Breast tomosynthesis and digital mammography: a comparison of diagnostic accuracy

[1]T M SVAHN, PhD, [2]D P CHAKRABORTY, PhD, [3]D IKEDA, MD, [4]S ZACKRISSON, PhD, [3]Y DO, PhD, [1]S MATTSSON, PhD and [4]I ANDERSSON, MD

[1]*Medical Radiation Physics, Department of Clinical Sciences Malmö, Lund University, Skåne University Hospital, Malmö, Sweden, [2]Department of Radiology, University of Pittsburgh, Pittsburgh, PA, USA, [3]Department of Radiology, Stanford University, Stanford Advanced Medicine Center, Stanford, CA, USA, and [4]Medical Radiology, Department of Clinical Sciences Malmö, Lund University, Skåne University Hospital, Malmö, Sweden*

Objective: Our aim was to compare the ability of radiologists to detect breast cancers using one-view breast tomosynthesis (BT) and two-view digital mammography (DM) in an enriched population of diseased patients and benign and/or healthy patients.
Methods: All participants gave informed consent. The BT and DM examinations were performed with about the same average glandular dose to the breast. The study population comprised patients with subtle signs of malignancy seen on DM and/or ultrasonography. Ground truth was established by pathology, needle biopsy and/or by 1-year follow-up by mammography, which retrospectively resulted in 89 diseased breasts (1 breast per patient) with 95 malignant lesions and 96 healthy or benign breasts. Two experienced radiologists, who were not participants in the study, determined the locations of the malignant lesions. Five radiologists, experienced in mammography, interpreted the cases independently in a free-response study. The data were analysed by the receiver operating characteristic (ROC) and jackknife alternative free-response ROC (JAFROC) methods, regarding both readers and cases as random effects.
Results: The diagnostic accuracy of BT was significantly better than that of DM (JAFROC: $p=0.0031$, ROC: $p=0.0415$). The average sensitivity of BT was higher than that of DM (~90% *vs* ~79%; 95% confidence interval of difference: 0.036, 0.108) while the average false-positive fraction was not significantly different (95% confidence interval of difference: −0.117, 0.010).
Conclusion: The diagnostic accuracy of BT was superior to DM in an enriched population.

About 1 in 8–10 females develop breast cancer during their lifetime [1, 2]. Screening mammography plays a key role in the detection of breast cancer at an early stage. Based on incidence of interval cancers it has been suggested that a radiologist reading screen-film mammograms might miss 16–30% of cancers detectable on the mammograms [3]. Mammography cancer detection varies widely: estimates of sensitivity have been reported from 68% (or as low as 48% for extremely dense breasts) to 88%, with specificities ranging from 82% to 98%. These results suggest that there is considerable room for improvement in mammography [4, 5]. Digital mammography (DM) was expected to improve the performance of breast cancer detection compared with screen-film mammography (SFM). In most clinical trials the overall sensitivity has been higher for DM, but, since the specificities have also been lower, only a few studies have been statistically significant in favour of DM [5]. In a subset of females under 50 years of age in the Digital Mammographic Imaging Screening Trial study, there was a significantly improved diagnostic accuracy in DM compared with SFM [5].

Because a mammogram is a two-dimensional (2D) projection of the breast onto the detector plane, over-projected healthy tissue (anatomical noise) can hamper breast cancer detectability. Anatomical noise is known to have a greater impact than quantum noise on the detection of certain breast cancers (*e.g.* masses) [6, 7]. Two views—mediolateral oblique (MLO) and craniocaudal (CC)—can partially compensate for the overlapping anatomical noise, but this depends on the radiologist's ability to mentally fuse the two images.

Breast tomosynthesis (BT) collects 2D projection views over a limited angular range, which allows reconstruction of thin slices of the breast volume. Reduced anatomical noise from superimposed tissues is expected to improve breast cancer detection compared with DM. In CT where hundreds of projection images are acquired

covering 360°, the anatomical noise can be reduced to a larger degree, but it is difficult to image the entire breast volume using CT, particularly close to the chest wall. Moreover, the average glandular dose is higher with CT, as is imaging time and the cost of the device. While there is ongoing research that may solve these issues [8–10], BT has a number of potential advantages and there are currently commercialised units.

