hypoechoic, markedly hypoechoic, anechoic), internal structure (cystic, mixed cystic-solid, solid), and calcifications (none, microcalcifications, macrocalcifications).

Classification systems include:
(1) ATA guidelines [6] risk stratification: Categories include high suspicion, intermediate suspicion, low suspicion, very low suspicion, and benign.
- **High suspicion:** Solid hypoechoic or mixed cystic-solid nodules with hypoechoic solid components demonstrating one or more of the following: irregular margins (lobulated, spiculated, infiltrative), microcalcifications, taller-than-wide shape, interrupted peripheral calcification with extrusive hypoechoic tissue, extrathyroidal extension.
- **Intermediate suspicion:** Solid hypoechoic nodules with smooth, regular margins, without microcalcifications, taller-than-wide shape, or extrathyroidal extension.
- **Low suspicion:** Isoechoic or hyperechoic solid nodules or eccentric solid portions of mixed nodules without microcalcifications, irregular margins, taller-than-wide shape, or extrathyroidal extension.
- **Very low suspicion (<3% malignancy risk):** "Spongiform" nodules, mixed cystic-solid nodules with non-eccentric solid portions lacking microcalcifications, irregular margins, taller-than-wide shape, or extrathyroidal extension.
- **Benign:** Predominantly cystic nodules.

(2) Kwak et al. [8] TI-RADS classification: Five suspicious ultrasound features include solid composition, hypoechogenicity or marked hypoechogenicity, lobulated or irregular margins, microcalcifications, and taller-than-wide shape.
- TIRADS 1 (0% risk): Normal thyroid
- TIRADS 2 (0% risk): Benign, colloid nodules (cystic, spongiform, mixed isoechoic with punctate echogenic foci)
- TIRADS 3 (1.7% risk): No suspicious features, Hashimoto nodules
- TIRADS 4a (3.3% risk): 1 suspicious feature
- TIRADS 4b (9.2% risk): 2 suspicious features
- TIRADS 4c (44.4%-74.4% risk): 3-4 suspicious features
- TIRADS 5 (87.5% risk): 5 suspicious features

Study Participants and Teaching Process In December 2017, 12 trainees participating in advanced training at the Department of Ultrasound, Peking Union Medical College Hospital were enrolled. A questionnaire collected baseline information including gender, age, telephone number, and years of education.

All 12 participants completed the "Thyroid Nodule Ultrasound Examination System" via the WeChat-based online platform before training. The examination consisted of 20 thyroid nodule image interpretation cases, with seven multiple-choice questions per case (five ultrasound features and two classification systems:

ATA risk stratification and TI-RADS), totaling 140 questions. The system automatically recorded scores. Two hours later, participants attended the "Thyroid Nodule Ultrasound Training System" through live instruction, where instructors used numerous typical images to explain ultrasound features including composition, echogenicity, shape, margin, and echogenic foci (5-10 images per feature or sub-feature). Follow-up examinations were conducted at 12 hours and 1 month post-training using the same online system, with scores automatically recorded.

Statistical Analysis Statistical analysis was performed using SPSS 22.0 and SAS software. Continuous variables were expressed as mean ± standard deviation (± s). T-tests were used for normally distributed data with equal variance; otherwise, rank-sum tests were applied. Categorical data were analyzed using chi-square tests, Fisher' s exact test, and Kappa consistency tests. P<0.05 was considered statistically significant.

Results

Baseline Characteristics of Trainees Among the 12 participants, 1 was male and 11 were female, with a mean age of 33.0$±3.3*years*, *meaneducationdurationof*17.2±0.9*years*, *andmed* years. Two participants (16.7%) had imaging backgrounds, while 10 (83.3%) had clinical medicine backgrounds. All 12 participants (100%) had previously received thyroid ultrasound instruction.

Comparison of Scores Before and After Training Before training, the mean score on the "Thyroid Nodule Ultrasound Examination System" was 69.9$±4.0.*At*12*hourspost*−*training*, *themeanscoreincreasedto*80.7±7.3.*At*1*monthpost*− *training*, *themeanscorewas*78.0±9.2.*Comparedwithpre*−*trainingscores*, *thetrainingsystemsignificantlyimp* 0.01; 78.0±9.2*vs*69.9±4.0, *P* < 0.05).*Nosignificantdifferencewasobservedbetweenscoresat*12*hoursand*1*mon* *training*(80.7±7.3*vs*78.0±$9.2, P>0.05).

