SBI ACR Breast Imaging Symposium, April 2020

Can an Artificial Intelligence Decision Aid Decrease False Positive Breast Biopsies?

Retrospective Study With vs Without AI Decision Support

Yiming Gao, MD, New York University School of Medicine

- **AI decision support may help improve sonographic diagnostic accuracy, particularly in cases with low reader confidence**
- **65.2% of benign biopsies could have been avoided with no significant impact on sensitivity**

Objective

To evaluate the effect of an artificial intelligence (AI) support system (Koios DS Breast) on diagnostic accuracy of breast ultrasound.

Clinical Relevance

Ultrasound is routinely used for supplemental screening, but has limited specificity.

Materials and Methods

- IRB-approved retrospective study of 200 lesions (155 benign, 45 malignant) randomly selected from ultrasound-guided biopsied cases.
- Two readers evaluated the lesions with and without an FDA-approved AI software (Koios DS Breast) in randomized order, with four weeks between sessions.
- Margin, echotexture, reader BI-RADS rating (1–5), reader's confidence level (1–5), and system BI-RADS rating (via reader-drawn region of interest) were recorded for each lesion.
- Performed statistical analysis for diagnostic accuracy, NPV, PPV, sensitivity, specificity of reader BI-RADS rating vs. system BI-RADS rating.
- Used GEE analysis to evaluate reader vs. system accuracy with respect to lesion features and AI impact on low-confidence scores. AI effect on false positive biopsy rate was also determined. Statistical tests were conducted at the two-sided 5% significance level.

Results

- There was no significant difference in overall accuracy (73 vs. 69.8%), NPV (100% vs. 98.5%), PPV (45.5 vs. 42.4%), sensitivity (100% vs. 96.7%), specificity (65.2 vs. 61.9), p= 0.118-p=0.409) comparing AI to pooled reader assessment.
- AI was more accurate than readers for irregular margins (74.1% vs. 57.4%, p=0.002), hypoechoic lesions (76.2% vs. 69.8%, p=0.032), and fibrocystic change (72.3% vs. 60.8%, p=0.049).
- AI was less accurate for round shaped (26.5% vs. 50.0%, p=0.049) and anechoic lesions (62.9% vs. 91.4%, p=0.008).
- AI significantly improved diagnostic accuracy for reader-rated low confidence cases with increased PPV (24.7% vs. 19.3%, p=0.004) and increased specificity (57.8% vs. 44.6%, p=0.008).
- If AI had been used in clinical practice, 65.2% (101/155) of benign biopsies could have been avoided.

CONCLUSION

AI decision support may help improve sonographic diagnostic accuracy, particularly in cases with low reader confidence, thereby decreasing false positives without negatively impacting sensitivity.