



Evaluation of Screening Mammograms Comparing Artificial Intelligence with Traditional Computer Aided Detection

- 
- ▶ Authors: R Mayo, A Watanabe, L Chang Sen, J Leung
 - ▶ Disclosures: None

Target Audience

- ▶ Radiologists
- ▶ Software engineers
- ▶ Breast imaging technologists

Goals and Objectives

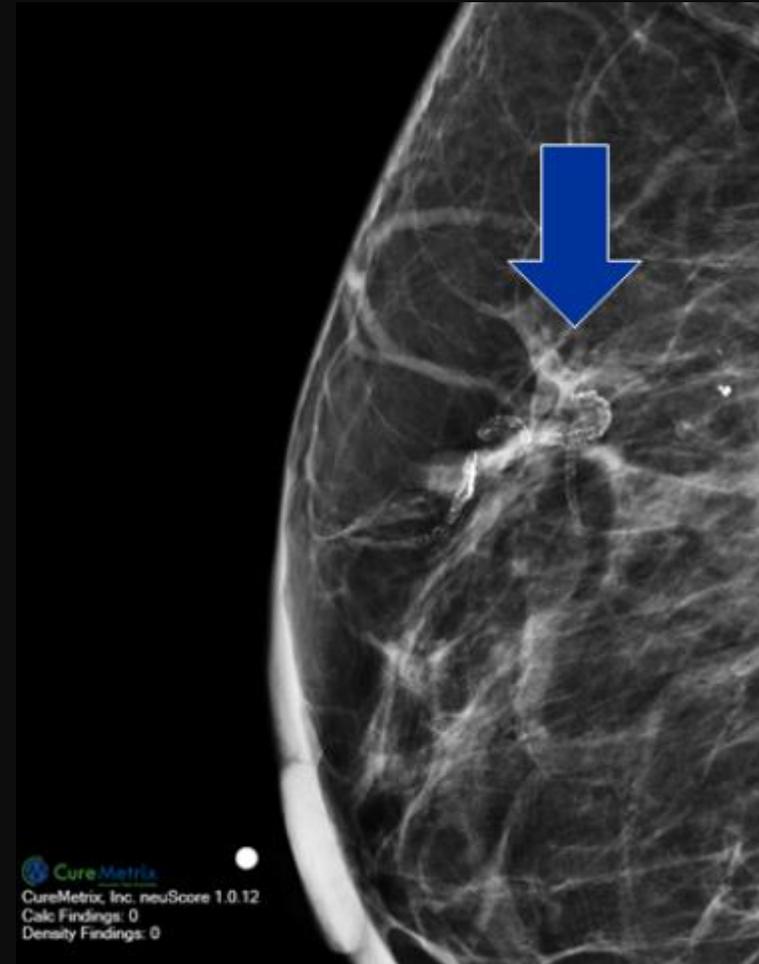
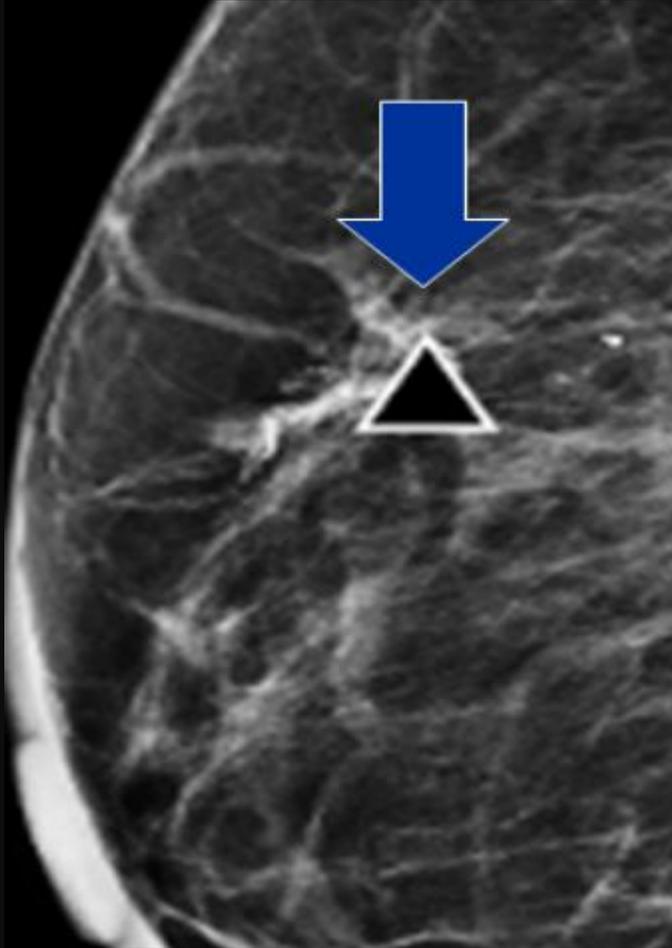
- ▶ Describe the current status of computer aided detection (CAD) in breast imaging
- ▶ Understand the goals of applying artificial intelligence (AI) to image interpretation
- ▶ Summarize the benefit of one currently available AI product compared with traditional CAD

