Can digital breast tomosynthesis replace full-field digital mammography?

By Dr. D Georgian-Smith & Dr. T Mertelmeier

This article summarizes the results of a recent trial to evaluate whether the performance characteristics of 2-view DBT alone, i.e. without an FFDM or synthesized 2D image, were superior to those of 2-view FFDM.

The addition of the 3D mammographic modality, digital breast tomosynthesis (DBT), to 2D mammography has had a dramatic effect on breast cancer screening as evidenced by studies which have shown that the addition of DBT can result in an improvement in sensitivity or specificity or both [1–3]. In contrast to CT, where, in terms of the equipment needed and the impression/look of the generated images, it is clear that CT is markedly different to a straightforward planar x-ray, DBT is actually quite similar to full-field digital mammography (FFDM), at least by these criteria. Thus DBT is still recognizable as mammography, but only better.

Against this background, it is reasonable to wonder whether a 3D tomosynthesis scan alone is sufficient for the detection and diagnosis of breast cancer.

The manufacturers of DBT systems have developed units with varying imaging geometries and technical characteristics, such as different sweep angles. While the early clinical studies of DBT were carried out on narrow-angle units (Hologic, Bedford, MA, US) for the simple reason that this was the first system to have U.S. Food and Drug Administration approval, more recent studies have been conducted using wide-angle DBT systems, e.g. from Siemens Healthineers (Erlangen, Germany). A wide-angle system was used in the study described in this article.

Almost all major clinical studies published to date on breast tomosynthesis were designed to compare the diagnostic accuracy of tomosynthesis in adjunct to a 2D image — be it an FFDM image or a synthesized 2D view, i.e. a single 2D image constructed by software from the individual slices of a complete tomosynthesis exam — and were carried out with a narrow angle tomosynthesis system. Currently the only example of a large-scale study carried out using a wide angle system is the Swedish Malmö screening trial [4] which compared DBT used as a stand-alone, to FFDM.

The objective of our study [5] was to evaluate whether 2-view DBT alone, without any additional information from FFDM or a synthesized overview 2 D image, was superior to a 2-view FFDM. If that were found to be the case, significant advantages could be expected in terms of radiation dose and reading time/workflow. To investigate this hypothesis, a multi-reader, multi-case study was performed using mammograms from asymptomatic and an enriched, population including patients with known malignancies [Figure 1] .

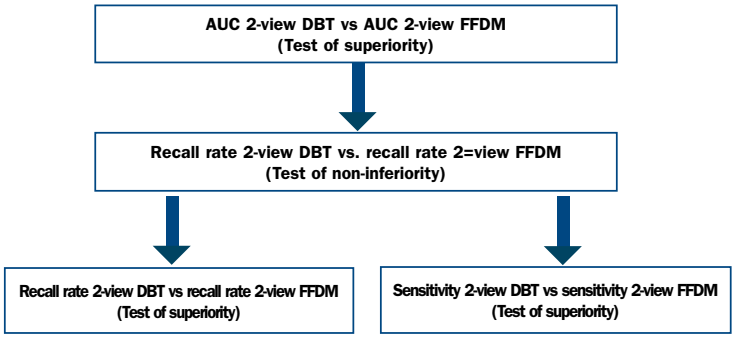

Figure. 1 - Hierarchical analysis plan. If the superiority of DBT versus FFDM is shown, the recall rates would then be compared in a test of non-inferiority of DBT. If non-inferiority is shown, then superiority of the recall rate and sensitivity of DBT would be tested. *Ilmage reproduced from Ref [5], courtesy of American Roentgen Ray Society.*

METHODS

Thirty one readers, selected from both dedicated breast imaging and general radiology practices in the U.S., interpreted 330 patient-cases, being shown either the DBT or the FFDM images in two separate sessions more than one month apart for wash-out reasons.

The study group was selected from a collection database of 764 diagnostic and screening cases collected from seven U.S. sites with DBT images having been acquired using the wide-angle tomosynthesis system. The study had several endpoints: mean area under the curve (AUC) for the receiver operating characteristics (ROC) per patient/subject and per breast; non-cancer recall rates; the sensitivity and specificity of breast cancer detection.

