Digital breast tomosynthesis in screening - does it reduce the recall rate?

By Dr A J Maxwell

Digital breast tomosynthesis (DBT) has rapidly been adopted by the breast screening community as a useful adjunct to, and possibly even a replacement for, conventional 2D digital mammography. Its superior sensitivity to 2D mammography for cancer detection is now well established but the evidence for its effect on screening recall rates is less consistent.

This article summarises the results of a recent randomised UK trial of DBT in screening the most challenging group, the younger women at increased risk. The study demonstrated no significant difference in recall rates between those who were screened with 2D mammography only and those screened with 2D plus DBT.

There was, however, evidence of a radiologist learning curve in reading screening DBT examinations. The potential future role of DBT in screening is discussed.

The Author

Dr Anthony J Maxwell,
Consultant Breast Radiologist,
University Hospital of South Manchester,
Manchester, M23 9LT,
UK
email: anthony.maxwell@manchester.ac.uk

DIGITAL BREAST TOMOSYNTHESIS

Since its commercial introduction several years ago, digital breast tomosynthesis (DBT) has quickly become established as an essential tool in many breast centres. During the DBT examination the x-ray tube of the mammography machine performs a series of low dose exposures of the compressed breast as it swings through an arc, and these are then reconstructed into a series of 1mm thick slices. It is not a true 3D technology, as the appearances on each slice are influenced by the composition of the tissues in the adjacent slices. The radiation dose is similar to that of a conventional 2D mammogram, which can additionally be performed under the same breast compression as the DBT. This dual examination therefore uses double the radiation of a single examination, and dose considerations have led to the development of software enabling reconstruction of a synthetic 2D mammogram from the DBT dataset as a potential replacement for the conventional 2D mammogram.

DBT has been shown to be at least as accurate as focal compression views in the assessment of screen detected abnormalities [1, 2, 3] and is useful in the assessment of symptomatic women [4, 5]. It provides a better assessment of the extent of malignancy than 2D mammography [6] and is increasingly used for image-guided biopsy [7].

DIGITAL BREAST TOMOSYNTHESIS IN PRIMARY SCREENING

Large studies of DBT in population screening such as the Oslo [8], Malmo [9] and STORM [10, 11] studies have shown a significantly higher sensitivity for the combination of 2D mammography and DBT compared to 2D mammography alone. The Oslo study reported a 27% higher cancer detection rate. As there was no increase in the detection of *in situ* disease, this equated to a 40% increase in invasive cancer detection. Similar findings have been seen in observational studies performed following the service introduction of DBT [12, 13].

Routine screening using DBT is now undertaken in an increasing number of centres in the USA and in some European countries. It is commonly performed together with 2D mammography, particularly as microcalcifications can be difficult to judge on DBT. There is now reasonably robust evidence that DBT plus synthetic 2D (S2D) mammography is diagnostically equivalent to DBT plus

conventional 2D mammography [11, 14], and the lower overall dose has led to S2D replacing 2D for screening in some centres.

"...The study showed no significant difference in either the overall recall rates between the groups .... or in the false positive recall rates..."

Composite densities are a common cause of false positive 2D mammograms. The ability of DBT to resolve these should, therefore, result in a reduction in the number of false positive recalls from screening, and indeed some studies have shown this to be the case [15, 16]. Other studies, however, have suggested that the recall rate is not affected by the addition of DBT, or may even increase [9] [11]. We have performed a study focusing on recall rates in perhaps the most challenging group of women undergoing screening, namely the younger, higher risk women.

DIGITAL BREAST TOMOSYNTHESIS IN SCREENING YOUNG HIGHER RISK WOMEN

This study [17] was performed in women aged 40 to 49 attending family history services and referred for, or currently undergoing, annual mammographic screening in two UK breast centres. These women are often particularly anxious because of their increased risk. Furthermore, their breast density is on average higher than that of older women, and this increases the likelihood of composite densities which may result in recall for further assessment.

The study was designed to compare false positive recall rates with 2D mammography alone and 2D plus DBT and is, to our knowledge, the first published prospective randomised controlled trial of DBT in screening. Mammography was performed on Hologic Selenia Dimensions equipment. Eligible and consenting women undergoing incident (second or subsequent round) screening were randomised to undergo either 2D mammography alone or 2D plus DBT. A year later those who remained in the study received the alternate examination, i.e. those who had 2D the first time then had 2D plus DBT, and *vice versa*. All examinations were double read by two radiologists with experience of DBT in screening assessment. Disagreement between the radiological opinions was resolved by consensus discussion, with arbitration by a third radiologist if necessary. Recruitment of women undergoing prevalent (first round) screening also took place but this was halted part way through the study due to a below target recruitment rate.