Previous studies of observer performance of BT compared with DM have shown contradictory results, varying from a statistically significant advantage for BT [11–13] to no clear advantage for BT [14–18]. Non-blinded pilot studies have been performed at our institution that suggest improved sensitivity of BT over DM [19, 20].

The aim of the current study was to compare the diagnostic accuracy of one-view BT with conventional two-view DM using an enriched population.

Methods and materials

Patient population

The study protocol was approved by the Regional Ethics Review Board at Lund University (Dnr 159/2006) and the local Radiation Safety Committee at Skåne University Hospital, Malmö. All patients provided informed consent. The study population included symptomatic as well as asymptomatic females examined at our institution during a 23-month period from 19 June 2006 to 21 May 2008. Patients with suspicious findings underwent fine needle or core biopsy followed by surgery and histopathological examination of the specimens. Patients without suspicious findings were followed for 1 year [18, 21] to establish absence of cancer. Of 185 patients (breasts), 89 were proven to have cancer (containing 95 abnormalities), whereas 95 of the cases were established to be benign or healthy. The average age of the patients was 60 years (range: 42–79 years). The study included one breast per patient and featured many difficult cases, as is normal in observer performance studies where the natural incidence of cancer is very low [22, 23]. The majority of the lesions were difficult to see in the DM views, but usually seen on ultrasonography [19].

Image acquisition

DM was performed using a Mammomat Novation DR unit (Siemens, Erlangen, Germany) and BT was performed using a prototype device based on the same type of unit. The detector used in this prototype BT system was an amorphous selenium flat-panel detector [24]. The BT and DM images were acquired using the same tube voltage and anode/filter (W/Rh) combination as determined by the automatic exposure control of the DM unit. BT examinations were performed immediately following the DM/ultrasound examinations. The BT examinations were performed using about twice the mAs of a single DM image for that patient in the same projection, resulting in approximately the same absorbed dose as a dual-view mammography examination [25]. For the BT examinations, 25 projection images were acquired over an angular range of approximately $\pm 20°$. The scan time using full-resolution mode of the BT unit was about 20 s [24]. The images were

reconstructed using filtered back-projection [26, 27]. Each DM examination consisted of the MLO and CC views. The MLO view was chosen for the BT examinations in the majority of the cases (88%), since the lesion was least visible in this DM view [19]. In the remaining 12% of the cases BT was performed in the CC view.

Ground truth lesion locations

Two radiologists (IA and DI, referred to as the ''truth panel''), who were not participants in the study and experienced in BT and DM, outlined the malignant regions in consensus using an electronic marker. They had access to all the available data from BT, DM, ultrasonography, needle biopsy and pathology. In the DM cases the corresponding malignant regions of the breast were outlined in the CC and the MLO view. In the BT image volumes the outlining was made in three slices: the initial, focus (central) and final slice where the lesion appeared. In cases where a lesion was not visible in BT or DM to the truth panel, but was seen and localised to a quadrant on ultrasound images and the histopathological examination, the lesion-containing quadrant was outlined (this event occurred twice). The breast cancer visibility was classified in the MLO and CC views of the DM cases into three categories: not visible, subtle and visible. The breast density was classified according to the Breast Imaging-Reporting and Data System (BI-RADS) [28]: 1, fatty or <25% dense; 2, scattered fibroglandular densities or 25–50% dense; 3, heterogeneously dense or 50–75% dense; 4, dense or >75% dense.