The Authors

Dr. Dianne Georgian-Smith[1] & Dr. Thomas Mertelmeier[2]

1. Associate Professor of Radiology, Harvard Medical School
Brigham and Women's HospitalBoston, MA, USA

2. Senior Principal Scientist, Siemens Healthcare GmbH
Erlangen, Germany

Corresponding Author

Dr. D Georgian-Smith, email: dgeorgiansmith@bwh.harvard.edu

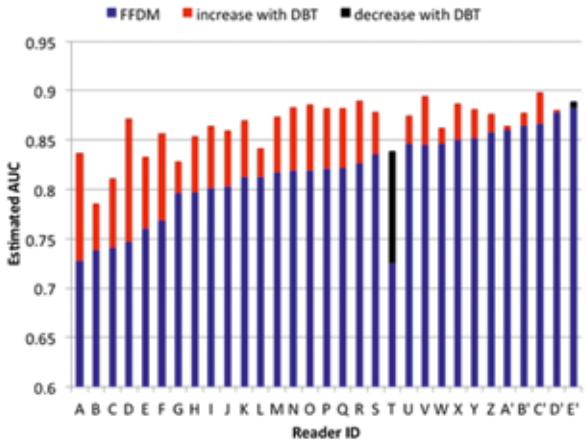

Figure. 2 - Estimated per breast AUC for 31 readers with FFDM (blue) and their improvement with DBT alone (red). The order of the readers is from lowest to highest AUC with FFDM. *Image reproduced from Ref [5], courtesy of American Roentgen Ray Society.*

DBT imaging was carried out on the Siemens Healthineers Mammomat Inspiration system (Siemens Healthcare GmbH, Erlangen, Germany). The system acquires 25 projections over an angle of 50 degrees in less than 25 s, and uses a tungsten/rhodium (W/Rh) anode/filter combination x-ray spectrum. The tube voltages are selected as a function of compressed breast thickness in the same manner as for 2D FFDM. For this trial, the automatic exposure control set the radiation dose of one tomosynthesis scan to 1.5 – 2 times that of a single digital mammogram acquired with the same system. The tomosynthesis slices of 1 mm slice separation were reconstructed with an advanced filtered backprojection algorithm adapted to the tomosynthesis data acquisition sampling scheme. The 2D FFDM images were acquired either with the Mammomat Inspiration in 2D mode or on various other commercial FFDM systems. The FFDM images and the tomosynthesis slices were read on a Siemens Healthineers Syngo MammoReport workstation (Siemens Healthcare GmbH, Erlangen, Germany).

RESULTS

The performance of the majority of the readers (29/31) was better with 2-view DBT — without a synthetic 2D overview — than with FFDM alone, regardless of breast density [Figures 2, 3]. The mean diagnostic accuracy with DBT in terms of the area under the ROC curve (AUC) [Table 1] was statistically significantly improved both per patient (subject-level) and per breast (breast-level): subject-level AUC increased from 0.765 (SE = 0.027) for FFDM to 0.835 (SE = 0.027) for DBT (p = 0.02); breast-level AUC increased from 0.818 (SE=0.019) FFDM to 0.861 (SE=0.019) DBT (p = 0.011). The non-cancer recall rate per patient was reduced by 19% using DBT (p<0.001). More masses/architectural distortions were detected with DBT (p < 0.001). The detection of microcalcifications using DBT tended to be lower than in FFDM (p = 0.136) but the difference was not statistically significant. The accuracy of the detection of invasive cancers was significantly higher with DBT (p < 0.001) [Table 2].

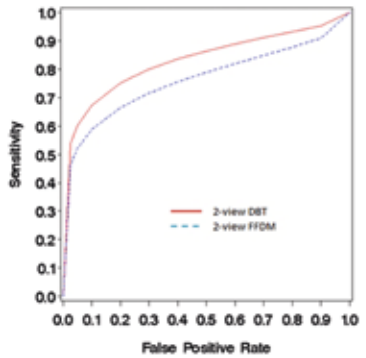

Figure 3. Summary parametric subject-level ROC curves. The AUC for FFDM (blue, dashed line) is 0.765; the AUC for DBT (red, solid line) is 0.835. *Image reproduced from Ref [5], courtesy of American Roentgen Ray Society.*

CONCLUSIONS

This reader study tested a "replacement" hypothesis that DBT alone would be superior to 2D FFDM. The study did not test the effect of synthetic 2D views. A wide-angle DBT unit was used exclusively for tomosynthesis imaging. The results clearly indicate the superior performance of wide-angle DBT alone vs. FFDM, although there was a non-statistically significant decrease in the detection of micro-calcifications, for which the study was not powered.

In a study involving a population of similar size and type, but comparing narrow-angle (15°) DBT together with FFDM to FFDM alone, a similar improvement of the DBT AUC was found [11]. This study however involved a higher radiation dose because a 2D acquisition was required in addition to the DBT scan. In the study the improvement in the detection of microcalcifications was not statistically significant either.