Purpose

- ▶ Determine whether an artificial intelligence (AI)-based, computer-aided detection (CAD) software can be used to reduce False-Positive-Per-Image (FPPI) on mammograms as compared to an FDA-approved conventional CAD.
- ▶ Head to head comparison of products using same data set
- ▶ No such studies currently in literature (at time of this writing)

Materials and Methods

- ▶ Retrospective study
- ▶ 250 (FFDM) performed Jan-Mar 2013
- ▶ Number of CAD marks recorded on each image for each CAD
- ▶ The count of false positive marks per image (FPPI) calculated
- ▶ Number of cases that were completely mark-free counted

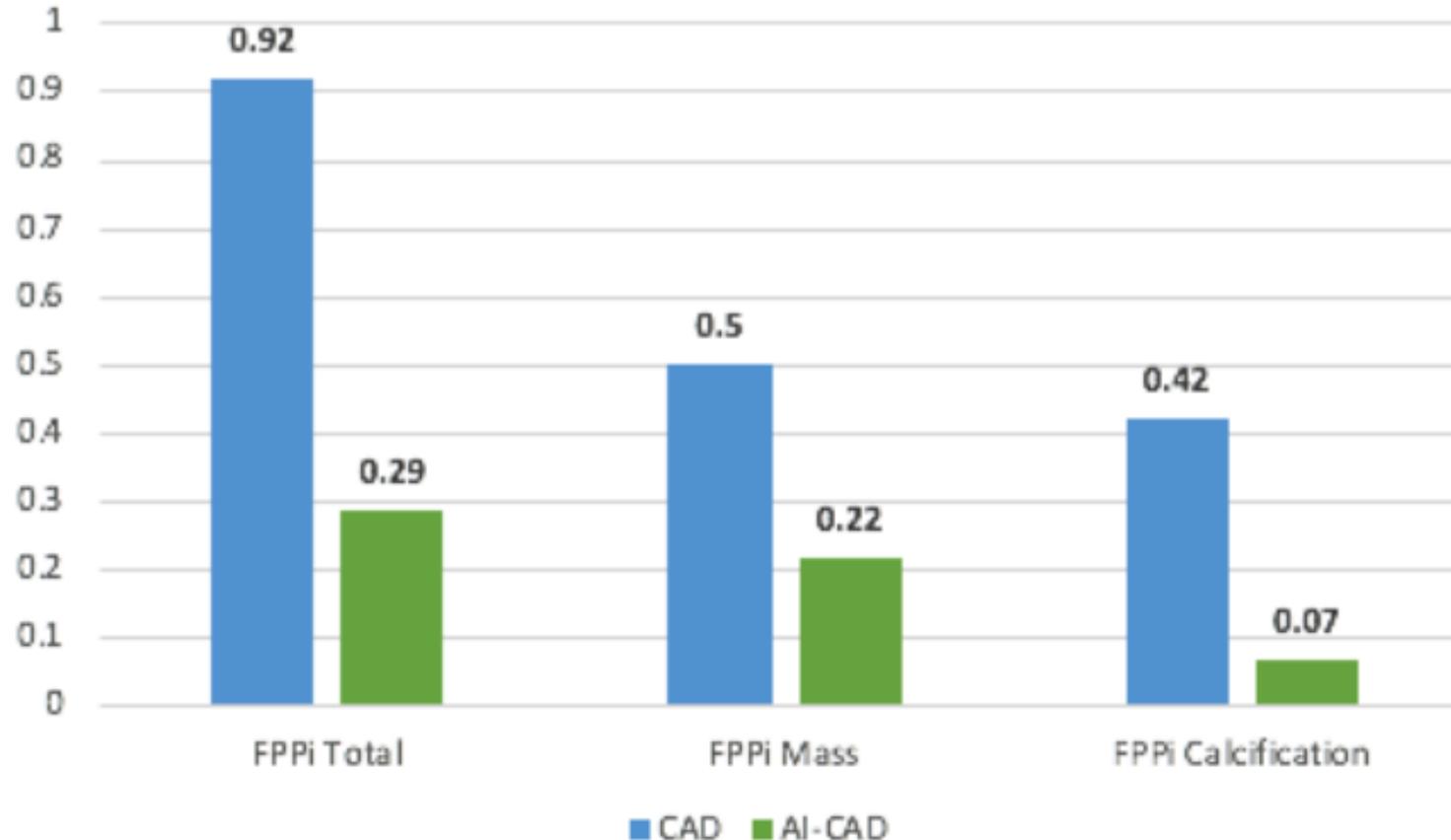


