Original paper

The threshold detectable mass diameter for 2D-mammography and digital breast tomosynthesis

Andria Hadjipanteli[a,b,*], Premkumar Elangovan[a], Alistair Mackenzie[a], Kevin Wells[c], David R. Dance[a,d], Kenneth C. Young[a,d]

[a] National Coordinating Centre for the Physics of Mammography, Royal Surrey County Hospital, Guildford, Surrey, UK
[b] Medical School, University of Cyprus, Nicosia, Cyprus
[c] Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, UK
[d] Department of Physics, University of Surrey, Guildford, UK

ABSTRACT

Digital breast tomosynthesis (DBT) is currently under consideration for replacement of, or combined use with 2D-mammography in national breast screening programmes. To investigate the potential benefits that DBT can bring to screening, the threshold detectable lesion diameters were measured for different forms of DBT in comparison to 2D-mammography. The aim of this study was to compare the threshold detectable mass diameters obtained with narrow angle (15°/15 projections) and wide angle (50°/25 projections) DBT in comparison to 2D-mammography. Simulated images of 60 mm thick compressed breasts were produced with and without masses using a set of validated image modelling tools for 2D-mammography and DBT. Image processing and re-construction were performed using commercial software. A series of 4-alternative forced choice (4AFC) experiments was conducted for signal detection with the masses as targets. The threshold detectable mass diameter was found for each imaging modality with a mean glandular dose of 2.5 mGy. The resulting values of the threshold diameter for 2D-mammography (10.2 ± 1.4 mm) were found to be larger (p < 0.001) than those for narrow angle DBT (6.0 ± 1.1 mm) and wide angle DBT (5.6 ± 1.2 mm). There was no significant difference between the threshold diameters for wide and narrow angle DBT. Implications for the introduction of DBT alone or in combination with 2D-mammography in breast cancer screening are discussed.

1. Introduction and background

Digital breast tomosynthesis (DBT) involves the acquisition of two-dimensional X-ray projections of the breast over a limited angular range and their reconstruction to image planes parallel to the detector [1–5]. It exposes the patient to similar dose levels to those of 2D-mammography [6] and is currently under consideration for its use in breast cancer screening in combination with 2D-mammography or alone in several countries in Europe. For DBT to be combined with 2D-mammography in breast screening it would require the additional dose due to DBT to be justified in terms of overall reduction in mortality and morbidity. For DBT to replace 2D-mammography in screening it would have to at least provide the same detectability of cancer lesions as 2D-mammography.

Some clinical observer studies have shown that with DBT, and especially for masses, there is an increase in sensitivity, and a significant reduction in specificity [7–10], while others have shown no change in sensitivity [11–13] and no significant increase in specificity [12,14]. Clinical observer studies have the advantage of high clinical relevance, however, there are conflicting conclusions in the literature. Furthermore, clinical studies can be time consuming, potentially subject to a number of confounding factors, expensive and do not always allow the evaluation of specific imaging technology metrics (for example, detectable lesion diameter), due to, for example, the limited number of sufficiently small lesions. The comparison of different imaging protocols using the same cohort of women can also be challenging, due to the ethical issues that exist in exposing one woman to different image acquisition methods.

Alternatively, simulation methods are sensitive and useful for investigating the imaging technology. Observer studies with lesions inserted into physical phantoms (anthropomorphic or structured) [15,16] or fully simulated breast images [17–20] have yielded useful results, while image metrics such as contrast-to-noise ratio (CNR), CNR/ASF (artefact spread function) [21–23] have also been used successfully.
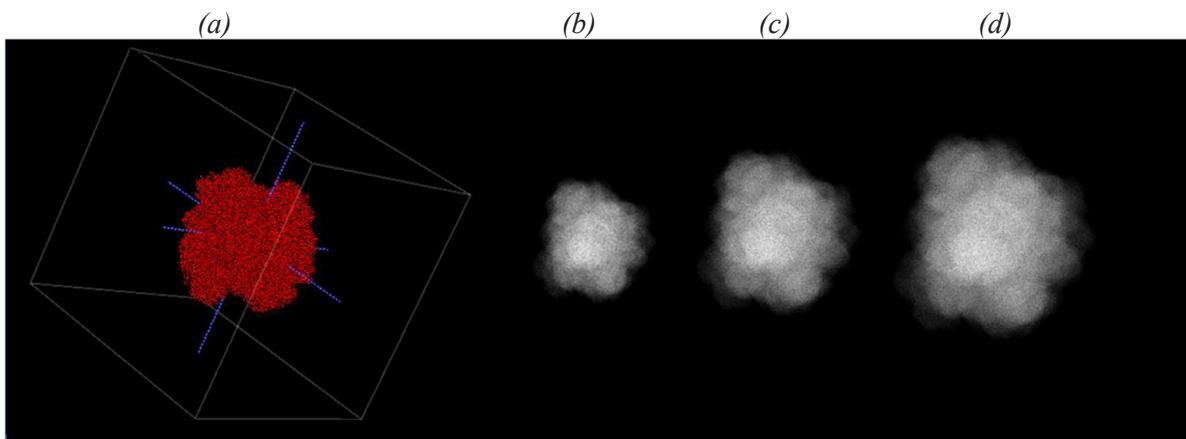
Fig. 1. Simulated mass: a) 3D rendering of mass; 2D projection images of isolated masses of average diameters (a) 6.6 mm, (b) 8.4 mm and (c) 10.3 mm before insertion into the mathematical breast phantom.

Simulation studies have shown that masses were visible at a lower contrast than with 2D-mammography in comparison to DBT [17,20]. In the recent work of Hadjipanteli *et al* [18] three commercially-implemented geometries (2D-mammography, narrow angle (15°/15 projections) and wide angle (50°/25 projections)) were compared. It was shown that 2D-mammography demonstrated a smaller minimum detectable calcification diameter than either DBT geometry, and narrow angle DBT showed a smaller minimum detectable calcification diameter than wide angle DBT. This might be explained by the higher resolution (higher MTF) that the 2D-mammography system exhibited compared to that for the DBT systems investigated [24], which was advantageous for the detection of small objects such as micro-calcifications. For the detection of larger objects DBT has the advantage of reducing the appearance of overlying structures in the image [4] but the benefit of this in terms of minimum detectable mass size against a clinically realistic background has not been quantified in an observer study.

