Can we make breast tomosynthesis feasible for screening?

By Dr Ioannis Sechopoulos

This article summarizes a recently published study comparing the detection performance of several different methods for the introduction of digital breast tomosynthesis to breast screening and gives insight into how screening tomosynthesis could be feasible without a significant increase in the resources required.

THE NEXT BREAST CANCER SCREENING TECHNOLOGY

Over the last decade, digital breast tomosynthesis (DBT) has increasingly been used in the clinic for the evaluation of certain mammographic breast screening findings and of clinical concerns [1]. DBT involves the acquisition of several low-dose mammogram-like images as the x-ray source rotates around the breast; the images are then combined into one pseudo-3D stack of slices [2, 3]. This stack, although having a limited resolution in the vertical direction, provides enough tomographic information on the relative position of normal and pathologic breast tissues so as to minimize the masking impact of tissue superposition. As a result, it has been shown in several studies that the use of DBT has superior detection performance than mammography [4-6].

To exploit this advantage, several prospective screening trials have been performed in Europe, and have again reported an increase in sensitivity when using DBT either as a replacement or an adjunct to digital mammography (DM) [7-9].

However, depicting the breast as a stack of slices as opposed to a single planar (2D) image, as in DM, greatly increases the amount of information shown to the interpreting radiologist. For example, for an average breast that is compressed to a thickness of 5 cm, DBT results in a stack of 50 slices. Even with the use of cine mode and other alternative methods for image display, e.g. user-selected combination of the slices into thicker slabs [10], the interpretation of DBT images has been repeatedly shown to, on average, double the reading time [8]. Especially for population screening programs, such a large increase in the interpretation time means that the introduction of DBT for screening is challenging.

WHAT IS A BREAST TOMOSYNTHESIS SCREENING EXAM?

Therefore, although the overall benefits of DBT in screening are generally well understood, for implementation of screening DBT to take place there are more practical aspects that need to be addressed. Chief among these is what exactly should a DBT screening examination entail? Specifically, screening DBT has been proposed as consisting of different combinations of views (cranio-caudal (CC) and mediolateral oblique (MLO)) of DBT and DM images. For example, in the Malmö Breast Tomosynthesis Screening Trial the DBT arm consisted of acquisition of a single DBT view per breast [7], while in the Oslo Tomosynthesis Screening Trial the DBT arm included the acquisition of both two-view DM and two-view DBT per breast [8]. Clearly, exactly how DBT is eventually implemented in screening will have a major impact on image interpretation time, and therefore on the increase, if any, in resources required.

COMPARISON OF MULTIPLE SCREENING PROTOCOLS

To provide insight into the trade-offs between different screening protocols, in terms of performance and reading time, a retrospective observer study was undertaken with an enriched case set of DBT and DM images [11]. In this study, 181 cases, including 76 malignant, 50 benign, and 55 normal cases were retrospectively reviewed by six radiologists over two sessions. In one session, the radiologist was shown the same case, sequentially using three different combinations of images: first only the MLO view DBT image, then the MLO DBT + the CC view DM, and finally the two-view DBT and two-view DM [Figure 1]. At each condition, the reader was asked to mark and rate all suspicious findings. The scoring involved providing the location and both a probability of malignancy rating (1 to 10, with 10 denoting the highest probability of malignancy) and a BI-RADS® score (1 to 5, as per American College of Radiology guidelines) for each finding. If the reader was convinced that a previously marked finding was incorrect due to the review

The Author

Dr Ioannis Sechopoulos

Department of Radiology and Nuclear Medicine,

Radboud University Medical Centre, Geert Grooteplein 10, Post 766, 6525 GA Nijmegen, The Netherlands

&

Dutch Expert Centre for Screening (LRCB), Wijchenseweg 101, 6538 SW Nijmegen, The Netherlands