Graphical user interface

A graphical user interface, ViewDEX [29], was used to display images and to record the data. The workstation consisted of a Sun Microsystem Ultra 24 Workstation (2.5 GHz, 4 Gb RAM) with two five-megapixel flat-panel calibrated monitors (SMD21500; EIZO GmbH, Karlsruhe, Germany). The minimum luminance was 0.4 cd m^{-2} and the maximum luminance was 355 cd m^{-2}, measured with recommended instrumentation [30, 31]. The ambient light level was less than 3 lux, which follows European guidelines [32]. Although window/level settings were pre-set by an experienced radiologist (IA) for the BT cases, and by the built-in software of the DM system, the radiologists were free to alter the settings during the study.

Reader study

Five dedicated breast radiologists with experience of breast imaging ranging between 3 and 25 years (average 16.6 years), who were blinded to the truth status (positive/negative) of the images, interpreted them using the free-response paradigm [33, 34]. The task was to mark the centres of all perceived cancers and assign malignancy ratings to them. The probability of malignancy was rated on a BI-RADS-based scale [35]: BI-RADS 1, 2, 3, 4A, 4B, 4C and 5. The zero and the sixth rating were not allowed since this was an experimental study. For the purpose of

the study the BI-RADS 2 (definitely benign) marks were ignored. The radiologists were instructed to mark/rate the lesions if visible on both views of the DM images and they were allowed to use the zoom and pan functions. The cine-loop mode was sequentially displaying the BT slices at a user-controlled rate.

The location of each mark was compared with the nearest malignant lesion, if any, and if it was inside the truth panel's outlined region (areal for DM and volumetric for BT) it was classified as a lesion localisation (LL) and if outside as a non-lesion localisation (NL). In DM, if a pair of marks in two views corresponded to the same physical location, in the opinion of the truth panel, the mark with the highest rating was used in the analysis. If one mark corresponded to the lesion but the other did not, the first was recorded as an LL and the other as an NL.

No limit was imposed on the viewing time. Prior to the study, the radiologists underwent a training session using 30 corresponding BT and DM cases, which were not used in the actual study. An accompanying expert radiologist familiarised each reader with the user interface and with the appearance of the healthy tissue and the various cancer types in BT. In each of the eight reading sessions, the cases were presented in random order in two blocks of 25 cases per modality (*i.e.* 25 one-view BT cases and 25 two-view DM cases). The modality presentation order was alternated and a period of 1–3 weeks separated consecutive viewings of the same case in the two modalities [22].

Statistical analysis

The LL and NL ratings were analysed by the jackknife alternative free-response receiver operating characteristic (JAFROC) method [33, 34]. The JAFROC figure of merit (θ) is the probability that the rating of the highest-rated and correct LL on a diseased case exceeds the rating of the highest-rated mark on a healthy/benign case (NL). This is equivalent to the non-parametric area under the AFROC curve [36, 37]. NLs on abnormal images are not used in the data representation of the AFROC curve. Significance testing was performed using the Dorfman–Berbaum–Metz multiple-reader multiple-case (DBM-MRMC) mixed model analysis of variance (ANOVA) procedure applied to θ (the ANOVA module was provided by Dr Kevin Berbaum). ANOVA yields an *F*-statistic and a *p*-value for rejecting the null hypothesis that the modalities have identical performance. Random-reader and random-case analysis was performed. Parametric AFROC curves (reader-averaged) fitted by search model [36, 37] are presented for illustration purposes. As a check, ROC analysis was also performed using the highest rating on a case as the inferred ROC rating. DBM-MRMC software [38] with the PROPROC [39] fitted area under the ROC curve (AUC; the probability that the value of the highest rated mark on an diseased case exceeds that on a healthy/benign case) was used as the figure of merit for ROC analysis. Note that the only difference between JAFROC analysis of FROC data and DBM-MRMC analysis of ROC data is in the definition of the figure of merit.

Table 1. The number of breast cancers by visibility on the digital mammography mediolateral oblique (MLO) and craniocaudal (CC) views

Visibility	MLO	CC	MLO+CC
Not visible	24[a]	22[b]	12
Subtle	33	16	21
Visible	38	57	62
Total	95	95	95

CC, craniocaudal; MLO, mediolateral oblique.
[a]In one case the breast cancer was not included in the MLO view.
[b]In five cases the breast cancer was not included in the CC view.