The aim of this work therefore was to use a simulation-based observer study to quantify the threshold detectable mass diameter for 2D-mammography, narrow and wide angle DBT. For this purpose we simulated the images of a voxel phantom [25] with inserted masses to create realistic simulated images. An advantage of this approach is that the same breast can be imaged using different acquisition methods, unlike clinical studies. We aimed to expand on the results of a previous observer study [18] conducted for the detection of calcifications and used the same methods to assess the same systems for the detection of masses. Knowledge of the threshold detectable mass diameter, in addition to calcification diameter, of the three imaging systems would help quantify the differences between DBT (narrow and wide angle) and 2D-mammography imaging and make a clinically relevant comparison between different imaging systems for different forms of breast cancer.

2. Methods and materials

In this study, validated simulation methods [26] were used to assess the detectability of masses in 2D-mammography and DBT. These methods involve the realistic simulation of breast images with and without masses and their use in 4-alternative forced choice (4-AFC) observer studies. The values of the threshold detectable mass diameter determined from the observer studies were compared for each modality. Dose, breast glandularity, breast thickness, mass insertion height and the processing or reconstruction software used (which can all affect lesion detectability) were fixed for the three modalities tested, so that only the effect of system geometry on mass detectability was assessed.

The simulation involved three stages: creation of voxel phantoms of the breast, creation and insertion of simulated masses into the phantoms, and the creation of images. These stages, together with the 4-AFC methodology and analysis, are described in Sections 2.1–2.5 below.

2.1. Mathematical breast phantom

For consistency with the methods followed by Hadjipanteli *et al* [18], we used six realistic mathematical breast phantoms validated for 4-AFC studies [25]. The breast phantoms were produced using a biological approach in which breast components were simulated using features extracted from clinical DBT images. The components simulated were the skin, adipose tissue, fibro-glandular tissue, Cooper's ligaments and blood vessels. Clinical breast readers were unable to distinguish simulated image patches produced using these phantoms from equivalent image patches from real mammograms and DBT images. Furthermore, the statistical properties of simulated images showed a strong match to those of real images [25].

The phantoms had a voxel size of $100\,\mu m \times 100\,\mu m \times 100\,\mu m$. The phantom shape was based on a real compressed breast and its 2D-size was 240 mm (chest wall direction) × 120 mm (nipple direction). Each breast phantom had a compressed breast thickness (CBT) of 60 mm and a glandularity of between 17% and 19% by volume in each patch. This volume glandularity was chosen as it is equivalent to the average glandularity of 21% by mass in the central portion of the breast for women of age 50 to 64 with CBT of 60 mm attending the UK breast screening programme [27].

2.2. Simulated masses

Simulated masses were produced using a fractal growth method known as diffusion limited aggregation [28]. Fig. 1 shows the 3D rendered mass and the 2D projections of the mass for different diameters. Eleven different masses of the same density were replicated three times each by rotation through 90°, 180° and 270°, creating 44 unique masses of a selected diameter. The attenuation properties of the masses were assumed to be the same as glandular tissue. The masses were then scaled and set to average diameters ranging from 4.7 to 10.3 mm and inserted into the breast phantoms by voxel replacement at a random location at least 10 mm from the skin edge and at a constant height, 30 mm, above the breast support.

2.3. Image simulation

The modelling tools used to calculate simulated images of the breast phantoms included a ray tracing tool, a scatter prediction tool and an image degradation tool [26]. These are briefly described below. More
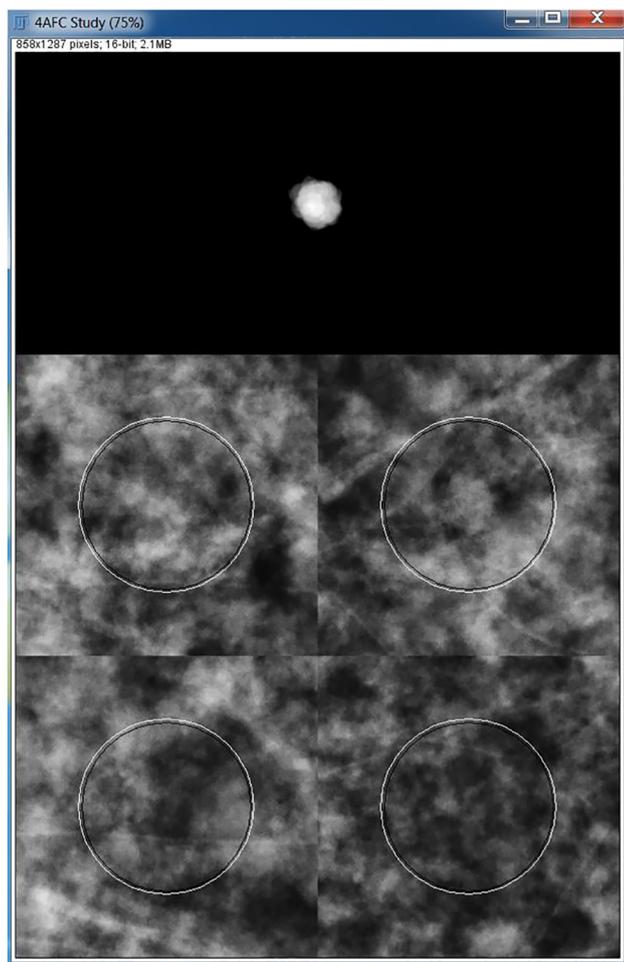
Fig. 2. Screenshot of software used in 4AFC study. The reference copy is shown above the four quadrants of which one contains a mass. The mass is in the top-right hand corner.

details can be found in Hadjipanteli et al [18].

The narrow angle DBT geometry tested used a 15°/15 projections configuration, based on the DBT geometry of the Hologic Selenia Dimensions X-ray set. The wide angle DBT geometry tested used a wide-angle and high number of projections (50°/25 projections) configuration, also based on an existing commercial geometry (Mammomat Inspiration, Siemens Medical Solutions, Erlangen, Germany). The number of projections and source movement blurring matched the commercial system. In this way, an attempt was made to compare the clinical geometry of Hologic (narrow angle DBT) to a possible wider-angular range geometry, which is realistic and is based on an existing geometry. The image receptor was represented as an amorphous selenium detector for both 2D-mammography and the two DBT geometries being simulated (Selenia Dimensions, Hologic Inc., Bedford, MA). Thus the same detector and X-ray spectrum were used for both DBT configurations. The physical pixel pitch of the detector was 0.07 mm for 2D-mammography and 0.14 mm for narrow and wide angle DBT following pixel binning.

