
AI Improves BI-RADS Accuracy in Mammography



Breast cancer continues to be one of the most prevalent and deadly diseases affecting women globally. Early detection remains crucial, and mammography has long been the cornerstone of breast cancer screening programmes. However, mammographic interpretation is susceptible to high false-positive and false-negative rates, particularly in women with dense breast tissue or indeterminate findings. The Breast Imaging Reporting and Data System (BI-RADS) provides a framework for interpreting and reporting mammograms, yet categories 0, 3 and 4 are especially prone to uncertainty. In response to these challenges, researchers have developed an artificial intelligence system (AIS) that leverages deep learning to assist radiologists in detecting malignancies and refining BI-RADS categorisation. This system has shown promise in improving diagnostic accuracy and reducing unnecessary procedures.

A Deep Learning System for Breast Cancer Detection

The newly developed AIS was designed using mammographic data from over 12,000 Asian women and incorporated a two-step process. The first stage involved dividing mammograms into smaller patches and analysing each for signs of malignancy. These analyses produced heatmaps indicating areas likely to contain cancerous lesions. In the second stage, the system combined full mammographic views with the heatmaps to generate a malignancy score using a deep learning model known as EfficientNet-B0.

Must Read: [Impact of Mammography on Late-Stage Diagnosis](#)

This dual-level approach allowed the AIS to evaluate both localised detail and broader structural context. The results were impressive, with area under the curve (AUC) scores of up to 0.995 in differentiating malignant from non-malignant breasts, and up to 0.988 in distinguishing malignant from benign cases. Such levels of accuracy suggest that the system could be a reliable tool for radiologists, especially in settings where interpretive experience may vary.

Reducing Errors in BI-RADS Classification

The BI-RADS categories 3 and 4 are among the most difficult to interpret, as they span a wide range of malignancy risk. This often results in unnecessary biopsies or, conversely, delayed diagnoses. In a stratified analysis, the AIS demonstrated its ability to correct these misclassifications. It accurately reassigned 83.1% of false positives (BI-RADS 4B and 4C) as benign, thereby potentially reducing unnecessary surgical interventions. Simultaneously, it reclassified 54.1% of false negatives (BI-RADS 3 and 4A) as malignant, which could allow for earlier treatment and improved outcomes.

These enhancements led to an overall reduction in misclassified BI-RADS 3 and 4 cases, and the AIS outperformed standard BI-RADS assessments in sensitivity, specificity and accuracy. For instance, the system achieved AUC scores between 0.892 and 0.988 in validation and test datasets, compared with 0.828 to 0.873 for radiologist-assigned BI-RADS categories. These results support the AIS's potential to provide more precise risk stratification, particularly in dense breast tissue where diagnostic uncertainty is higher.

The system also showed utility in reassessing BI-RADS 0 cases—those where findings are inconclusive and additional imaging is typically required. In a focused evaluation, AIS enabled radiologists to correctly identify malignancies in 7 of 43 confirmed cases, with a specificity of 96.7%. While sensitivity remained low, the ability to confidently rule out malignancy in a large proportion of cases suggests that the system can support more efficient patient management and reduce unnecessary follow-ups.

Enhancing Radiologist Performance through AI Collaboration

To further explore the clinical applicability of AIS, a counterbalanced reader study was conducted involving 10 radiologists of varying experience

levels. Each reader reviewed 1,302 mammograms both with and without AI support. Across the board, performance improved when AIS assistance was available. The average AUC across all readers rose from 0.870 to 0.888, and diagnostic accuracy also increased.

Senior radiologists demonstrated the most significant gains, likely reflecting their ability to integrate AI insights with extensive clinical knowledge. Notably, inter-reader agreement improved with AIS, indicating a harmonising effect on diagnostic interpretation. In complex cases—such as those involving dense glandular tissue—AIS helped readers more confidently upgrade or downgrade BI-RADS scores, resulting in more appropriate clinical recommendations.

These findings suggest that AI can enhance human expertise rather than replace it. By offering a second, data-driven perspective, AIS enables radiologists to make more informed decisions, particularly in ambiguous or borderline cases.

The study highlighted the potential of artificial intelligence to transform mammographic interpretation and breast cancer screening. The AIS demonstrated high diagnostic performance in distinguishing between benign and malignant lesions and significantly improved the accuracy of BI-RADS categorisation. It reduced both over-treatment and under-diagnosis, addressing longstanding limitations in mammography.

While promising, further validation across diverse populations, imaging systems and clinical contexts is needed. The system was developed using data from an Asian cohort and a single imaging vendor, which may limit generalisability. Future research should include broader datasets and prospective studies incorporating clinical variables to fully assess real-world impact.

Nonetheless, this AI model represents a meaningful step towards more precise, efficient and equitable breast cancer diagnostics. As radiology continues to embrace technological innovation, tools like AIS can play a vital role in supporting clinicians and improving patient outcomes.

Source: [Insights into Imaging](#)

Image Credit: [iStock](#)

Published on : Sun, 25 May 2025