A total of 1227 women undergoing incident screening were recruited, and of these 1170 had 2D only and 1175 2D plus DBT examinations. The study showed no significant difference in either the overall recall rates between the groups (2D 2.8%; 2D plus DBT 2.7%) or the false positive recall rates (2D 2.4%; 2D plus DBT 2.2%). One interesting finding was that in the first year of the study there were significantly more 2D plus DBT examinations which were scored as abnormal by one or both readers but were judged to be normal at consensus/arbitration (and the women therefore not recalled) than in the second year. This suggests an initial lack of confidence in reading DBT screening examinations (particularly in dismissing benign asymmetric densities) which then improved with experience.

The findings suggest that, at least in our practice, there is little or nothing to be gained by adding DBT in terms of reducing recall rate in these younger women. As recall rates tend to be lower in older women due to their lower breast density, one would expect the same conclusion to hold true for women in the UK's population screening age group (50 to 70 years), although this does require confirmation with further research. Centres with higher recall rates may obtain some reduction from the addition of DBT, although arguably double reading with robust consensus/arbitration (as in our centres) could achieve the same goal. The recall rates in this study were surprisingly low, even for the 2D only examinations. Most of our population screening is carried out using mammography machines from another manufacturer, and it is possible that 2D false positive rates vary according to the machine used.

UNANSWERED QUESTIONS ABOUT DBT IN SCREENING

Although DBT is being increasingly used in population screening, there remain a number of unanswered questions about its role, in particular whether there are defined subgroups of women that would benefit from its use and about its cost effectiveness.

The sensitivity of mammography is inversely related to breast density due to the masking effect of fibroglandular tissue [18]. Although the sensitivity of DBT has been reported to be higher than 2D mammography in breasts of all densities, there is evidence to suggest that women with denser breasts are most likely to benefit [11]. DBT may therefore have a role for screening these women in a stratified screening programme.

There has been almost universal disregard among the authors of the published DBT studies of the prevalent/incident screening status of the women studied. This is of fundamental importance, as women undergoing screening for the first time have a substantially higher false positive recall rate than those being screened for the second or subsequent time (7.9% v. 3.0% in the UK population screening programme) due to the absence of previous mammograms for comparison. DBT may be more useful in reducing false positive recalls in this group than in previously screened women. Furthermore, the workload resulting from screening only prevalent women with DBT (predominantly the longer reading times) would be much easier to manage than

if DBT were used for screening all women. This limited additional workload may be offset to a certain extent by a reduction in false positive recalls for assessment, thus also addressing one of the harms of screening.

"...There has been almost universal disregard among the authors of published DBT studies of the prevalent/incident screening status of the women studied..."

The cancer detection performance for women undergoing repeat screening with DBT remains to be seen. The high detection performance reported in many of the large DBT studies is almost certainly due to some extent to the 'prevalent screen effect', whereby the first use of the higher sensitivity new technology detects additional cancers that wouldn't have been detected with the standard technology. The detection rate for DBT may then drop towards the pre-DBT level for subsequent rounds, although there may be an overall downshift in the size of screen-detected cancers. One would hope to see fewer interval cancers in women screened with DBT, but robust data on interval cancer rates are not yet available.

"... The high detection performance reported in many of the large DBT studies is almost certainly due to some extent to the 'prevalent screen effect'..."

The high sensitivity of DBT for the detection of small spiculate masses (typically low grade invasive cancers) has raised the question of whether screening with DBT increases overdiagnosis. This may be an argument for concentrating DBT usage in younger women, in whom slow growing cancers would otherwise eventually present symptomatically and would therefore not be overdiagnosed.

The multicentre PROSPECTS study (Prospective Randomised Trial of Digital Breast Tomosynthesis (DBT) Plus Standard 2D Digital Mammography (2DDM) or Synthetic 2D Digital Mammography (S 2D) Compared to Standard 2D Digital Mammography in Breast Cancer Screening) will randomise 100,000 women undergoing population screening in the UK and will hopefully provide answers to some of these important questions.