Clinically used tube voltage and target/filter materials were simulated for the X-ray source: (i) 2D-mammography: 31 kVp W/Rh; (ii) DBT: 33 kVp W/Al. Additional aluminium attenuation was used to attenuate the spectra to match the experimentally measured half value layers of 0.55 mm Al for 2D-mammography and 0.59 mm Al for narrow and wide angle DBT.

For each spectrum, the incident air kerma was calculated at the entrance of the breast and the mean glandular dose (MGD) was computed using conversion factors from Dance et al [27,29], based on the standard methods adopted in UK and EU breast dosimetry protocols in mammography [30,31] and DBT [32]. MGD was fixed at 2.5 mGy for all three modalities [6].

The primary images/projections were produced using a ray tracing tool developed for 2D-mammography and DBT [26]. The ray tracing simulation included transmission through the anti-scatter grid (in 2D-mammography only), geometric blurring due to focal spot finite size ($0.4 \times 0.4\,mm^2$). Blurring due to tube movement was taken into account by stretching the focal spot in the direction of movement, such that the focal spot size was $0.4 \times 1.4\,mm^2$ for the narrow angle DBT and $0.4 \times 2.2\,mm^2$ for wide angle DBT in the direction of tube movement.) Focal spot size values were based on experimental measurements of the exposure time and tube movement. The attenuation by the breast support and compression paddle was also taken into account. Breast movement was not included.

A pre-calculated table of the scatter-to-primary ratios (SPR) derived from Monte Carlo measurements of scatter was used for the calculation and incorporation of scatter in the images. This had been previously validated as part of an image simulation framework [26].

The image degradation tool was based on the methods of Mackenzie et al [33,34]. These use measurements of signal transfer properties (STP), pre-sampled modulation transfer function (MTF), noise power spectrum (NPS) and flat field correction map to adapt the image quality of the simulated images.

It was not possible to use vendor specific reconstruction software for both the narrow and wide angle DBT. Therefore, the images were processed and reconstructed using generic software from Real Time Tomography, LLC, (Philadelphia, USA), due to the flexibility in processing different DBT geometry configurations. The 2D images were processed using *Adara* and the DBT planes were reconstructed using *Briona*, which used filtered back-projection for reconstruction of the tomosynthesis image planes. After post processing, the 2D-mammography images and DBT planes were cropped into 30 mm × 30 mm image patches, randomly selected away from the edges of the whole breast image, to be used in the human observer experiments. For DBT, 12 planes were used, centred on the centre of the mass volume. A pilot study was undertaken to aid the selection of three mass diameters to be inserted into the images for each modality to give detection rates between 30% and 95% for each. To achieve the required detection rates, the mass diameters used were: 2D-mammography: 8.4 mm, 9.4 mm and 10.3 mm; and narrow and wide angle DBT: 4.7 mm, 5.6 mm and 6.6 mm. In total, 396 cropped images (44 images × 3 diameters × 3 modalities) each with a mass were produced. In addition, 1768 background cropped images were produced without an inserted mass for each modality.

2.4. Observation and 4AFC study

Five physicists participated in a series of 4AFC human observer experiments. As this was a forced choice study and does not include the effect of 'searching' or interpretation, non-radiologists as observers were acceptable. Elangovan et al [20] found that although radiologists had a slightly better detection rate in 4-AFC studies than non-radiologists, there were no differences in the overall conclusions about the effect of differences in technology. An in-house Java-based plug-in (ImageJ 1.50i, NIH, USA) was used to run the 4AFC study (Fig. 2). Sets of four 30 mm × 30 mm breast phantom 2D-mammography cropped images or DBT planes were randomly selected and presented in turn to each observer. In each set, one image contained a mass in the centre and the other three did not. The observers were required to identify which of the four images contained the mass and register their decision by selecting the relevant quadrant. A 2D projection of the inserted 3D mass (reference copy of the signal or "image cue") was also shown isolated from the background. In the DBT 4AFC studies the observer was able to scroll through 12 planes. Each observer was shown 132 groups of four images for each modality. Within these 132 groups, 44

cases were presented for each of the three different mass diameters to the observer in a random order.

As in clinical practice, a high-resolution reporting quality monitor (Barco, B-8500, 5MP, Belgium) was used for the experiments. All images were displayed at 100% magnification (one-to-one pixel mapping between the image and the monitor), low lighting levels were used and no time limit was imposed.

2.5. Analysis

A linear relationship was assumed between the mass diameter and detectability, based on the Rose Model [35,36]. The detectability was expressed numerically in terms of the detectability index (d'), a quantity related to the visibility of a lesion. The d' can be found in look-up tables in Macmillan and Creelman [37] from the percentage of correct detection decisions (PC) for an observer in a 4-AFC study. The threshold detectable mass diameter was taken as the size at which the observer makes 90.7% correct decisions, and which corresponds to a d' of 2.5 [37]. For each observer a linear least square fit for d' versus mass diameter was used to find the threshold detectable mass diameter (at a d' of 2.5) and the overall mean threshold detectable mass diameter for each modality was calculated. For comparison, the threshold detectable mass diameter for each modality at a PC of 62.5% (corresponds to a d' of 1.2), the value used for reading the CDMAM phantom [38], was also calculated. If the measured detection rate was above 62.5% for all mass diameters, the calculated threshold diameter required an extrapolation of the fitted straight line between d' and mass diameter.

Errors in PC were calculated from the 95% confidence intervals on the measured PC values assuming a binomial distribution and N = 44 for each observer. This provided a range of uncertainty in PC for each observer, which was translated (through the relation of PC to d') to a range of uncertainty in d' and the threshold detectable mass diameter value for each observer (through the linear least square fit). The error in the overall mean threshold detectable mass diameter for each modality was set equal to the root mean square error of the uncertainty in threshold detectable mass diameters of the five observers, over the square root of five. Analysis of variance (ANOVA) was performed to identify any statistically significant differences between the threshold detectable mass diameters for the three imaging modalities and to test for differences between the observers.

3. Results

Fig. 3 shows examples of the simulated images for (a) 2D-mammography, (b) narrow and (c) wide angle DBT, used in the observer studies.