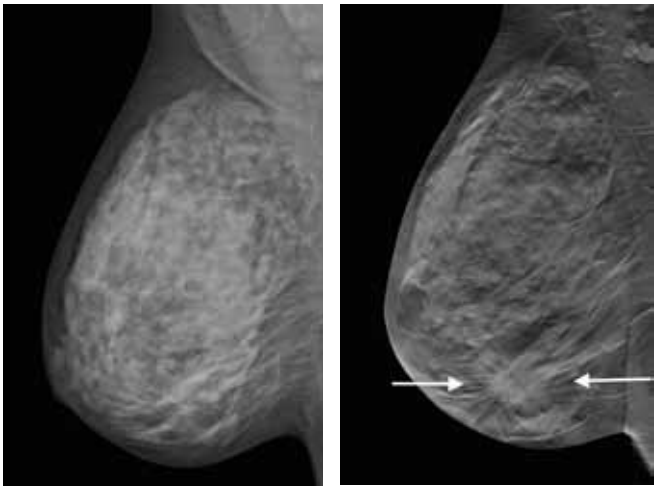

Figure 4. Left panel. FFDM image. Right Panel. A single slice from the DBT exam of the same patient. The spiculated mass (indicated by the arrows) is clearly more visible than on FFDM. (*Images courtesy of the Department of Translational Medicine, Diagnostic Radiology, Lund University, Skåne University Hospital Malmö, Sweden.*)

In the meantime, the higher performance of the wide-angle acquisition has also been confirmed by the real-life Malmö screening trial [4]. The results from Sweden are comparable to those of earlier studies using narrow-angle technology combined with 2D images, albeit at a much lower radiation dose as only the DBT MLO-view was used in the Malmö trial. The high diagnostic performance of the single DBT view might be explained by the higher mass conspicuity that comes with a wide-angle tomosynthesis system [6–9] and by the experience of the readers in reading one-view DBT [10]. Even with such strong positive results suggesting that DBT alone may replace an FFDM image, there is a body of literature based on visual and perception studies arguing for the advantages of first seeing an overview 2D image — either a direct FFDM or possibly a synthetic 2D view. Eye-tracking studies have shown that the efficiency of experienced, expert radiologists is improved if they

	2-view FFDM	2-view DBT	Δ p
Breast-level non-parametric AUC*	0.818 (0.019)	0.861 (0.019)	0.043 0.011
Subject-level parametric AUC*	0.765 (0.027)	0.835 (0.027)	0.070 0.002
Subject-level non-cancer recall rate*	0.479 (0.027)	0.386 (0.025)	-0.093 <0.001
Subject-level sensitivity BIRADS 3, 4, 5[(b)]	0.774 (0.032)	0.819 (0.030)	0.045 0.122
Breast-level sensitivity BIRADS 4, 5[(a)]*	0.715 (0.034)	0.772 (0.032)	0.057 0.047
Breast-level specificity BIRADS 4, 5[(a)]*	0.807 (0.017)	0.848 (0.015)	0.041 <0.001
Breast-level sensitivity BIRADS 3, 4, 5[(b)]	0.762 (0.031)	0.807 (0.030)	0.045 0.109
Breast-level specificity BIRADS 3, 4, 5[(b)]*	0.718 (0.019)	0.773 (0.017)	0.055 <0.001

Table 1: Results data of the trial. Area Under ROC Curves, Recall Rates, Sensitivity and Specificity for DBT compared to FFDM. Differences are denoted by Δ, and the p-value by p. The numbers in brackets are standard errors. (a) sensitivity and specificity counting BIRADS 4 and 5 as positive test (b) sensitivity and specificity counting BIRADS 3, 4, and 5 as positive test . * denotes statistically significant at p< 0.05 level. *Table adapted from Ref [5], courtesy of American Roentgen Ray Society.*

know where to focus their gaze as a result of first being presented with the overview. That first impression— sometimes known as the "gist" — takes less than a second and occurs subconsciously. Symmetry is also rapidly assessed.

The value of a 2D view synthesized from tomosynthesis slices is that it provides a holistic overview without the need for a separate FFDM examination and the associated radiation exposure. The increased lesion conspicuity is still available in the individual tomosynthesis slices. Future studies with synthesized images are warranted.