Comparison of Ultrasound Feature Identification Scores Before and After Training Before training, mean scores for individual features were: margin/boundary 92.5$±6.6, *shape*82.5±6.9, *echogenicity*69.2±12.4, *internalstructure*80.8±7.0, *calcification* *RADSclassification*39.6±11.2.*At*12*hourspost*−*training*, *scoresimprovedsignificantlyto* : *margin*/*boundary*96.3±5.7, *shape*88.8±7.7, *echogenicity*78.3±9.8, *internalstructure*88.8±9.1, *calcifications*9 *RADSclassification*61.7±$16.7 (P=0.04, P=0.04, P=0.04, P=0.03, P<0.01, P=0.04, P<0.01, respectively) .

Inter-observer Variability in Ultrasound Feature Identification Before and After Training This study demonstrated moderate to good inter-observer consistency among trainees for thyroid nodule feature identification both before and after training. However, no significant improvement in inter-observer variability was observed at 12 hours post-training .

Discussion

This study demonstrates that trainees' thyroid nodule image interpretation scores improved significantly after completing the thyroid nodule ultrasound training, indicating that the training-examination system enhances diagnostic capabilities. Furthermore, diagnostic performance did not decline at 1 month compared with 12 hours post-training. Thyroid ultrasound is an image-based discipline with inherent observer dependency and subjectivity, contributing to high misdiagnosis rates among less experienced physicians. Traditional thyroid ultrasound teaching primarily uses PowerPoint presentations to cover standard scanning techniques and normal/abnormal images, making it difficult for trainees to achieve standardized scanning and diagnosis within short timeframes. Our training system addresses this by incorporating the latest standardized guidelines and providing clear, representative images of thyroid ultrasound features with detailed explanations, facilitating comprehension and mastery.

Before training, trainees demonstrated relatively low accuracy in identifying "echogenicity"(69.2$±12.4)*amongthefiveultrasoundfeatures*(*composition*, *echogenicity*, *shape*, *margin/boun* P<0.05). Accurate identification of echogenicity levels in solid thyroid nodules is clinically crucial: according to ATA guidelines, solid hypoechoic nodules are classified as high or intermediate suspicion (malignancy rates of 70-90% and 10-20%, respectively), while isoechoic or hyperechoic solid nodules are low suspicion (malignancy rate only 5-10%). Therefore, accurate differentiation between hypoechoic and iso/hyperechoic solid nodules is essential in clinical practice.

Before training, trainees showed relatively high accuracy in identifying composition, shape, margin/boundary, and echogenic foci, suggesting adequate recognition of malignant features such as irregular margins, microcalcifications, and taller-than-wide shape. However, performance in ATA risk stratification and TI-RADS classification was poor pre-training (47.5$±17.8*and*39.6±11.2, *respectively*), *likelybecausetraineesreliedonconventionaldiagnostichabitsratherth*
basedstratification.*Sinceclinicalmanagementdecisionsregardingbiopsy*, *follow*−
up, *andtreatmentdependonaccurateriskstratification*, *standardizedtrainingisparamount*.*Post*−
trainingaccuracyinATAstratificationandTI−*RADSclassificationimprovedsignificantly*(63.8±18.1, *P* <
0.05; 61.7±$16.7, P<0.05) but remained suboptimal, suggesting that additional training sessions may be needed to achieve satisfactory proficiency [6-7].

A study by Chang et al. involving five radiologists demonstrated that inter-observer agreement reached acceptable levels when using unified risk stratification criteria [9]. Similarly, M. Naren et al. reported inter-observer agreement for TI-RADS classification among six American radiologists ranging from 0.82 to 1.0 [10]. Our study found moderate to good inter-observer consistency before training, but at a relatively low accuracy level. After training, while inter-observer variability did not change substantially, diagnostic accuracy improved significantly, suggesting that training corrected common misconceptions among trainees, enabling higher accuracy while maintaining consistency.