Results

Pathological findings

Of 95 breast cancers, 57 (60%) were invasive ductal carcinomas (IDCs), 19 (20%) invasive lobular carcinomas (ILCs), 6 (6.3%) were tubular carcinomas (Tubs), 6 (6.3%) were ductal carcinomas *in situ* (DCISs), 1 (1.1%) was lobular carcinoma *in situ* (LCIS), 1 (1.1%) was IDC+ILC, 1 (1.1%) was IDC+LCIS, 2 (2.1%) were mucinous carcinomas, 1 (1.1%) was mucinous carcinoma+ILC and 1 (1.1%) was intracystic papillary carcinoma. The largest average lesion diameter was 19.9 mm (range 2–90 mm, median 15 mm), as measured at the histopathological examination.

Breast cancer visibility and breast density

In 12 out of 95 breast cancer cases the cancer was considered neither visible in the MLO nor CC of the DM views (Table 1). In another 21 cases the cancer was considered subtle in one of both DM views and clearly visible in 62 cases. The breast cancers are shown with regard to visibility and histopathology group (Table 2). The breast density of the diseased cases was classified as BI-RADS 1 in 9 (10.1%) cases, BI-RADS 2 in 26 (29.2%) cases, BI-RADS 3 in 50 (56.2%) cases and BI-RADS 4 in 4 (4.5%) cases. The breast cancers are listed according to BI-RADS and histopathology group (Table 3). The breast

Table 2. The numbers of breast cancers per histopathology group by visibility (MLO+CC)

Histopathology group	Visible	Subtle	Not visible
IDC	39 (41.1)	11 (11.6)	7 (7.4)
ILC	8 (8.4)	7 (7.4)	4 (4.2)
Tub	3 (3.2)	2 (2.1)	1 (1.1)
DCIS	6 (6.3)	0	0
Other groups	6 (6.3)	1 (1.1)	0
Total	62 (65.3)	21 (22.1)	12 (12.6)

CC, craniocaudal; DCIS, ductal carcinoma *in situ*; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; MLO, mediolateral oblique; Tub, tubular carcinoma.
The percentages of the total numbers of breast cancers are given in the parentheses.

Table 3. The numbers of breast cancers per histopathology group with a specific BI-RADS breast density value

Histopathology group	BI-RADS 1	BI-RADS 2	BI-RADS 3	BI-RADS 4
IDC	7 (7.4)	18 (18.9)	31 (32.6)	1 (1.1)
ILC	1 (1.1)	4 (4.2)	14 (14.7)	0
Tub	0	2 (2.1)	2 (2.1)	2 (2.1)
DCIS	0	3 (3.2)	3 (3.2)	0
Other groups	2 (2.1)	2 (2.1)	2 (2.1)	1 (1.1)
Total	10 (10.5)	29 (30.5)	52 (54.7)	4 (4.2)

BI-RADS, Breast Imaging-Reporting and Data System; CC, craniocaudal; DCIS, ductal carcinoma *in situ*; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinomas; MLO, mediolateral oblique; Tub, tubular carcinoma.
The percentages of the total numbers of breast cancers are given in the parentheses.

density of the healthy/benign cases was classified as BI-RADS 1 in 19 (19.8%) cases, BI-RADS 2 in 27 (28.1%) cases, BI-RADS 3 in 40 (41.7%) cases and BI-RADS 4 in 10 (10.4%) cases.

Observer performance

All five radiologists had higher θ values (the non-parametric area under the AFROC curve; Table 4) and four of the five radiologists had higher ROC figure of merits (AUC) for the BT than for the DM modality (Table 5). For both methods the reader-averaged modality differences were statistically significant at the 5% level—that is, the 95% confidence interval (95% CI) of the difference did not include 0, 0.103 (95% CI: 0.039, 0.167) and 0.094 (95% CI: 0.004, 0.183) for θ and AUC, respectively. The average AFROC curve plots of lesion localisation fraction (LLF) *vs* false-positive fraction (FPF; Figure 1a), and the average ROC curve plots of true-positive fraction (TPF) *vs* FPF (Figure 1b) are shown for all readers. Note that the x-axes of the AFROC curves are identical to those of the ROC curves.

