AIBx, Artificial Intelligence Model to Risk Stratify Thyroid Nodules

Johnson Thomas, Tracy Haertling

***Thyroid*. 2020**;[published online ahead of print, 2020 Mar 11]. doi:10.1089/thy.2019.0752

Summary of the paper:

Artificial intelligence (AI) algorithms have been used to decrease subjectivity in medical image interpretation. In this study, authors described an image similarity algorithm based on deep learning for thyroid nodule risk stratification. Ultrasound (US) images of thyroid nodules from patients who underwent either biopsy or thyroid surgery from February 2012 to February 2017 in authors institution were used to create AI models. Nodules were excluded if there was no definitive diagnosis of it being benign or malignant. A total of 482 nodules met the inclusion criteria and all available images from these nodules were used to create the AI models. Later, these AI models were used to test 103 thyroid nodules that underwent biopsy or surgery from March 2017 to July 2018. When compared with published results of ACR-TIRADS and ATA classification system, AIBx, the image similarity model had comparable NPV (93.2%) with better sensitivity (87.8%), specificity (78.5%), and PPV (65.9%). Hence, authors concluded that by using image similarity AI models, we can decrease subjectivity and decrease the number of unnecessary biopsies. This algorithm may also aid in the management of indeterminate and nondiagnostic thyroid nodules. Using image similarity AI model, authors were able to create an explainable AI model that encourages physician's confidence in model predictions.

Commentary:

Thyroid lesions are common and owing to increasing prevalence of thyroid cancer worldwide the effective management of thyroid nodules with accurate selection for surgery has been an ongoing challenge (1). Ultrasound is widely accepted as first-line imaging modality in thyroid diseases as it is noninvasive, cheap, easily available, and cost-effective diagnostic modality. The major obstacle of this technique is that it is observer dependent and its accuracy may differ between different specialties who undertake US examination including radiologists, endocrinologists, and surgeons. In recent years well-recognized scientific societies like American Association of Clinical Endocrinologists (AACE), American Thyroid Association (ATA), American College of Radiologists (ACR), Korean Society of Thyroid Radiology , European Thyroid Association (ETA), and others introduced guidelines for ultrasound malignancy risk stratification of thyroid nodules (2). In general, these guidelines categorize the risk of malignancy in relation to a combination of several US features as no single feature can reliably predict malignancy (1,2). Utility of these guidelines lies in providing a practical image guide for clinical usage allowing for a more accurate selection criteria for fine needle aspiration (FNA) cytology assessment based on standardized US risk of malignancy evaluation. In addition, these guidelines proposed a structured format of US thyroid lesions reporting (1). A growing awareness of these guidelines of all physicians involved in thyroid nodules management resulted in optimized patients pathways.

In recent few years, owing to technological progress in building AI an ultrasound-based computer-aided diagnosis (CAD) system based on semi-automated US image analysis techniques has been developed and introduced to commercially available US machines software. Its utility has been initially validated in breast and thyroid tumors examined by radiologists with promising results (3). Recently, we performed a prospective study of 50 patients who underwent surgeon-performed thyroid US (basic US skills without CAD *vs.* with CAD *vs.* expert US skills without CAD) in the out-patient office as part of the preoperative workup. The real-time CAD system software using AI (S-Detect™ for Thyroid; Samsung Medison Co.) which was used in this study was integrated into the RS85 US system. CAD system added-value to thyroid assessment by a surgeon with basic US skills was equal to 6% (overall accuracy of 82% for evaluation with CAD *vs.* 76% for evaluation without CAD system; $P<0.001$), and final diagnosis was different than predicted by US assessment in 3 patients (1 more true-positive and 2 more true-negative results). However, CAD system was inferior to thyroid assessment by a surgeon with expert US skills in 6 patients who had false-positive results ($P<0.001$) (4).

The current limitations of AI testing in the diagnosis of thyroid lesions and nodules are due to several facts: most of the published studies are retrospective, single-center ones, comprised of small patient cohorts, and the model of malignancy is papillary thyroid cancer. To date, no data have been published indicating the usefulness of these systems in the differential diagnostics of indeterminate FNA cytology (e.g. Bethesda III-IV). Also, the assessment of the added value of these systems for diagnoses made by experienced radiologists or other specialists performing ultrasound examinations is still pending to be defined by independent groups of clinicians.

Current artificial intelligence development has a diagnostic performance that is comparable with medical experts, especially in image recognition-related fields like US thyroid imaging (5). However, future technical improvements in automatic image recognition and diagnosis systems based on deep learning using the neural network may shift the current format of CAD into a brand new dimension of real AI leading to increased accuracy as well as diagnostic efficiency of the CAD imaging modality (6,7).

References:

1. 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. *Thyroid*. 2016;26:1-133.
2. Kim PH, Suh CH, Baek JH et al. Diagnostic Performance of Four Ultrasound Risk Stratification Systems: A Systematic Review and Meta-analysis [published online ahead of print, 2020 Apr 17]. *Thyroid*. 2020;10.1089/thy.2019.0812. doi:10.1089/thy.2019.0812
3. Choi YJ, Baek JH, Park HS et al. Computer-Aided Diagnosis System Using Artificial Intelligence for the Diagnosis and Characterization of Thyroid Nodules on Ultrasound: Initial Clinical Assessment. *Thyroid*. 2017;27:546-52.
4. Barczyński M, Stopa-Barczyńska M, Wojtczak B, Czarniecka A, Konturek A. Clinical validation of S-Detect^{TM} mode in semi-automated ultrasound classification of thyroid lesions in surgical office. *Gland Surg*. 2020;9(Suppl 2):S77-S85. doi:10.21037/gs.2019.12.23
5. Prochazka A, Gulati S, Holinka S et al. Classification of Thyroid Nodules in Ultrasound Images Using Direction-Independent Features Extracted by Two-Threshold Binary Decomposition. *Technol Cancer Res Treat*. 2019;18:1533033819830748.
6. Nguyen DT, Pham TD, Batchuluun G, Yoon HS, Park KR. Artificial Intelligence-Based Thyroid Nodule Classification Using Information from Spatial and Frequency Domains. *J Clin Med*. 2019;8(11):1976. Published 2019 Nov 14. doi:10.3390/jcm8111976
7. Zhou H, Jin Y, Dai L, et al. Differential Diagnosis of Benign and Malignant Thyroid Nodules Using Deep Learning Radiomics of Thyroid Ultrasound Images. *Eur J Radiol*. 2020;127:108992. doi:10.1016/j.ejrad.2020.108992

Reviewer:

Marcin Barczyński, MD, PhD, FEBS-ES,

Deapertment of Endocrine Surgery, Third Chair of General Surgery,

Jagiellonian University Medical College,

Kraków, Poland