Fig. 4 shows the dependence of the detectability index, d' on the mass diameter for wide angle DBT for each observer. For the least squares fits shown in the figure, the goodness of fit, in terms of the square of the Pearson product moment correlation coefficient, R^2 was found to be between 0.80 and 0.99 for all modalities and observers.

The threshold detectable mass diameters for the five observers and the three imaging modalities, at a MGD of 2.5 mGy and with a height of insertion of 30 mm above the breast support, are presented in Fig. 4 (a). ANOVA analysis produced a *p*-value of 0.65 for the effect of the variation between the observers, showing confidence in the rejection of the hypothesis that there is a significant difference in the threshold detectable mass diameters estimated for the five observers. The difference between the three modalities for each observer is shown in Fig. 5(a). Fig. 5(b) shows the averages of the observers' results for each modality at PC values of 62.5% and 90.7%. It can be seen that DBT performs better than 2D-mammography and wide angle DBT performs marginally better than narrow angle DBT (for both a PC of 90.7% and 62.5%). There was a highly significant difference between the threshold mass diameter that can be detected by 2D-mammography and DBT ($p < 0.0001$). When comparing the results at a PC of 90.7% to those for a PC of 62.5% a difference of 2.1–2.5 mm was found depending on modality.

4. Discussion

In this study, the threshold detectable mass diameter for three different existing geometries of breast imaging systems have been quantified through 4AFC observer studies for the same 60 mm thick breasts. The comparison methods used offered the key advantage of allowing precise control over the variables under investigation, which would be impossible under a conventional clinical trial. We used human observers viewing images of the same breast tissue for the comparison of different systems, thereby eliminating variability in subject background and target appearance, over and above the particular variable under consideration: threshold diameter. Moreover such a simulation-based approach is far less time and resource hungry than clinical trials. However, the 4AFC methodology does limit the number of acquisition methods that can be practicably studied. Other authors [21,23] used
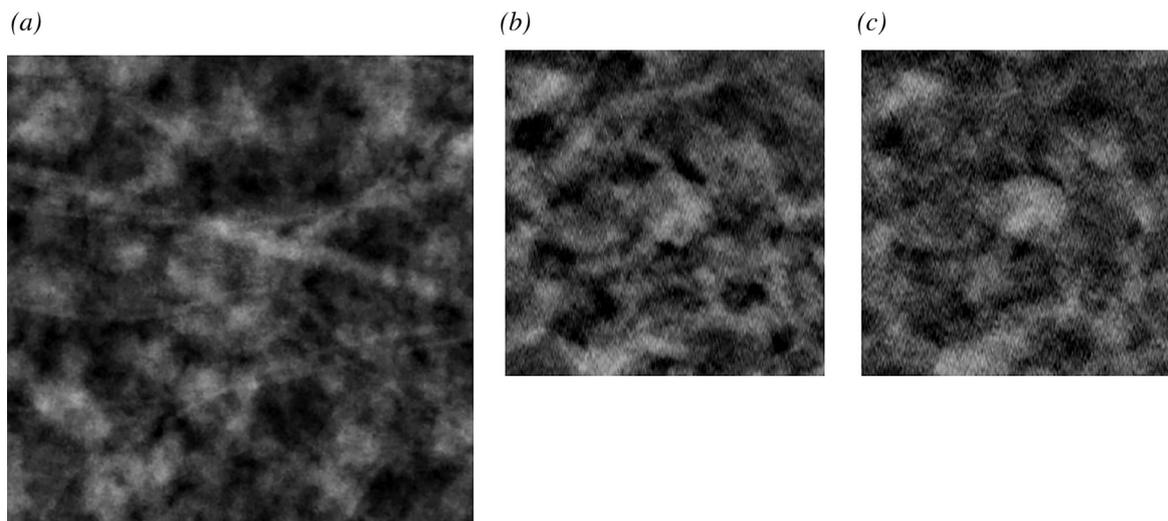
(a) *(b)* *(c)*

Fig. 3. Example images of masses inserted into mathematical breast phantoms and following the image simulation of (a) 2D-mammography, (b) narrow and (c) wide-angle DBT. The sizes of the masses are (a) 6.6 mm (b) 10.3 mm (c) 10.3 mm.
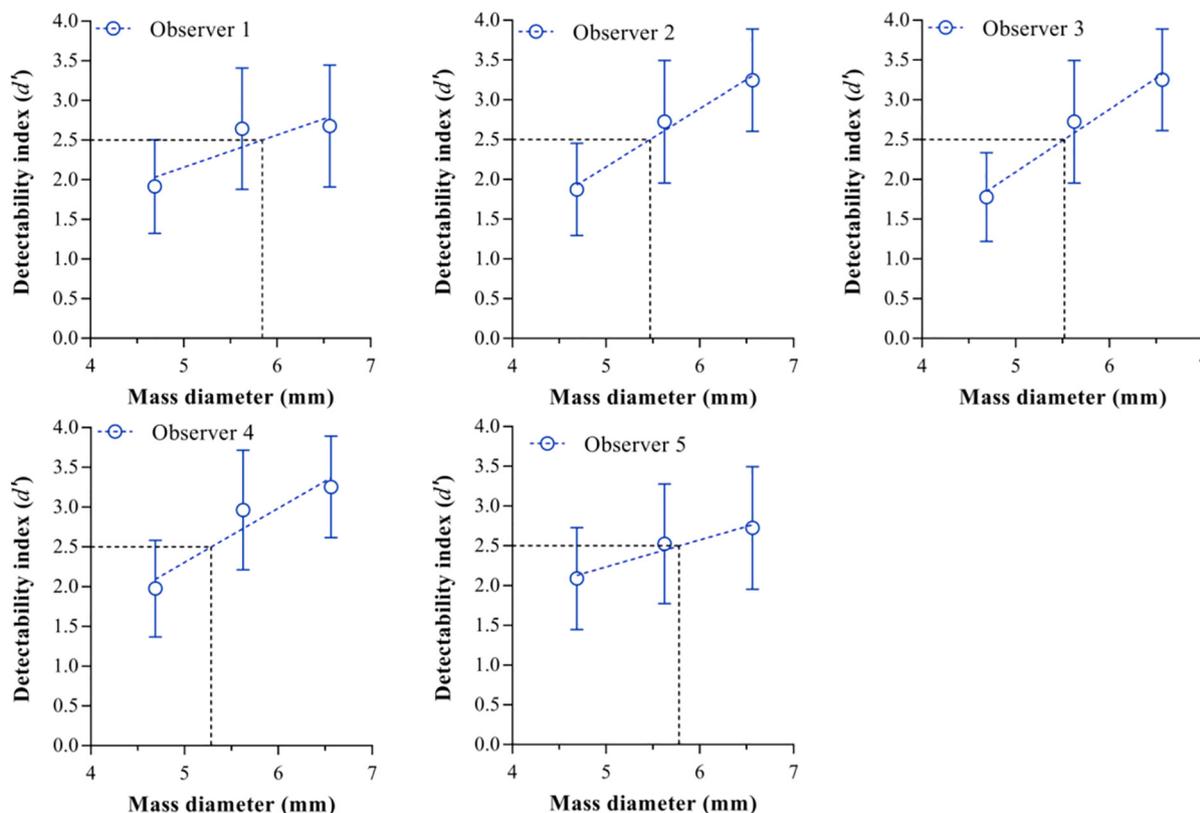
Fig. 4. The dependence of the detectability index, d', on mass diameter for wide angle DBT for the five observers. The vertical lines show the mass diameters at a detectability index of 2.5, which was taken as the threshold detectable mass diameter. Vertical dotted lines show the corresponding mass diameter for each observer where d' = 2.5. Error bars were calculated from the 95% confidence intervals on the measured PC values assuming a binomial distribution and N = 44 for each observer and translated to d' values.