An average of 10.4 more breast cancer(s) were detected per reader, with the correct malignant region localised, on BT than on DM, as shown by the average LLFs (LLF = number of localised lesions divided by the total number of lesions): 0.809 for BT and 0.706 for DM (Table 6). The sensitivity (TPF) of BT was significantly higher than that for DM (95% CI: 0.036, 0.108). The difference in FPF (1–specificity) was not significantly different (95% CI: −0.117, 0.010) but varied substantially between readers; Readers 1 and 3 had higher FPF values in both modalities (between 60% and 72%) relative to the others (between 22% and 42%).

Breast cancer distribution and detection by histopathology

The breast cancer distribution was dominated by the IDC and ILC groups (Figure 2a). In general, more of all lesion types, except for the DCIS, were detected with BT (Figure 2b), which is reflected in the higher TPF and narrower confidence intervals. Two examples of cases are shown in Figures 3 and 4, representing the histopathology group IDC (detected by more of the readers on BT than on DM).

Discussion

This study showed a higher diagnostic accuracy of one-view BT compared with two-view DM and was demonstrated by JAFROC as well as ROC analyses. The improved performance with BT in relation to DM is consistent with the result of prior studies [11–13, 19, 20]. The sensitivity of BT was higher than that of DM (average ~90% *vs* average ~79%), while the difference in average FPF (1–specificity) values was not significantly different. Out of 95 breast cancers, an average of 10.4 more cancers per reader were detected with BT than with DM. The observed sensitivity of DM (average 79%) is in the range of reported values (48–88%), but the FPF (45%) was higher than reported values (2–18%) [4, 5]. There are several possible explanations for the higher FPF. (i) The cases were more difficult than the average case encountered in clinical practice. (ii) The case set was enriched with diseased cases (study prevalence ~45% *vs* 0.5% in the general population). (iii) In the clinic the radiologist has access to clinical history including the reason for the referral, the family history and so on. Images from previous screening examinations are also available. In

Table 4. Jackknife alternative free response receiver operating characteristic (JAFROC) figures of merit (θ) by reader and modality

Radiologist	BT (θ)	DM (θ)	BT-DM (Δ_θ)
1	0.759 (5)	0.641 (5)	0.118 (3)
2	0.838 (1)	0.704 (3)	0.134 (2)
3	0.813 (3)	0.736 (2)	0.077 (4)
4	0.807 (4)	0.769 (1)	0.038 (5)
5	0.828 (2)	0.681 (4)	0.147 (1)
Mean	0.809[a] (0.756, 0.862)	0.706[a] (0.638, 0.774)	0.103 (0.039, 0.167)

BT, one-view breast tomosynthesis; DM, two-view digital mammography.
The last row lists the averages and 95% confidence intervals in parentheses.
[a]The *p*-value was 3.1×10^{-3}, Fisher *F*-statistic=10.9, numerator and denominator degrees of freedom equal to 1 and 23.1, respectively.

Table 5. The receiver operating characteristic (ROC) figures of merit (AUC) by reader and modality

Radiologist	BT (AUC)	DM (AUC)	BT-DM (Δ_{AUC})
1	0.801 (5)	0.672 (5)	0.130 (3)
2	0.895 (1)	0.754 (3)	0.140 (2)
3	0.853 (4)	0.793 (2)	0.059 (4)
4	0.858 (3)	0.874 (1)	−0.016 (5)
5	0.891 (2)	0.736 (4)	0.155 (1)
Mean	0.860[a] (0.804, 0.915)	0.766[a] (0.673, 0.859)	0.094 (0.004, 0.183)