Enlarged cropped RMLO for CAD (left) and AI-CAD (right). False CAD mark (black triangle) of vascular calcifications (blue arrow). AI-CAD (Figure 7b) does not flag the benign vascular calcifications (blue arrow). Elimination of false positive marks reduces distractions to the radiologist and reduces reading time.

Results

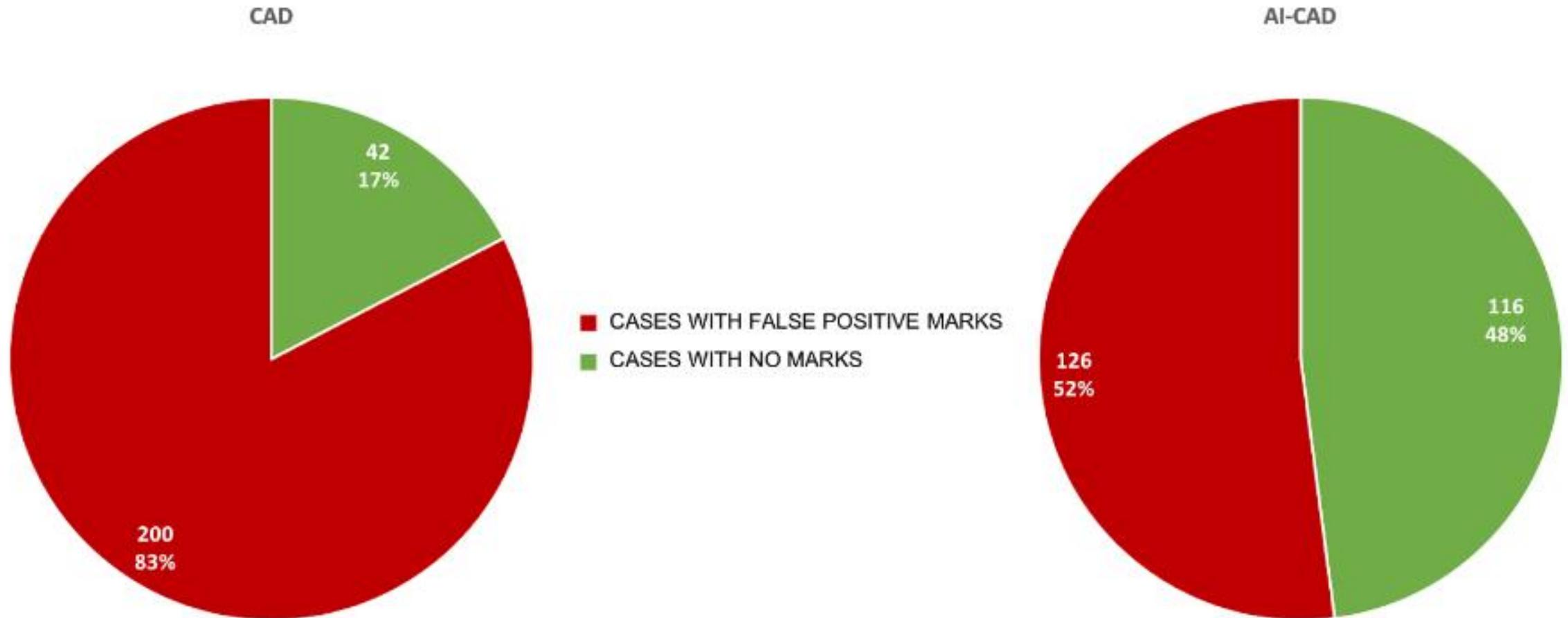
- ▶ Significant reduction in false marks with use of AI-CAD vs CAD (confidence interval = 95%) with no reduction in sensitivity.
- ▶ 69% reduction in false positive per image (FPPI) using the AI-CAD as compared to CAD
 - ▶ 83% reduction for calcifications
 - ▶ 56% reduction for masses.
 - ▶ Reduction in FPPI was consistent across all tissue densities.
- ▶ Almost half (48%) of cases showed no AI-CAD markings while only 17% show no conventional CAD marks.
- ▶ Both systems had 100% sensitivity for cancer detection in this study.