image metrics (e.g. contrast to noise ratio (CNR), CNR/ASF (artefact spread function)) instead of observer experiments to assess the visibility of calcifications and tested more geometries. However, such an approach lacks the link to the detection performance of a real observer.

DBT was found to have a lower threshold detectable mass diameter than 2D-mammography (p-value of < 0.0001), whichever choice of PC was made; 90.7% or 62.5%, although the threshold values are higher for the higher PC choice. The order of performance of the systems agrees with a previous 4AFC observer study [20], carried out using methods similar to those used in the current study, but which used Hologic reconstruction software instead of Briona reconstruction software, and which evaluated the signal contrast needed to correctly identify a mass, instead of the threshold detectable mass diameter. Elangovan et al [20] showed that observers needed over three times the signal contrast to correctly identify a mass in 2D-mammography compared with DBT (narrow angle). The work by Mackenzie et al [19] used similar methods to quantify the minimum detectable mass diameter for 2D-mammography and DBT for the Siemens Inspiration system (which has the wide-angle geometry used in this study). Their conclusions agree with our work in that DBT has a smaller detectable mass diameter than 2D-mammography. However, numerically, there are differences in the results (2D-mammography 6.3 mm, wide-angle DBT: 4.9 mm, for a PC of 90.7%), from these latest results, which might be partly explained by the using a thinner CBT of 53 mm and that the simulated system was based entirely on the Siemens Inspiration, including the manufacturer's processing and reconstruction software and pixel size. In our study, Briona processing and reconstruction software were used, instead of the manufacturer's software used by Mackenzie et al.

Our results clearly confirm the advantage of DBT over 2D-mammography for mass detection. However, they also raise a concern for either DBT or 2D-mammography being used alone in breast screening

as each better visualises a different form of breast cancer. 2D-mammography appears to offer higher detectability of small calcifications [18] and DBT appears to offer superior performance for the detection of small masses. Even though the DBT images may be less sharp than 2D-mammography images, as the modulation transfer function (MTF) for some systems is lower [24], DBT still favours the detectability of masses, due to the ability to remove or reduce the appearance of overlapping structures in image planes of the breast tissue being imaged.

The results of this study suggest that if 2D-mammography were used alone in screening, it could miss masses that would be visible with DBT. However, by also considering the results in Hadjipanteli et al [18], if DBT alone is used it could miss small calcifications, which would be visible with 2D-mammography. 2D-mammography could be used alone at an increased dose, but there is evidence to support that changing dose does not significantly change mass detection levels [39]. Alternatively, if DBT alone is used at double the standard dose level (5 mGy), so that an increased calcification detection is provided, the threshold detectable calcification diameter would still not be as small as that that can be provided by 2D-mammography [18]. The choice of the imaging methods seems more important than dose. On the other hand, the combined use of both 2D-mammography and DBT in screening has the advantage of adding the newest technology available for a further benefit in early diagnosis, but at a cost of increased radiation dose to a largely healthy screening population. A potential solution may lie in the use of synthetic imaging, which provides the generation of a synthesised 2D image from DBT data (and thus eliminating the use of 2D-mammography), and is available from some manufacturers. There are on-going studies on the use of synthetic imaging in screening [40–43] and the identification of its strengths and weaknesses [44].