BT, one-view breast tomosynthesis; DM, two-view digital mammography.
The last row lists the averages and 95% confidence intervals in parentheses.
[a]The *p*-value was 0.0415; Fisher *F*-statistic=5.54; numerator and denominator degrees of freedom were equal to 1 and 9.56, respectively.

this study images of only one breast were provided, whereas in the clinical setting images of both breasts are available. The additional information gives the radiologist more confidence in the diagnosis. These circumstances may cause the operating point to move towards higher FPF [37, 40]. However, since other studies have not found a significant effect of higher prevalence on the diagnostic accuracy [37, 40], the shift in the operating point is not expected to affect the overall conclusion of this study (*i.e.* higher performance of one-view BT relative to two-view DM).

Gur et al compared the performance of a combined modality consisting of two-view BT and two-view DM with two-view DM, and found significantly higher specificity for the combined modality, but the difference in sensitivity was not significant [16]. The current study findings, namely a significantly higher sensitivity using BT but a non-significant difference in specificity, are inconsistent with the study of Gur et al [16], whose findings can be explained by a shift in reporting threshold. Had the readers in the prior study adopted a laxer reporting criterion in BT, then the specificity difference might have been smaller and the sensitivity higher. Furthermore, our study demonstrated a significantly higher overall performance, while the prior study found a significant difference in specificity.

Another study [18] comparing one-view BT with two-view DM found a non-significant advantage for one-view BT (Δ_{AUC}=0.015). In that study the readers had relatively limited experience with the BT modality and the equipment was not fully optimised. It might also be that our study population contained more difficult cases. These factors might explain the different outcome.

Recall bias due to the relatively short time interval of 1–3 weeks between sessions is expected to be minimal since the cases were displayed in random order. Moreover, it has been shown that memory washout occurs within days [41]. Also, since the appearance of the images by the different techniques was markedly different (2D *vs* 3D), memory effects are expected to be further attenuated.

It is encouraging that BT performed better than DM even though the radiologists had much less experience in the BT modality than in DM, and with only one view available in BT compared with two views in DM. This suggests that with increased experience, and perhaps with the addition of the corresponding ipsilateral DM view, BT would have performed even better, as shown in a recent study [12]. As noted previously, for the currently used exposure parameter settings, the system used in this study delivered about the same dose as a two-view DM [25]. The addition of the corresponding ipsilateral DM view would increase the dose by about 50%. Moreover, further improvements are to be expected in the evolving BT technology.
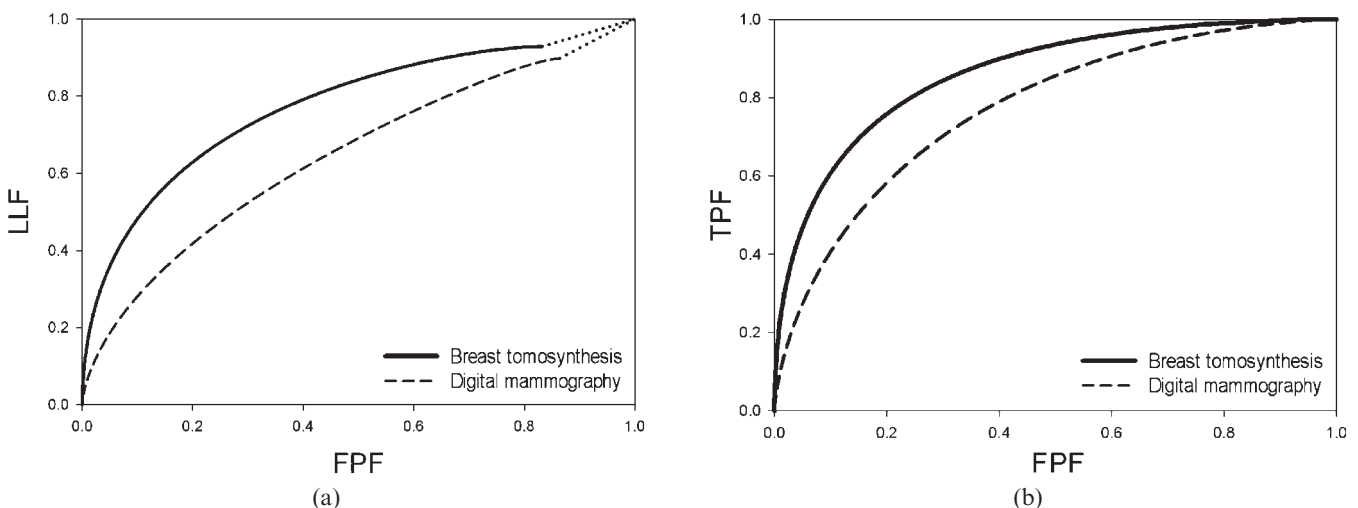