Email: Ioannis.sechopoulos@radboudumc.nl

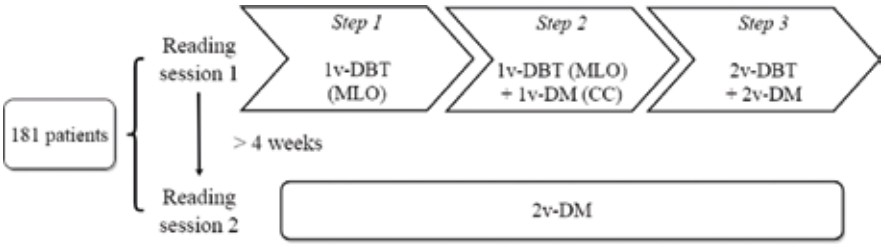

Figure 1. Sequence of image interpretation. Three of the conditions were read sequentially since they implied only addition of images from step to step, while the last condition, involving only two-view DM was interpreted in a separate reading session.
Abbreviations:
1v-DBT : One-view digital breast tomosynthesis (MLO);
1v-DBT +1v-DM: One-view digital breast tomosynthesis (MLO) plus one-view digital mammography (CC);
2v-DM: Two-view digital mammography (CC+MLO);
2v-DBT+2v-DM: Two-view digital breast tomosynthesis (CC + MLO) plus two-view digital mammography (CC + MLO)

Figure reproduced from: Rodriguez-Ruiz A, et al. (2018) Eur Radiol 28: 1938–1948.© the Authors

of additional images at the next step, he/she was able to remove that finding from the collection of findings for the new reading condition. In a second session, separated from the first by at least four weeks, the radiologist was shown the two DM views of the same 181 cases. Again, the observer had to mark and rate any suspicious findings.

The marked locations of the suspicious findings and their corresponding probability of malignancy ratings were analyzed using the jack-knife alternative free-response receiver operating characteristic (JAFROC) method. The benefit of this method over the standard ROC analysis is that location is taken into account. That is, if a true positive case is marked with a high probability of malignancy but only because the reader is concerned about a different, actually normal, area, then the finding is not considered a true positive. Hence, the reader cannot "*be right for the wrong reason*." However, to also allow for the more traditional interpretation of the results, the standard case-based ROC analysis was also performed. Using the BI-RADS® ratings, the actual operating sensitivity and specificity were also calculated, using the threshold of a BI-RADS® 3 or higher rating for the most-suspicious finding in the case as a positive interpretation of the case. Finally, during all interpretations, the observer study software automatically recorded the time taken for interpretation.

Two groups of breast radiologists participated in the trial; three from the Netherlands and three from Sweden. All six readers had experience with DM and DBT, but only the Swedish radiologists had previous experience in interpreting single-view DBT, since they had participated in the Malmö Breast Tomosynthesis Screening Trial. We called this latter group the "single-view-DBT experienced readers".

SO, IS ONE-VIEW DBT ENOUGH?

It seems to be. There was no statistical difference in performance between the single-view DBT interpretation and any of the other three protocols. As can be seen in Figure 2, both the ROC and the JAFROC results showed no statistical difference among protocols. Likewise, using these metrics no difference was found when comparing the results of radiologists unused to making decisions based only on the single-view DBT (i.e. the Dutch radiologists) and those having previously performed single-view DBT interpretations (i.e. the Swedish radiologists).

Interestingly, although there was no difference in sensitivity and specificity across protocols, a difference was found when evaluating the results according to the level of experience in single-view DBT-interpretations. For the single-view-DBT-experienced readers, the inclusion of additional views did not affect their sensitivity or specificity. However, for the readers for whom making decisions based on only the single DBT view was a new experience, their sensitivity did increase with the addition of more views, while their specificity decreased slightly.

Importantly, interpretation of single-view DBT took, on average, only 25% longer (44 s vs. 55 s, per case) than reading of two-view DM, a difference that was found to be statistically significant. Although the use of DBT, even with a single view, instead of DM, still resulted in an increase in reading time, this relatively small increase is a considerable improvement over the doubling in reading time previously reported for DBT compared to DM screening. For three of the readers, there was no statistical difference in the reading time, and no difference was found between the experienced and the inexperienced readers.

Finally, the average glandular dose per case was equal in the single-view DBT and the two-view DM acquisitions.

WHAT DO THESE RESULTS MEAN?

