Improving Reading Time of Digital Breast Tomosynthesis with Concurrent Computer Aided Detection
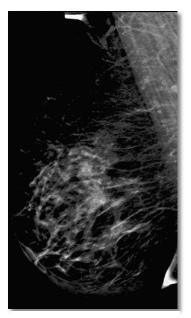
WHITE PAPER

Abstract

PowerLook® Tomo Detection, a concurrent computer-aided detection (CAD) system for digital breast tomosynthesis (DBT) developed by iCAD, Inc., Nashua, NH, was tested to determine if the software can shorten the reading time of DBT without affecting radiologist interpretation or performance. The results of this study showed a 29.2% average reduction in reading time (up to 36.5% reduction) with the use of PowerLook Tomo Detection while maintaining radiologists' performance and accuracy.

Introduction

DBT is one of the most accurate screening methods available for detecting breast cancer. However, this method produces significantly more images than full-field digital mammography (FFDM), therefore DBT requires considerably more time for radiologists to review. A major advancement in DBT would shorten radiologist reading time, while maintaining clinical high clinical performance.
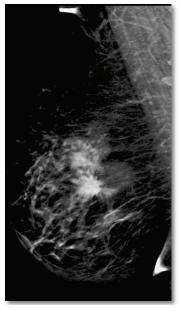
PowerLook Tomo Detection uses DBT images to produce a CAD-enhanced 2D synthetic image. The system detects soft tissue densities (masses, architectural distortions and asymmetries) using the 3D tomosynthesis planes. The detected regions are blended into a corresponding 2D synthetic image to produce the resulting CAD-enhanced synthetic image (Figure 1). The enhanced image assists radiologists in identifying soft tissue densities and provides navigation capabilities from the suspected lesion to the corresponding 3D DBT planes; the lesions are then confirmed or dismissed by the radiologist.

This study [Benedikt et al. 2017] was conducted to evaluate the radiologist's reading time, accuracy and performance with and without PowerLook Tomo Detection. The standalone performance assessed system performance without a radiologist.

Standard 2D Synthetic Image   CAD-Enhanced 2D Synthetic Image   DBT Planes

Figure 1. Use of Concurrent DBT CAD System

Materials and Methods

Study Design

A retrospective multi-reader, multi-case crossover study was conducted in which all readers reviewed images from 240 women in two sessions (Figure 2). Each session was separated by a 4-week washout period. All cases were read by each reader twice, once with PowerLook Tomo Detection and once without. Cases read with Tomo Detection comprised the experimental group while cases read without served as controls. The case reading order was individually randomized for each radiologist. Radiologists independently reviewed exams from randomly selected cases.
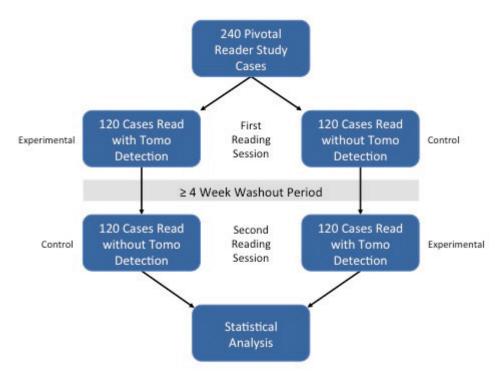

Figure 2: Study Design

Cases

All 240 cases were from women undergoing bilateral screening or diagnostic tomosynthesis exams. Women were excluded from this study if they had a personal history of breast cancer or images indicating prior breast surgery.

Cases were selected based on case type (negative, recalled, benign, or cancer), mammographic appearance, size and histopathology of lesions, breast density, and ACR BI-RADS® assessment categories [Sickles et al. 2013]. The lesions present in all cases were verified by a radiologist with expertise in breast imaging.

Readers

Twenty board-certified radiologists with a range of experience served as readers for this study. All readers were qualified to interpret mammograms under the Mammography Quality Standards Act (MQSA), and had interpreted more than 500 DBT exams in the last two years. Eleven readers devoted less than 75% of their professional time to breast imaging for the last 3 years, while nine readers devoted more than 75% of their professional time to breast imaging. All readers received training on the study reading protocol with 30 additional DBT cases used for practice.