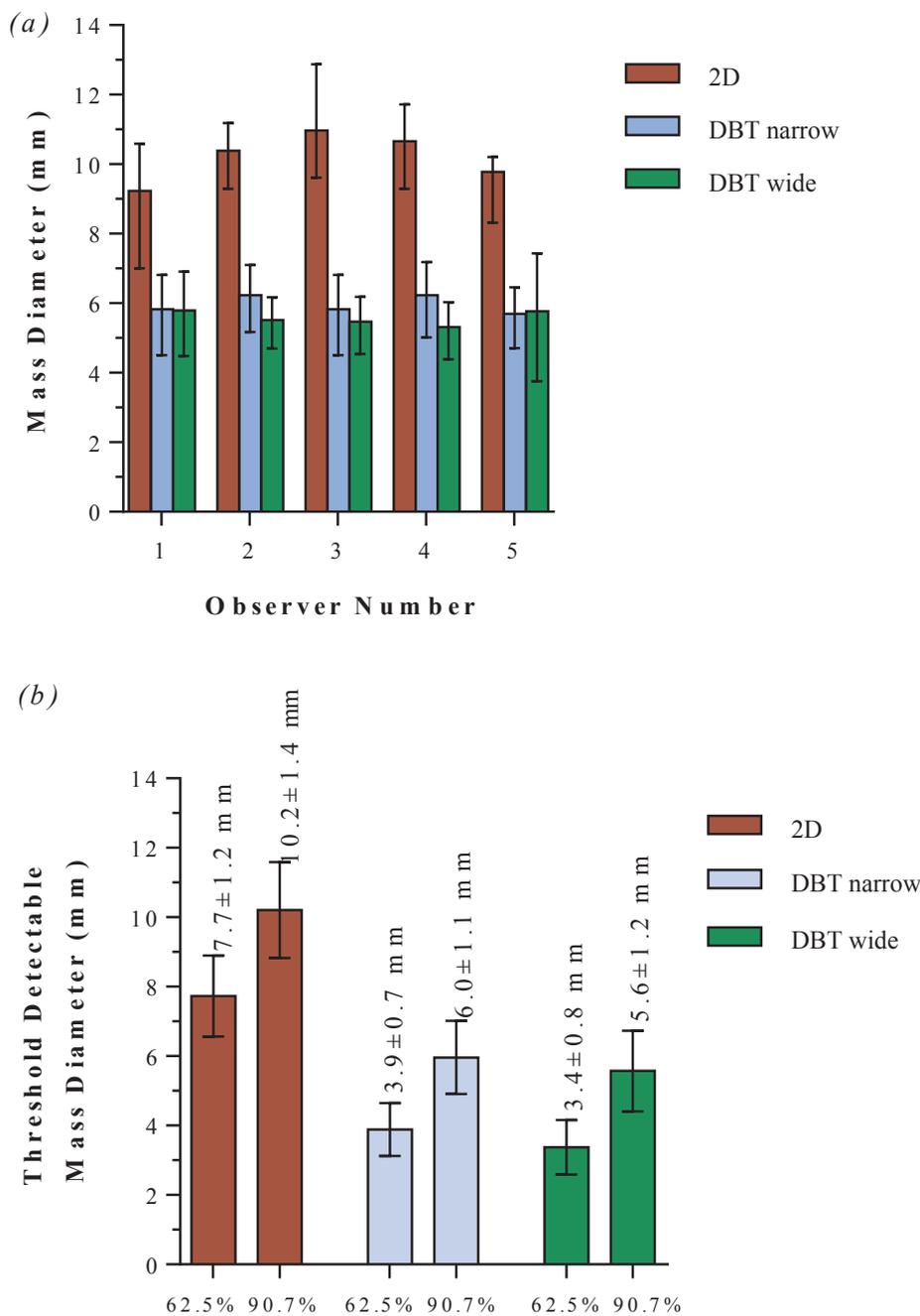This study also provided a comparison between an existing narrow

(a)

(b)

Fig. 5. (a) The threshold detectable mass diameter for each observer for 2D-mammography, narrow and wide angle DBT, for an MGD of 2.5 mGy and mass insertion 30 mm above the breast support and PC = 90.7%. Errors in PC were calculated from the 95% confidence intervals on the measured PC values assuming a binomial distribution and N = 44 for each observer. This provided a range of uncertainty in PC for each observer, which was translated (through the relation of PC to d') to a range of uncertainty in d' and the threshold detectable mass diameter value for each observer (through the linear least square fit). (b) The average of the observers results in (a) for each imaging modality and with PC values of 62.5% and 90.7%. The error bars represent the error in the overall mean d' for each modality, calculated as the root mean square error of the uncertainty in threshold detectable mass diameters of the five observers.

angle DBT and a wide angle DBT system when the same image receptor and X-ray spectrum are used. Wide-angle DBT was found to have a marginally smaller (but not significantly different) threshold mass diameter than narrow-angle DBT. Goodsitt *et al* [22] found that observers preferred images acquired with the wide angle (60°, 21 projections) as opposed to narrow angle geometries (16°, 17 projections) for low contrast objects such as masses. Similarly, Sechopoulos and Ghetti [21] and Tucker *et al* [45] found best performance for mass-like objects at the widest angle in their studies. Even though the wide angle DBT geometry might introduce an increased geometric blurring as the projection angle becomes wider [46], the dose per projection is decreased and the relative quantum and electronic noise in each projection is increased, it was still advantageous over narrow angle DBT. By increasing the DBT scan angular range, depth resolution improves [47]. Wider scan angles with more projections (for an adequate sampling of image data and fewer tomosynthesis reconstruction artefacts), increased relative quantum noise and insufficient angular sampling can

have a negative effect on the detectability of small-scale signals [48]. However, for the detectability of large-scale objects like masses, this effect might not be important.

As with all virtual clinical trials to date, limitations of this study include omitting any breast movement in the simulations of the three imaging modalities. Also, the image simulation did not include any lag or ghosting effects for the DBT images. It is not expected that lag and ghosting would significantly affect the conclusions of this study, as lag has negligible effect on image noise, and it is not expected that the effect it has on MTF [24] will affect the ability to resolve objects with the size of masses we used in this study. The simulation could be improved by reducing the size of the voxels and increasing the number of simulated rays per pixel. Similarly, Monte Carlo simulation could have been used to calculate scatter in the breast phantoms used, instead of assuming a constant scatter contribution across the whole image patch. For reasons of practicality we used a simple approach to estimate scatter as in all cases the patches were small and positioned away from

the edges of the whole image. Any systematic error that might have been introduced in the total signal of the image is expected to be similar for both the DBT geometries considered, but could lead to a small difference in the performance comparison between 2D and DBT. Another potential limitation is that we used Briona for the image and reconstruction software rather than the system manufacturer's software. Therefore it could be claimed that it does not fully represent clinically applicable processing. However, quantitative analysis (full width at half maximum, contrast-to-noise ratio (CNR), contrast degradation factor) and qualitative tests on DBT images that included objects reconstructed using Briona showed that it produced images comparable to those produced by reconstruction and processing software used clinically (Hologic Selenia Dimensions).

5. Conclusions

This study used a quantitative approach with observers viewing simulated masses imaged in clinically realistic backgrounds to compare different DBT geometries and 2D imaging. The methodology employed a set of validated tools to ensure clinically relevant conclusions. The results showed no significant difference between the performance of narrow and wide angle DBT for the detection of masses and that both DBT acquisition methods had a smaller threshold detectable mass diameter than 2D-mammography. This raises concerns on the use of 2D-mammography alone in screening, while the addition of DBT could add some benefit, but also, a risk of higher dose effects in the screening population. Further studies are required for the assessment of the overall benefit of using DBT in breast screening, either alone or in conjunction with 2D imaging.

Acknowledgements

References

[1] Niklason LT, Christian BT, Niklason LE, Kopans DB, Castleberry DE, Opsahl-Ong BH, et al. Digital tomosynthesis in breast imaging. Radiology 1997;205:399–406. https://doi.org/10.1148/radiology.205.2.9356620.