	2-view FFDM	2-view DBT	Δ p-value
Dense Breasts	0.802 (0.027)	0.844 (0.026)	0.043 0.106
Non-dense Breasts*	0.826 (0.026)	0.873 (0.026)	0.047 0.024
Masses*	0.858 (0.018)	0.923 (0.018)	0.065 <0.001
Calcifications	0.796 (0.042)	0.749 (0.041)	-0.047 0.136
Architectural distortion*	0.773 (0.038)	0.930 (0.035)	0.157 <0.001
Invasive Cancers*	0.836 (0.018)	0.912 (0.018)	0.076 <0.001

Table 2: Sub-analyses per Breast Density, Mammographic Findings, and Invasive Cancers, Breast-level AUC (SE) stratified for dense (BIRADS categories c or d) and non-dense breasts (BIRADS categories a or b), and for masses and calcifications. Differences Δ and p values are shown in the last column. * denotes statistically significant at p< 0.05 level. *Table adapted from Ref [5], courtesy of American Roentgen Ray Society.*

Since our study was a reader study, neither symptoms, patient history nor prior imaging results were considered. The study was not powered to show small differences in the detection of specific features, such as microcalcifications. Whether these results can be generalized to narrow-angle units has yet to be determined, and is the subject of on-going investigation.

Future developments of CAD systems based on artificial intelligence (AI) and deep learning approaches open the potential that the information content provided by a tomosynthesis scan can be fully exploited without the need for an additional 2D overview image. Deep learning algorithms have already been shown to be as good as an average radiologist in the interpretation of 2D FFDM images [12]. It has also been shown that the use of interactive decision support leads to an increase in cancer detection without increasing the reading time [13], although both these studies were limited, in that the population studied was enriched by cases with known malignancies.

CONCLUSION

This reader study shows that, for most radiologists, wide-angle 2-view DBT alone has a higher diagnostic accuracy than 2-view FFDM, verifying the robustness of wide-angle DBT as a sole view.

REFERENCES

1. Skaane P, Bandos AI, Niklason LT, Sebuødegård S, Østerås BH, Gullien R et al. (2019) Digital Mammography versus Digital Mammography Plus Tomosynthesis in Breast Cancer Screening: The Oslo Tomosynthesis Screening Trial. Radiology 291(1):23–30.
2. Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fanto C, Ostillio L et al. (2016) Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. The Lancet. Oncology 17(8):1105–13.
3. Rose SL, Tidwell AL, Bujnoch LJ, Kushwaha AC, Nordmann AS, Sexton R JR (2013) Implementation of breast tomosynthesis in a routine screening practice: an observational study. AJR. American journal of roentgenology 200(6):1401–8.
4. Zackrisson S, Lång K, Rosso A, Johnson K, Dustler M, Förnvik D et al. (2018) One-view breast tomosynthesis versus two-view mammography in the Malmö Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study. The Lancet Oncology 19(11):1493–503.
5. Georgian-Smith D, Obuchowski NA, Lo JY, Brem RF, Baker JA, Fisher PR et al. (2019) Can Digital Breast Tomosynthesis Replace Full-Field Digital Mammography? A Multireader, Multicase Study of Wide-Angle Tomosynthesis. AJR. American journal of Roentgenology. doi:10.2214/AJR.18.20294.
6. Scaduto DA, Huang H, Liu C, Rinaldi K, Hebecker A, Mertelmeier T, Vogt S, Fisher P, Zhao W (2018) Impact of angular range of digital breast tomosynthesis on mass detection in dense breasts. Proc. SPIE 10718, 14th International Workshop in Breast Imaging (IWBI 2018).
7. Zhao B, Zhou J, Hu Y-H, Mertelmeier T, Ludwig J, Zhao W (2009) Experimental validation of a three-dimensional linear system model for breast tomosynthesis. Medical physics 36(1):240–51.
8. Mertelmeier T, Ludwig J, Zhao B, Zhao W (2008) Optimization of Tomosynthesis Acquisition Parameters: Angular Range and Number of Projections. In: Krupinski EA (ed) Digital Mammography. Springer Berlin Heidelberg. Berlin, Heidelberg, pp 220–227.
9. Zhao B, Zhao W (2008) Three-dimensional linear system analysis for breast tomosynthesis. Medical physics 35(12):5219–32.
10. Rodriguez-Ruiz A, Gubern-Merida A, Imhof-Tas M, Lardenoije S, Wanders AJT, Andersson I et al. (2018) One-view digital breast tomosynthesis as a stand-alone modality for breast cancer detection: do we need more? European radiology 28(5):1938–48.
11. Rafferty EA, Park J-M, Philpotts LE, Poplack SP, Sumkin JH, Halpern EF et al. (2013) Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial. Radiology 266(1):104–13.
12. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P et al. (2019) Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. Journal of the National Cancer Institute. doi:10.1093/jnci/djy222.
13. Rodríguez-Ruiz A, Krupinski E, Mordang J-J, Schilling K, Heywang-Köbrunner SH, Sechopoulos I et al. (2019) Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. Radiology 290(2):305–14.