This study provides insight into how screening performance and reading time could be affected by implementing DBT

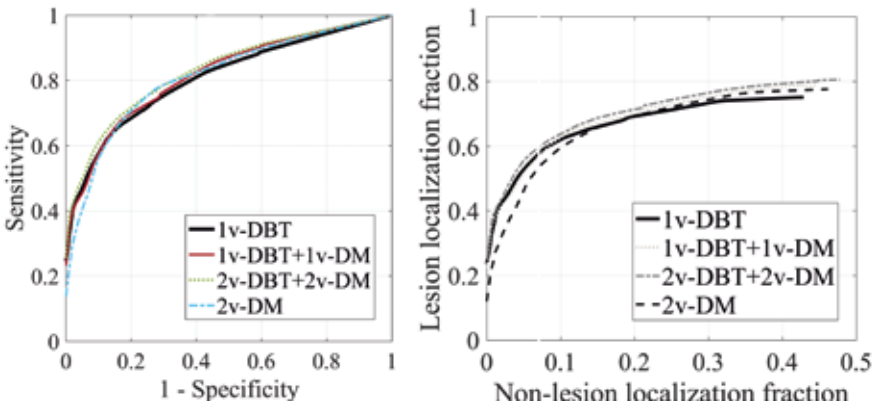

Figure 2. ROC (left panel) and JAFROC (right panel) results of the performance of the six radiologists interpreting all cases under the four different conditions evaluated. None of the conditions studied yielded a statistical significant difference in their performance.
Figure reproduced from: Rodriguez-Ruiz A, et al. (2018) Eur Radiol 28: 1938–1948.© the Authors

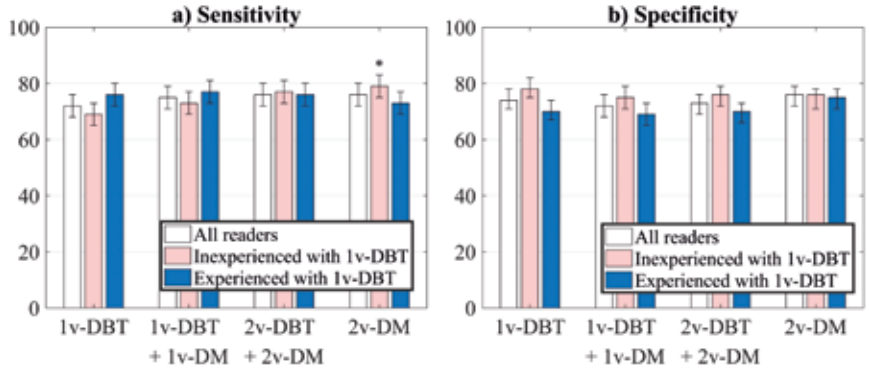

Figure 3. Sensitivity and specificity of each protocol studied, both for all readers combined and separated by experience in interpreting single-view DBT.
Figure reproduced from: Rodriguez-Ruiz A, et al. (2018) Eur Radiol 28: 1938–1948.© the Authors

for screening in different various ways. Most importantly, also as evaluated by the Malmö Breast Tomosynthesis Screening Trial, MLO view-only DBT results in a performance similar to that of standard two-view DM. It seems that the benefit of reduced tissue superposition because of the partial tomographic information provided by DBT outweighs, or at least is equal to, the information introduced by a second view. Although some lesions might be present only within the field of view of the CC compression, the main reason for performing two-view mammography is for the observer to be able to discern lesions from overlapped tissue, increasing sensitivity and specificity. If DBT reduces the tissue overlap effect enough, then the benefit of the second view is limited. This seems to be shown by the current study.

"... MLO view-only DBT results in a performance similar to that of standard two-view DM..."

It should be noted, however, that the images used in this study were acquired with the same DBT system, namely the Siemens Mammomat Inspiration. This DBT system has an angular range for acquisition of the DBT projections of 50° [2], which is the widest angular range currently commercially available. This most probably means it has the highest vertical spatial resolution [12], and therefore the most prominent reduction of the tissue superposition masking effect. Therefore, it is not clear if the results obtained in the current study, just as with the Malmö trial, which also used Siemens systems, is generalizable to all DBT imaging. Further studies with other equipment are needed to evaluate the impact of system design on the possibility of performing single-view DBT.

