Combining Image Similarity and Predictive AI Models to Decrease Subjectivity in Thyroid Nodule Diagnosis and Improve Malignancy Prediction

Govind Nair, Saint Louis University; Aishwarya Vedula, Princeton High School; Johnson Thomas, MD, FACE

Background/Problem to be solved
Nodularity of the thyroid tissue is extremely common. Up to 67% of older adults might harbor thyroid nodules. Current classification systems for thyroid nodules are very subjective. Definitive way to exclude malignancy is through invasive procedures such as Fine Needle Aspiration or surgery. This creates significant emotional burden to the patient and can contribute to rising health care costs. Hence, there exists a need for noninvasive objective classification.

Intervention(s)
Artificial intelligence (AI) algorithms have been used to decrease subjectivity in medical imaging interpretation. However, since some thyroid cancers have benign appearance on ultrasonography, we used an AI algorithm that used image similarity algorithm rather than an image classification algorithm to generate diagnosis. Then we used a TI-RADS prediction algorithm to generate human understandable descriptors. This is unique in that image similarity and predictive models have not been combined before for thyroid nodules. Thereby providing explanations for the AI decisions which can be verified by humans. Publicly available Stanford dataset consisting of ultrasound images of 192 biopsy-confirmed thyroid nodules between April 2017 to May 2018 was used to determine the efficacy of AI in predicting a histopathological diagnosis and TI-RADS recommendations. Prevalence of malignancy was 8.85%. Image with the largest nodule area was used for the analysis. AI algorithm examined regions of interest to predict the benignity and malignancy risk of the thyroid nodules based on similar images. Descriptions and TI-RADS score was generated using a CNN.

Outcome
There were 17 malignant nodules and 175 benign nodules. Compared to ground truth, the AI algorithm predicted malignancies with a sensitivity of 1.0 (95% CI 1 to 1) and a specificity of 0.55 (CI: 0.48,0.63). The PPV was 0.18 and the NPV was 1.0. The AUROC was 0.78 (95% CI: 0.68, 0.87). Our AI algorithm did not miss any cases of cancer. TI-RADS based clinical recommendation had a polychoric correlation of 0.67.

Conclusion
In this external validation study, our model was able to predict the histopathologically proven presence of malignancy accurately. The AI algorithm was able to reduce unnecessary biopsy in 50.5% of the nodules.

Statement of Impact
This AI algorithm which combines image similarity and predictive models can potentially decrease subjectivity in thyroid nodule interpretation, leading to less unnecessary invasive procedures by about 50% and thereby reducing healthcare spending.
Diagram of Process

**Keywords**
Thyroid Nodules; TI-RADS; Image Similarity; Thyroid Cancer; Ultrasound; Cost Effectiveness