ePoster
(educational)

A-236

Designing an AI-CAD Performance and Robustness Validation Protocol for Dutch Breast Cancer Screening

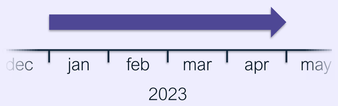
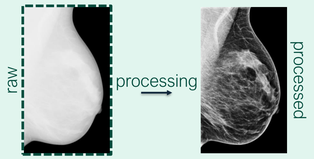
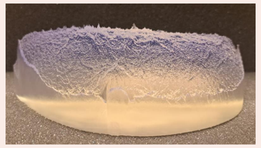
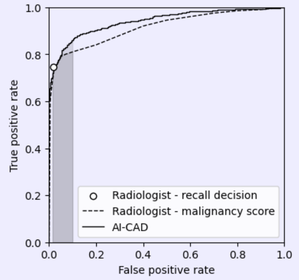
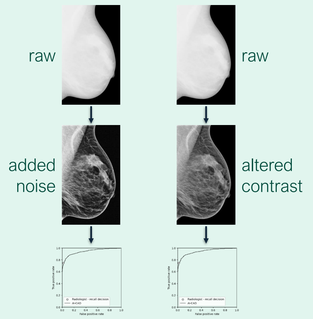
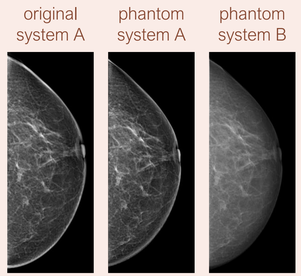
Verboom, Sarah D.[1,2]; Sechopoulos, Ioannis[1,2,3]; Van Engen, Ruben E.[1]; Goris, Liselot C.[2,3]; Smid-Geirnaerdt, Maartje J.A.[1]

[1]Dutch Expert Centre for Screening (LRCB), Nijmegen, the Netherlands – [2]Radboud University Medical Center, Department of Medical Imaging, Nijmegen, the Netherlands – [3]University of Twente, Multi-Modality Medical Imaging (M3I), Enschede, the Netherlands



s.verboom@lrcb.nl

The protocol aims to evaluate whether the AI-CAD performance is non-inferior to that of a breast radiologist under varying circumstances. These circumstances include changes in acquisition or post-processing of the mammograms, since research has shown that these can affect the performance of an AI-CAD model [5, 6].

	1. Performance evaluation	2. Robustness to acquisition and post-processing	3. Robustness across mammography systems
Goal			
Data	Large historic screening cohort  240,000 examinations	Small raw mammogram dataset  1,500 screening examinations	3D-printed patient-based breast phantoms 
Method	Compared Receiver Operating Characteristic curve analysis to performance of an average radiologist  	Simulated acquisition and post-processing differences and extrapolate performance  	Evaluated AI-CAD consistency across matched cases across systems and settings 
Metric	Non-inferiority of sensitivity (at equal specificity) and pAUC to average radiologist	Non-inferiority of sensitivity for different acquisition and post-processing	Stable abnormal interpretation across mammography systems and settings

- How to validate Artificial Intelligence Computer Aided Detection (AI-CAD) performance when used as an independent second reader.
- How to validate AI-CAD robustness when image characteristics vary across post-processing pipelines and acquisition differences.

State-of-the-art AI-CAD models for breast cancer detection in mammography show comparable or superior performance to radiologists [1, 2, 3]. Many European institutions are implementing or planning to implement AI-CAD, yet practical validation guidelines remain limited [4]. In this poster we describe the development of a validation protocol for AI-CAD used as an independent second reader with arbitration for the Dutch Breast Cancer Screening Program. Other forms of implementation might require different methods or limiting values.

AI-CAD models must be externally validated for both performance and robustness before implementation. The LRCB validation protocol addresses this using historical cohorts, acquisition and post-processing simulations, and patient-based physical phantoms.

[1] Taib, A.G., et al., Keeping AI on Track: Regular monitoring of algorithmic updates in mammography. European Journal of Radiology, 2025.

[2] van Winkel, S.L., et al., AI as an independent second reader in detection of clinically relevant breast cancers within a population-based screening programme in the Netherlands: a retrospective cohort study. The Lancet Digital Health, 2025.

[3] Yoon, J.H., et al., Standalone AI for Breast Cancer Detection at Screening Digital Mammography and Digital Breast Tomosynthesis: A Systematic Review and Meta-Analysis. Radiology, 2023.

[4] Brady, A.P., et al., Developing, purchasing, implementing and monitoring AI tools in radiology: practical considerations. A multi-society statement from the ACR, CAR, ESR, RANZCR & RSNA. Insights into Imaging, 2024.

[5] Verboom, S.D., et al. The impact of common image quality issues in screening mammograms on commercial AI CAD models. SPIE Medical Imaging, 2026.

[6] de Vries, C.F., et al., Impact of Different Mammography Systems on Artificial Intelligence Performance in Breast Cancer Screening. Radiology Artificial Intelligence, 2023.

LRCB
dutch expert centre for screening