[2] Dobbins JT, Godfrey DJ. Digital x-ray tomosynthesis: current state of the art and clinical potential. Phys Med Biol 2003;48:R65–106. https://doi.org/10.1088/0031-9155/48/19/R01.

[3] Dobbins JT. Tomosynthesis imaging: At a translational crossroads. Med Phys 2009;36:1956–67. https://doi.org/10.1118/1.3120285.

[4] Sechopoulos I. A review of breast tomosynthesis. Part I. The image acquisition process. Med Phys 2013;40:014301-1–11. doi:10.1118/1.4770279.

[5] Sechopoulos I. A review of breast tomosynthesis. Part II. Image reconstruction, processing and analysis, and advanced applications. Med Phys 2013;40:014302-1–17. doi:10.1118/1.4770281.

[6] Bouwman RW, van Engen RE, Young KC, den Heeten GJ, Broeders MJM, Schopphoven S, et al. Average glandular dose in digital mammography and digital breast tomosynthesis: comparison of phantom and patient data. Phys Med Biol 2015;60:7893–907. https://doi.org/10.1088/0031-9155/60/20/7893.

[7] Wallis MG, Moa E, Zanca F, Leifland K, Danielsson M. Two-view and single-view tomosynthesis versus full-field digital mammography: high-resolution X-ray imaging observer study. Radiology 2012;262:788–96. https://doi.org/10.1148/radiol.11103514.

[8] Svahn TM, Chakraborty DP, Ikeda D, Zackrisson S, Do Y, Mattsson S, et al. Breast tomosynthesis and digital mammography: a comparison of diagnostic accuracy. Br J Radiol 2012;85:e1074–82. https://doi.org/10.1259/bjr/53282892.

[9] Rafferty EA, Park JM, Philpotts LE, Poplack SP, Sumkin JH, Halpern EF, 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;266:104–13. https://doi.org/10.1148/radiol.12120674.

[10] Gilbert FJ, Tucker L, Gillan MGC, Willsher P, Cooke J, Duncan KA, et al. Accuracy of digital breast tomosynthesis for depicting breast cancer subgroups in a UK retrospective reading study (TOMMY Trial). Radiology 2015;277:697–706.

[11] Gur D, Abrams GS, Chough DM, Ganott MA, Hakim CM, Perrin RL, et al. Digital Breast Tomosynthesis: Observer Performance Study. Am J Roentgenol 2009;193:586–91. https://doi.org/10.2214/AJR.08.2031.

[12] Gennaro G, Toledano A, di Maggio C, Baldan E, Bezzon E, La Grassa M, et al. Digital breast tomosynthesis versus digital mammography: a clinical performance study. Eur Radiol 2010;20:1545–53. https://doi.org/10.1007/s00330-009-1699-5.

[13] Teertstra HJ, Loo CE, van den Bosch MAAJ, van Tinteren H, Rutgers EJT, Muller SH, et al. Breast tomosynthesis in clinical practice: initial results. Eur Radiol 2010;20:16–24. https://doi.org/10.1007/s00330-009-1523-2.

[14] Zuley ML, Bandos AI, Abrams GS, Cohen C, Hakim CM, Sumkin JH, et al. Time to diagnosis and performance levels during repeat interpretations of digital breast tomosynthesis: preliminary observations. Acad Radiol 2010;17:450–5. https://doi.org/10.1016/J.ACRA.2009.11.011.

[15] Ikejimba LC, Salad J, Makeev A, Graff C, Ghammraoui B, Glick SJ. Anthropomorphic breast phantoms for evaluation of FFDM/DBT and breast CT using inkjet printing. In: Chen G-H, Lo JY, Gilat Schmidt T, editors. Proc. SPIE Med. Imaging, vol. 10573, SPIE; 2018, p. 7. doi:10.1117/12.2294534.

[16] Cockmartin L, Marshall NW, Zhang G, Lemmens K, Shaheen E, Van Ongeval C, et al. Design and application of a structured phantom for detection performance comparison between breast tomosynthesis and digital mammography. Phys Med Biol 2017;62:758–80. https://doi.org/10.1088/1361-6560/aa5407.

[17] Timberg P, Båth M, Andersson I, Mattsson S, Tingberg A, Ruschin M. Visibility of microcalcification clusters and masses in breast tomosynthesis image volumes and digital mammography: a 4AFC human observer study. Med Phys 2012;39:2431–7. https://doi.org/10.1118/1.3694105.

[18] Hadjipanteli A, Elangovan P, Mackenzie A, Looney PT, Wells K, Dance DR, et al. The effect of system geometry and dose on the threshold detectable calcification diameter in 2D-mammography and digital breast tomosynthesis. Phys Med Biol 2017;62:858–77. https://doi.org/10.1088/1361-6560/aa4f6e.

[19] Mackenzie A, Kaur S, Elangovan P, Dance DR, Young KC. Comparison of synthetic 2D images with planar and tomosynthesis imaging of the breast using a virtual clinical trial. In: Nishikawa RM, Samuelson FW, editors. Proc. SPIE Med. Imaging, vol. 10577, SPIE; 2018, p. 105770H–1–9. doi:10.1117/12.2293070.

[20] Elangovan P, Mackenzie A, Dance DR, Young KC, Wells K. Lesion detectability in 2D-mammography and digital breast tomosynthesis using different targets and observers. Phys Med Biol 2018;63:095014 https://doi.org/10.1088/1361-6560/aabd53.

[21] Sechopoulos I, Ghetti C. Optimization of the acquisition geometry in digital tomosynthesis of the breast. Med Phys 2009;36:1199–207. https://doi.org/10.1118/1.3090889.

[22] Goodsitt MM, Chan H-P, Schmitz A, Zelakiewicz S, Telang S, Hadjiiski L, et al. Digital breast tomosynthesis: studies of the effects of acquisition geometry on contrast-to-noise ratio and observer preference of low-contrast objects in breast phantom images. Phys Med Biol 2014;59:5883–902. https://doi.org/10.1088/0031-9155/59/19/5883.

[23] Petersson H, Dustler M, Tingberg A, Timberg P. Monte Carlo simulation of breast tomosynthesis: visibility of microcalcifications at different acquisition schemes. In: Hoeschen C, Kontos D, Flohr TG, editors. Proc SPIE Med. Imaging, vol. 9412, SPIE; 2015, p. 94121H. doi:10.1117/12.2081942.