In addition, at least at the time of image acquisition, the Siemens system used double the dose per DBT acquisition as that of a DM acquisition. This is why even with the acquisition of a single view there were no dose savings. Other DBT systems have a different dose relationship between DBT and DM acquisitions, with some systems even having the same dose for DBT and DM [13]. For such systems acquiring a single DBT would result in a reduction in dose compared to acquiring two DM views. However, dose saving should not be a reason to introduce single-view DBT for screening.

The variation in the trade-off between sensitivity and specificity experienced by the radiologists not used to interpreting only single-view DBT is interesting and also provides some insight as to how observers behave when confronted with a new reading situation. The increase in sensitivity accompanied by the decrease in specificity when more views were added means that the radiologists were actually operating along the same ROC curve, just with a different decision threshold. Their performance was no worse or better than that of the experienced radiologists, they just (most probably unconsciously) needed to be a little more certain about a finding before they decided if it was suspicious enough. However, this means that if it is desired, working at a different sensitivity/specificity trade-off level is only a matter of training or simply getting used to having to interpret only one image per breast. This is the experience that the Swedish radiologists had probably already gone through during the Malmö trial. Although the difference could also be explained by the possibility that there were institutional and national differences in practice, the fact that the operating performance of the experienced readers did not vary when more views were shown — whereas it did for the inexperienced readers — points to there being an additional effect, rather than just simply differences in usual practice methods.

SCREENING WITH DIGITAL BREAST TOMOSYNTHESIS IS FEASIBLE

In short, this study investigated one of the strategies that could be used to introduce DBT for population screening for breast cancer in the near future without fundamentally increasing the resources required to do so. Various prospective screening trials have shown the positive impact of introducing DBT on detection performance, even with a single view per breast, and this study provides further insight on how and why this is feasible.

REFERENCES

1. Vedantham S, *et al.* Digital Breast Tomosynthesis: State of the Art. Radiology, 2015; 277: 663.

2. Sechopoulos I. A review of breast tomosynthesis. Part I. The image acquisition process. Medical Physics, 2013. 40(1).

3. Sechopoulos I. A review of breast tomosynthesis. Part II. Image reconstruction, processing and analysis, and advanced applications. Medical Physics. 2013; 40(1): 014302.

4. Gilbert FJ *et al.* Accuracy of Digital Breast Tomosynthesis for Depicting Breast Cancer Subgroups in a UK Retrospective Reading Study (TOMMY Trial). Radiology, 2015. 277: 697.

5. Friedewald SM *et al.* Breast cancer screening using tomosynthesis in combination with digital mammography. JAMA. 2014 ; 311(24): 2499-2507.

6. Rafferty EA *et al.* Assessing Radiologist Performance Using Combined Digital Mammography and Breast Tomosynthesis Compared with Digital Mammography Alone: Results of a Multicenter, Multireader Trial. Radiology. 2013; 266(1): 104-113.

7. Lång K *et al.* Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study. European Radiology. 2016; 26(1): 184-190.

8. Skaane P *et al.* Comparison of Digital Mammography Alone and Digital Mammography Plus Tomosynthesis in a Population-based Screening Program. Radiology, 2013; 267(1): 47-56.

9. Ciatto S *et al.* Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. The Lancet Oncology, 2013; 14(7): 583-589.

10. Diekmann F *et al.* Thick Slices from Tomosynthesis Data Sets: Phantom Study for the Evaluation of Different Algorithms. Journal of Digital Imaging, 2009; 22(5): 519-526.

11. Rodriguez-Ruiz A *et al.* One-View Digital Breast Tomosynthesis as a Stand-Alone Modality for Breast Cancer Detection: Do We Need More? European Radiology, 2018; 28(5): 1938-1948.

12. Sechopoulos I and Ghetti C. Optimization of the acquisition geometry in digital tomosynthesis of the breast. Medical Physics. 2009; 36(4): 1199-1207.

13. Svahn TM *et al.* Review of radiation dose estimates in digital breast tomosynthesis relative to those in two-view full-field digital mammography. Breast. 2015; 24(2): 93-99.