Image Interpretation

Each radiologist read each case, assessed breast density and determined if there were questionable or suspicious lesions that would require recall, short-interval follow-up, or tissue diagnosis. If no questionable lesions were found, the reader recorded the level of suspicion (LOS) score and assigned a BI-RADS assessment category. For each suspicious lesion, the reader recorded the LOS score, BI-RADS category, mammographic appearance including soft tissue density and/or calcifications and marked the location of each lesion using the workstation.

Reading Time

Reading time was measured (in seconds) from the time the reader started examining the images to the time the reader stopped viewing the images. Reading time did not include the time needed for documentation. The reading time timer was hidden from view of all readers.

Standalone Performance

Standalone performance assessed the ability of Tomo Detection to detect malignant lesions without a radiologist. Specifically, this study determined enhanced 2D synthetic image sensitivity for malignant soft tissue densities and mixed lesions at the case-level and lesion-level. Malignant lesions were determined to be correctly detected when the centroid of a CAD detection was within the 3D truth volume established by the expert breast imaging (truthing) radiologist. True positives for enhanced 2D synthetic image sensitivity were lesions correctly detected by PowerLook Tomo Detection and blended into the 2D synthetic image and lesions not detected by Tomo Detection but determined to be visible on the standard synthetic image by the expert radiologist.

Statistical Methods

Radiologist performance was assessed by measuring the area under the curve for the detection of malignant lesions. The area under the curve and reading times were evaluated using a multi-reader multi-case mixed effects analysis of variance method [Obuchowski and Rockette, 1995] with degrees of freedom adjusted for estimation [Hillis, 2007]. A statistical test was used to determine if percent differences in reading time were normally distributed. Because these differences were not normally distributed, a normalizing transformation for percent difference was used and the resulting percent difference was back-transformed to the original scale [Balleyguier et al. 2017]. Sensitivity, specificity and recall rate in non-cancers were also assessed.

Results

Area Under the Curve

The average area under the curve across readers without Tomo Detection was 0.841 and the average across readers with Tomo Detection was 0.850. The readers plots with and without Tomo Detection tended to overlap and cross each other. This study demonstrated non-inferior area under the curve, which indicates similar radiologist performance with and without Tomo Detection.

Reading time improved with Tomo Detection as indicated by 19 out of 20 readers who decreased their reading time with Tomo Detection (Figure 3). The average reading time for the same case was 19 seconds (p<0.01) faster with Tomo Detection compared to the same case without. These data demonstrate a 12.6 to 26.6 second reduction in reading time with Tomo Detection. For percent difference, the average reduction in reading time was 29.2% (p < 0.01) with Tomo Detection. This corresponds to average reading times of 46.3 seconds with Tomo Detection compared to 65.3 seconds without. Radiologist reading time with Tomo Detection was shorter than radiologist reading time without CAD.
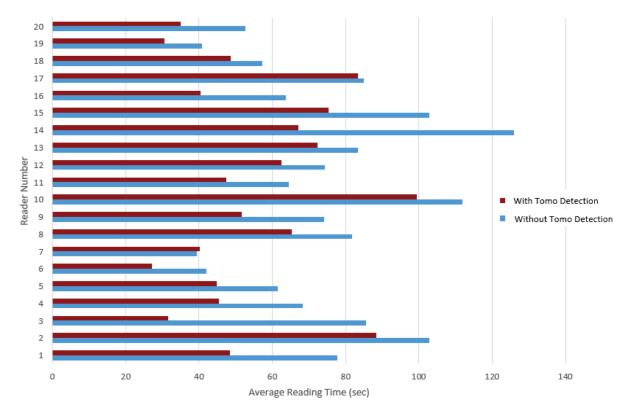
Figure 3: Average Reading Times for Each Reader With and Without Tomo Detection.

Sensitivity, Specificity and Recall Rate

Per-subject sensitivity increased with Tomo Detection overall (from 84.7% without to 87.1% with Tomo Detection) and in cancer cases with at least one soft tissue density or mixed lesion (from 83.7% without to 87.1% with Tomo Detection). For cancer cases with only calcifications, per-subject sensitivity slightly decreased (from 88.1% without to 87.3% with Tomo Detection). The average lesion-level sensitivity with Tomo Detection was 84.5%, which was higher than without Tomo detection (82.3%). The average specificity with Tomo Detection was 50.9%, which was slightly lower than without (52.7%). Average recall rate for non-cancers increased a small amount with Tomo Detection (49.2%) compared to without (47.4%).