Traditional thyroid ultrasound training is costly and constrained by time and space limitations. A key advantage of our system is its online accessibility, allowing trainees to complete assessments remotely. Another weakness of conventional teaching is the lack of standardized image interpretation, as individual instructors' diagnostic experiences vary. Our system addresses this by providing standardized, systematic training based on ACR TI-RADS guidelines, standardizing the identification of each ultrasound feature.

Furthermore, this study demonstrates that standardized training improves accurate risk stratification through precise feature identification, though it does not substantially enhance inter-observer agreement. This suggests that trainees shared common misinterpretations before training, which were corrected through standardized instruction, resulting in high-accuracy consistency.

This study is limited by its small sample size constrained by space and participant availability. Future research with larger cohorts is needed to validate the effectiveness of this training-examination system. In conclusion, the "Standardized Thyroid Nodule Ultrasound Training-Examination System" effectively enhances ultrasound physicians' diagnostic capabilities for thyroid nodules in a short period, providing methodology and evidence for standardized thyroid nodule ultrasound education.

Funding

National International Science and Technology Cooperation Program (2015DFA30440); Capital Health Development Research Fund (2016-2-40110); Peking Union Medical College Education and Teaching Reform Project (2016zlgc0108)

References

[1] Guth S, Theune U, Aberle J, et al. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination [J]. Eur J Clin Invest. 2009,39:699-706.
[2] Tan G H, Gharib H. Thyroid Incidentalomas: Management Approaches to Nonpalpable Nodules Discovered Incidentally on Thyroid Imaging [J]. Annals of Internal Medicine. 1997, 126:226-31.
[3] Hegedüs, L. The thyroid nodule. The New England Journal of Medicine [J]. 2004,351: 1764-1771
[4] Leenhardt L, Erdogan M F, Hegedus L, et al. European Thyroid Association Guidelines for Cervical Ultrasound Scan and Ultrasound-Guided Techniques in the Postoperative Management of Patients with Thyroid Cancer[J]. European

Thyroid Journal, 2013, 2:147-159.
[5] Perros P, Boelaert K, Colley S, et al. British Thyroid Association Guidelines for the Management of Thyroid Cancer[J]. Clinical Endocrinology, 2014, 81:1-122.
[6] Tessler F N, Middleton W D, Grant E G, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee[J]. Journal of the American College of Radiology, 2017, 14:587.
[7] Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer [J]. Thyroid,2016, 26:1-133.
[8] Jin Young Kwak, Kyung Hwa Han, Yoon Hyun Yoon, et al. Thyroid Imaging Reporting and Data System for US Features of Nodules: A Step in Establishing Better Stratification of Cancer Risk[J]. Radiology. 2011, 260:892-9
[9] Park C S, Kim S H, Jung S L, et al. Observer variability in the sonographic evaluation of thyroid nodules[J]. Journal of Clinical Ultrasound, 2010, 38:287.
[10] Satya S M N, Amogh V N, Satya G M, et al. A Prospective Study to Evaluate the Reliability of Thyroid Imaging Reporting and Data System in Differentiation between Benign and Malignant Thyroid Lesions[J]. Journal of Clinical Imaging Science, 2016,

---

Table 1 Comparison of Trainee Scores Before and 12 Hours After Thyroid Nodule Ultrasound Training

<table>
<tr><th>Parameter</th><th>Pre-training Score (± s, points)</th><th>12 Hours Post-training Score (± s, points)</th><th>P-value</th></tr>
<tr><td>Margin/Boundary</td><td>92.5$±6.6|96.3±5.7|0.04||Shape|81.5±6.9|88.8±7.7|0.04||Echogenicity|69.2±12.4|78.3±9.8|0.04||Int 0.01||ATARiskStratification|47.5±17.8|63.8±18.1|0.04||TI−RADSClassification|39.6±11.2|61.7±16.7| < 0.01|| ∗∗TotalScore∗∗| ∗∗69.9±4.0∗∗| ∗∗80.7±$7.3**</td><td><0.01</td><td></td></tr>
</table>

Figure 1 Partial display of the thyroid nodule ultrasound examination system interface

Note: Figure translations are in progress. See original paper for figures.