Assessment and Validation of an Artificial Intelligence Tool for Breast Cancer Detection in Screening Mammograms

Introduction

Breast cancer is the leading cause of cancer-related death in women worldwide. Mammographic screening is the primary tool for early cancer detection, thereby increasing the survival rate of patients. This study aims to evaluate and validate an Artificial Intelligence (AI) tool designed for the detection of breast cancer in digital screening mammograms. The assessment focuses on the tool's performance, comparing it with other systems and datasets across different populations.

Materials and methods

The evaluation involves three distinct studies. The first utilizes data from Argentina, comprising 155 cases, with 41 recorded findings, including 22 benign and 19 malignant cases based on histopathology. The second study assesses the generalization capability of the AI algorithm using the MBTST dataset from Sweden (University of Lund), consisting of 14,666 cases, including 94 with confirmed malignant findings through biopsy. Lastly, a comparative study with an FDA-approved AI tool involved 388 patients and confirmed diagnostic mammograms.

Results and discussion

In the retrospective analysis of the Argentinean dataset, the AI algorithm exhibited high sensitivity and acceptable specificity for detecting malignant findings achieving a sensitivity of 90% and specificity of 64%, with an AUROC of 0.89 for detecting malignant findings. The evaluation on the Swedish dataset confirmed the tool's generalization capabilities, maintaining performance levels across different populations with an AUROC of 0.87.Furthermore, the comparative study showcased Entelai Mammo's superior sensitivity while still maintaining a reasonable level of specificity compared to an FDA-approved tool, revealing a 53% higher sensitivity ($P<0.0001$) and a 25% lower specificity ($P=0.4$) for breast cancer diagnosis. Both tools significantly reduced workload by 60-72%, emphasizing their potential impact on clinical efficiency.

Conclusions

The findings of this study support the effectiveness of an AI algorithm in breast cancer detection. Notably, the tool demonstrates regional adaptability, generalization capabilities, and comparable performance to other tools. These results underscore the potential of AI as a valuable tool in breast cancer screening, with implications for improving patient outcomes and reducing the burden on healthcare professionals.