[24] Mackenzie A, Marshall NW, Hadjipanteli A, Dance DR, Bosmans H, Young KC. Characterisation of noise and sharpness of images from four digital breast tomosynthesis systems for simulation of images for virtual clinical trials. Phys Med Biol 2017;62:2376–97. https://doi.org/10.1088/1361-6560/aa5dd9.

[25] Elangovan P, Mackenzie A, Dance DR, Young KC, Cooke V, Wilkinson L, et al. Design and validation of realistic breast models for use in multiple alternative forced choice virtual clinical trials. Phys Med Biol 2017;62:2778–94. https://doi.org/10.1088/1361-6560/aa622c.

[26] Elangovan P, Warren LM, Mackenzie A, Rashidnasab A, Diaz O, Dance DR, et al. Development and validation of a modelling framework for simulating 2D-mammography and breast tomosynthesis images. Phys Med Biol 2014;59:4275–93. https://doi.org/10.1088/0031-9155/59/15/4275.

[27] Dance DR, Skinner CL, Young KC, Beckett JR, Kotre CJ. Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol. Phys Med Biol 2000;45:3225–40. https://doi.org/10.1088/0031-9155/45/11/308.

[28] Rashidnasab A, Elangovan P, Yip M, Diaz O, Dance DR, Young KC, et al. Simulation and assessment of realistic breast lesions using fractal growth models. Phys Med Biol 2013;58:5613–27. https://doi.org/10.1088/0031-9155/58/16/5613.

[29] Dance DR, Young KC, van Engen RE. Estimation of mean glandular dose for breast tomosynthesis: factors for use with the UK, European and IAEA breast dosimetry protocols. Phys Med Biol 2011;56:453–71. https://doi.org/10.1088/0031-9155/56/2/011.

[30] Moore AC, Dance DR, Evans DS, Lawinski CP, Pitcher EM, Rust A, et al. IPEM report 89. Commissioning and Routine Testing of Mammographic X-Ray Systems. York, UK: Institue of Physics and Engineering in Medicine; 2005.

[31] European Commission. European guidelines for quality assurance in breast cancer screening and diagnosis (4th Edition). Brussels, Belgium: European Commission; 2006.

[32] Van Engen RE, Bosmans H, Bouwman RW, Dance DR, Heid P, Lazzari B, et al. Protocol for the quality control of the physical and technical aspects of digital breast tomosynthesis system (v1.03). 2018.

[33] Mackenzie A, Dance DR, Workman A, Yip M, Wells K, Young KC. Conversion of mammographic images to appear with the noise and sharpness characteristics of a different detector and x-ray system. Med Phys 2012;39:2721–34. https://doi.org/10.1118/1.4704525.

[34] Mackenzie A, Dance DR, Diaz O, Young KC. Image simulation and a model of noise power spectra across a range of mammographic beam qualities. Med Phys 2014;41:121901 https://doi.org/10.1118/1.4900819.

[35] Burgess A. Image quality, the ideal observer, and human performance of radiologic decision tasks. Acad Radiol 1995;2:522–6. https://doi.org/10.1016/S1076-6332(05)80411-8.

[36] Burgess AE. The Rose model, revisited. J Opt Soc Am A 1999;16:633. https://doi.org/10.1364/JOSAA.16.000633.

[37] Macmillan NA, Creel CD. Detection Theory, A User's Guide. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates Inc.; 2004.

[38] Karssemeijer N, Thijssen MAO. Determination of contrast-detail curves of mammography systems by automated image analysis. In: Doi K, Giger ML, Nishikawa RM, Scmidt RA, editors. Digit. Mammogr., vol. 96, Amsterdam: Elsevier; 1996, pp. 155–160.

[39] Ruschin M, Timberg P, Båth M, Hemdal B, Svahn T, Saunders RS, et al. Dose dependence of mass and microcalcification detection in digital mammography: free response human observer studies. Med Phys 2007;34:400–7. https://doi.org/10.1118/1.2405324.

[40] Hofvind S, Hovda T, Holen ÅS, Lee CI, Albertsen J, Bjørndal H, et al. Digital Breast Tomosynthesis and Synthetic 2D Mammography versus Digital Mammography: Evaluation in a Population-based Screening Program. Radiology 2018;287:787–94. https://doi.org/10.1148/radiol.2018171361.

[41] Murphy MC, Coffey L, O'Neill AC, Quinn C, Prichard R, McNally S. Can the synthetic C view images be used in isolation for diagnosing breast malignancy without reviewing the entire digital breast tomosynthesis data set? Ir J Med Sci 2018:1–5.

https://doi.org/10.1007/s11845-018-1748-7.

[42] Ambinder EB, Harvey SC, Panigrahi B, Li X, Woods RW. Synthesized Mammography: The New Standard of Care When Screening for Breast Cancer with Digital Breast Tomosynthesis? Acad Radiol 2018;25:973–6. https://doi.org/10.1016/J.ACRA.2017.12.015.

[43] Ikejimba LC, Glick SJ, Choudhury KR, Samei E, Lo JY. Assessing task performance in FFDM, DBT, and synthetic mammography using uniform and anthropomorphic physical phantoms. Med Phys 2016;43:5593–602. https://doi.org/10.1118/1.4962475.

[44] Ratanaprasatporn L, Chikarmane SA, Giess CS. Strengths and Weaknesses of Synthetic Mammography in Screening. RadioGraphics 2017;37:1913–27. https://doi.org/10.1148/rg.2017170032.

[45] Tucker AW, Lu J, Zhou O. Dependency of image quality on system configuration parameters in a stationary digital breast tomosynthesis system. Med Phys 2013;40:031917 https://doi.org/10.1118/1.4792296.

[46] Mainprize JG, Bloomquist AK, Kempston MP, Yaffe MJ. Resolution at oblique incidence angles of a flat panel imager for breast tomosynthesis. Med Phys 2006;33:3159–64. https://doi.org/10.1118/1.2241994.

[47] Hu Y-H, Zhao B, Zhao W. Image artifacts in digital breast tomosynthesis: investigation of the effects of system geometry and reconstruction parameters using a linear system approach. Med Phys 2008;35:5242–52. https://doi.org/10.1118/1.2996110.

[48] Reiser I, Nishikawa RM. Task-based assessment of breast tomosynthesis: effect of acquisition parameters and quantum noise. Med Phys 2010;37:1591–600. https://doi.org/10.1118/1.3357288.