False Positive Marks Per Image



FPPI of CAD vs AI-CAD. There is a significant 69% reduction in FPPI using AI-CAD (green bars) compared to CAD (blue bars). The AI-CAD FPPI reduction was significant for both mass and calcification.

Complete Absence of False Positive Marks



The left graph shows number and percentage of cases with false positive marks for CAD. The right graph shows the same for AI-CAD. 17% of CAD cases were mark-free compared to 48% of AI-CAD cases.

Discussion

- ▶ Any method to improve the performance of mammography interpretation could tremendously affect patient care, radiologist workflow, and system costs
- ▶ Multiple studies in the early and mid-2000's showed variable ability of computer aided detection (CAD) to improve diagnostic performance
- ▶ More recent studies show that CAD may not improve the diagnostic ability of mammography
- ▶ One weakness of currently available CAD is a high rate of false-positive marks.
- ▶ False-positive marks may lead to unnecessary workups and biopsies

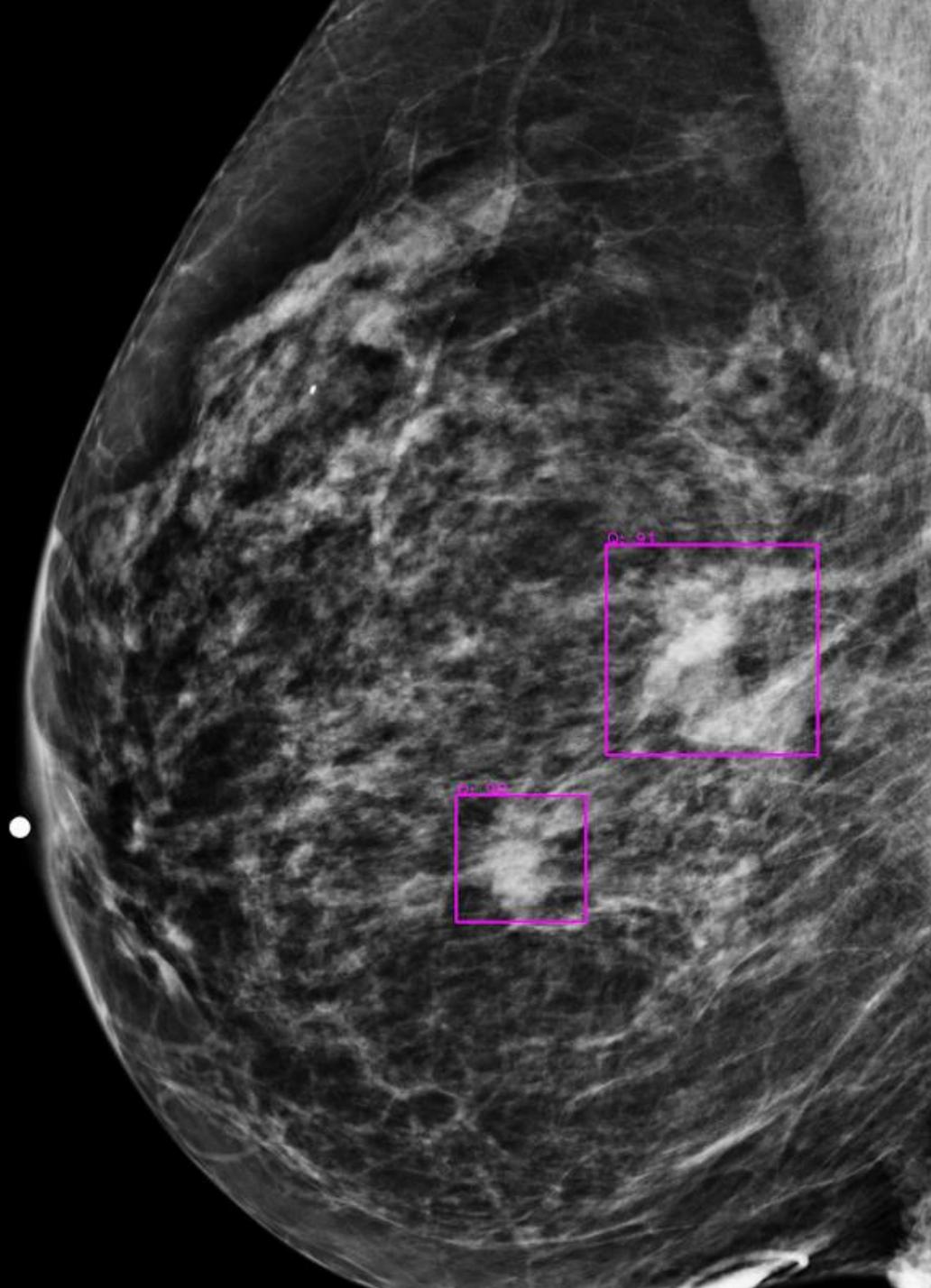
Discussion

- ▶ Artificial intelligence (AI) has recently undergone tremendous advancement with increasing influence throughout the scientific community.
- ▶ Advanced analytics such as deep learning and convolutional neural networks are driving new waves of progress.
- ▶ One of these programs which was designed to assess mammograms is evaluated in this study.
- ▶ To our knowledge and at the time of the study, there are no other head to head comparisons in the peer reviewed literature comparing evaluating FPPI of an AI-CAD to a conventional CAD using the same data set.

Discussion