Standalone Study

Results of the standalone study indicate the proportion of biopsy-proven malignant lesions detected by PowerLook Tomo Detection as soft tissue densities were 91.7% (44 out of 48) at the case-level and 92.3% (48 out of 52) at the lesion-level (Table 1). Tomo detection sensitivity for malignant soft tissue densities and calcifications was lower, 72.1% (44 out of 61) at the case-level and 70.6% (48 out of 68) at the lesion-level, because the system was not designed to detect calcifications. In contrast, the sensitivity of the standard 2D system was much lower for soft tissue densities without (case-level 52.1%, 25 out of 48; lesion-level 51.9%, 27 out of 52) and with calcifications (case-level 62.3%, 38 out of 61; lesion-level 63.2%, 43 out of 68) at both the case- and lesion-levels.

The proportion of biopsy-proven malignant lesions visible with the standard 2D synthetic images for soft tissue densities was 52.1% (25 out of 48) at the case-level and 51.9% (27 out of 52) at the lesion-level. In contrast, the proportion of biopsy-proven malignant lesions visible in the enhanced

2D synthetic image for soft tissue densities was considerably higher: 91.7% (44 out of 48) at the case-level and 90.4% (47 out of 52) at the lesion-level

Although PowerLook Tomo Detection was designed for soft tissue densities only, performance with calcifications was included and enhanced 2D synthetic image sensitivity for soft tissue densities and calcifications was 91.8% (56 out of 61) at the case-level and 91.2% (62 out of 68) at the lesion-level.

Performance Measure	Case-Level with 95% CI*	Lesion-Level with 95% CI*
Tomo Detection Sensitivity for Soft Tissue Densities	0.917 (44/48) (0.804, 0.967)	0.923 (48/52) (0.849, 0.997)
Tomo Detection Sensitivity for Soft Tissue Densities and Calcifications[†]	0.721 (44/61) (0.598, 0.818)	0.706 (48/68) (0.588, 0.824)
Standard 2D Synthetic Sensitivity for Soft Tissue Densities	0.521 (25/48) (0.383, 0.655)	0.519 (27/52) (0.372, 0.667)
Standard 2D Synthetic Sensitivity for Soft Tissue Densities and Calcifications[†]	0.623 (38/61) (0.497, 0.734)	0.632 (43/68) (0.506, 0.758)
Enhanced 2D Synthetic Image Sensitivity for Soft Tissue Densities	0.917 (44/48) (0.804, 0.967)	0.904 (47/52) (0.824, 0.984)
Enhanced 2D Synthetic Image Sensitivity for Soft Tissue Densities and Calcifications[†]	0.918 (56/61) (0.822, 0.964)	0.912 (62/68) (0.844, 0.979)

* CI represents confidence intervals.
† Calcifications are not detected by PowerLook Tomo Detection.

Table 1: PowerLook Tomo Detection Standalone Performance

Conclusion

PowerLook Tomo Detection creates a CAD-enhanced 2D synthetic image in which a radiologist can recognize the malignant soft tissue densities with 91.7% sensitivity at the case level with only the Tomo Detection-enhanced image (standalone data). With a radiologist also interpreting the DBT planes, the system enables the review of DBT images 29.2% faster, with similar performance and accuracy.

References

1. Obuchowski NA, Rockette HE. Hypothesis testing of diagnostic accuracy for multiple readers and multiple tests an anova approach with dependent observations. *Communications in Statistics - Simulation and Computation* 1995; 24:285-308
2. Hillis SL. A comparison of denominator degrees of freedom methods for multiple observer ROC analysis. *Statistics in Medicine* 2007; 26:596-619
3. Sickles E, D'Orsi C, Bassett L, et al. ACR BI-RADS® Atlas Breast Imaging Reporting and Data System. *Reston, VA: American College of Radiology* 2013

Acknowledgements

iCAD

DMM232 Rev. A