REFERENCES

1. Morel JC, Iqbal A, Wasan RK, Peacock C, Evans DR, Rahim R, *et al.* The accuracy of digital breast tomosynthesis compared with coned compression magnification mammography in the assessment of abnormalities found on mammography. Clin Radiol 2014; 69: 1112–6. doi:10.1016/j.crad.2014.06.005.
2. Cornford EJ, Turnbull AE, James JJ, Tsang R, Akram T, Burrell HC, *et al.* Accuracy of GE digital breast tomosynthesis versus supplementary mammographic views for diagnosis of screen-detected soft-tissue breast lesions. Br J Radiol 2016; 89: 20150735. doi:10.1259/bjr.20150735.
3. Whelehan P, Heywang-Köbrunner SH, Vinnicombe SJ, Hacker A, Jänsch A, Hapca A, *et al.* Clinical performance of Siemens digital breast tomosynthesis versus standard supplementary mammography for the assessment of screen-detected soft-tissue abnormalities:a multi-reader study. Clin Radiol 2017; 72: 95.e9–95.e15. doi:10.1016/j.crad.2016.08.011.
4. Peppard HR, Nicholson BE, Rochman CM, Merchant JK, Mayo RC III, Harvey JA. Digital Breast Tomosynthesis in the Diagnostic Setting: Indications and Clinical Applications. Radiographics 2015; 35: 975–90. doi:10.1148/rg.2015140204.
5. Raghu M, Durand MA, Andrejeva L, Goehler A, Michalski MH, Geisel JL, *et al.* Tomosynthesis in the Diagnostic Setting: Changing Rates of BI-RADS Final Assessment over Time. Radiology 2016; 281: 54–61. doi:10.1148/radiol.2016151999.
6. Mun HS, Kim HH, Shin HJ, Cha JH, Ruppel PL, Oh HY, *et al.* Assessment of extent of breast cancer: Comparison between digital breast tomosynthesis and full. Clin Radiol 2013; 68: 1254–9. doi:10.1016/j.crad.2013.07.006.
7. Waldherr C, Berclaz G, Altermatt HJ, Cerny P, Keller P, Dietz U, *et al.* Tomosynthesis-guided vacuum-assisted breast biopsy: A feasibility study. Eur Radiol 2016: 26: 1582 doi: 10.1007/s00330-015-4009-4
8. Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U, et al. Comparison of Digital Mammography Alone and Digital Mammography Plus Tomosynthesis in a Population-based Screening Program. Radiology 2013;267:47–56. doi:10.1148/radiol.12121373.
9. §Lång K, Andersson I, Rosso A, Tingberg A, Timberg P, Zackrisson S. 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. Eur Radiol 2016; 26: 184–90. doi:10.1007/s00330-015-3803-3.
10. Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S, *et al.* Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. Lancet Oncol 2013; 14: 583–9. doi:10.1016/S1470-2045(13)70134-7.
11. Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fantò C, Ostillio L, *et al.* 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. Lancet Oncol 2016; 17: 1105–13. doi:10.1016/S1470-2045(16)30101-2.
12. Durand MA, Haas BM, Yao X, Geisel JL, Raghu M, Hooley RJ, *et al.* Early clinical experience with digital breast tomosynthesis for screening mammography. Radiology 2015; 274: 85–92. doi:10.1148/radiol.14131319.
13. Sharpe RE Jr, Venkataraman S, Phillips J, Dialani V, Fein-Zachary VJ, Prakash S, *et al.* Increased Cancer Detection Rate and Variations in the Recall Rate Resulting from Implementation of 3D Digital Breast Tomosynthesis into a Population-based Screening Program. Radiology 2016;278:698–706. doi:10.1148/radiol.2015142036.
14. Skaane P, Bandos AI, Eben EB, Jebsen IN, Krager M, Haakenaasen U, *et al.* Two-View Digital Breast Tomosynthesis Screening with Synthetically Reconstructed Projection Images: Comparison with Digital Breast Tomosynthesis with Full-Field Digital Mammographic Images. Radiology 2014; 271: 655–63. doi:10.1148/radiol.13131391.
15. Destounis SV, Arieno A, Morgan R. Initial Experience with Combination Digital Breast Tomosynthesis Plus Full Field Digital Mammography or Full Field Digital Mammography Alone in the Screening Environment. J Clin Imaging Sci 2014;4:9. doi:10.4103/2156-7514.127838.
16. Lourenco AP, Barry-Brooks M, Baird GL, Tuttle A, Mainiero MB. Changes in Recall Type and Patient Treatment Following Implementation of Screening Digital Breast Tomosynthesis. Radiology 2015; 274: 337–42. doi:10.1148/radiol.14140317.
17. Maxwell AJ, Michell M, Lim YY, Astley SM, Wilson M, Hurley E, *et al.* A randomised trial of screening with digital breast tomosynthesis plus conventional digital 2D mammography versus 2D mammography alone in younger higher risk women. Eur J Radiol 2017. doi:10.1016/j.ejrad.2017.06.018.
18. Wanders JOP, Holland K, Veldhuis WB, Mann RM, Pijnappel RM, Peeters PHM, *et al.* Volumetric breast density affects performance of digital screening mammography. Breast Cancer Res Treat 2016; 162: 95–103. doi:10.1007/s10549-016-4090-7.