- ▶ Ultimately the radiologist will make the final decision regarding BIRADS assessment, but their decisions are influenced by CAD marks
- ▶ The significant reduction in FPPI with AI-CAD should translate into fewer false recalls, improved workflow, and decreased costs.
- ▶ CAD marks on average added an additional 23 seconds to the reading time in one study
 - ▶ 7.3 seconds to review each mass flag
 - ▶ 3.2 seconds for each calcification flag

Further analysis of marked lesions through the use of deep learning techniques gives AI-CAD the benefit of providing a probability of malignancy with a **neuScore™** which ranges from 0-100. In this case scores of 98 and 91 were given for two locations. Both sites revealed IDC at biopsy



Conclusion

- ▶ There was a significant reduction in FPPI with AI-CAD as compared to CAD for both masses and calcifications at all tissue densities.
- ▶ A 69% decrease in FPPI could result in a 17% decrease in radiologist reading time per case based on prior literature of CAD reading times.
- ▶ Decreasing false positive recalls in screening mammography has many direct social and economic benefits.
- ▶ Additional studies are warranted as AI techniques advance

References

Lehman CD, et al. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Intern Med* 2015;175(11).

Fenton JJ, Taplin N. Influence of computer-aided detection on performance of screening mammography. *Engl J Med* 2007;356(14):1399–409.

Alcusky M, Philpotts L, Bonafede M, et al. The patient burden of screening mammography recall. *J Womens Health* 2014;23(S1).

Tchou, P et al. Interpretation Time of Computer-aided Detection at Screening Mammography. *Radiology* 2010; 257(1):40-46