Figure 1. (a) The average alternative free-response receiver operating characteristic curves: breast tomosynthesis (BT)=0.809; digital mammography (DM)=0.706; Δ_{BT-DM}=0.103. (b) The average receiver operating characteristic curves: BT=0.860; DM=0.766; Δ_{BT-DM}=0.094. FPF, false-positive fraction; LLF, lesion localisation, fraction; TPF, true-positive fraction.
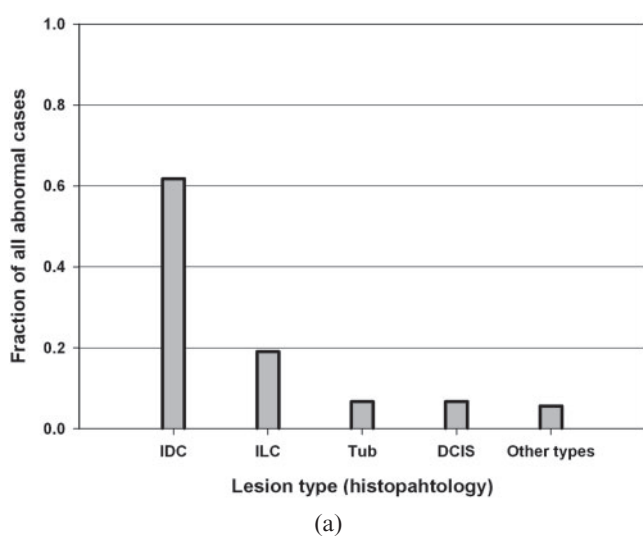
Table 6. The lesion localisation fractions (LLFs), true-positive fractions (TPFs; sensitivity) and false-positive fraction (FPF; 1–specificity) by reader and modality

Radiologist	LLF (BT/DM)	TPF (BT/DM)	FPF (BT/DM)
1	0.821/0.758	0.888/0.809	0.667/0.667
2	0.863/0.758	0.910/0.787	0.365/0.417
3	0.916/0.821	0.966/0.888	0.719/0.604
4	0.832/0.663	0.876/0.719	0.313/0.156
5	0.811/0.695	0.843/0.742	0.219/0.396
Mean	0.848/0.739	0.897/0.789	0.456/0.448

BT, one-view breast tomosynthesis; DM, two-view digital mammography.

The distribution of cancer types was dominated by IDC (61.8%) and ILC (19.1%; Figure 2a). The proportion of ILC (19.1%) was higher than in published materials [42], probably owing to the case selection. It is promising that the TPF for this histopathology group was higher for BT than for DM (Figure 2b) since a proportion of the ILCs are known to have a growth pattern making them difficult to detect with even the best diagnostic techniques [42–44]. The TPF was higher for BT than for DM for all histopathology groups, except for the DCIS group (Figure 2b). However, since the number of lesions was limited for this group (Figure 2a) the detectability of DCIS needs to be further investigated. Microcalcifications are a dominant indicator for malignancy for the DCIS group. Early impressions on microcalcifications on BT in relation to DM have been that they are equally visible, but the morphological details of the individual calcification have not been as well visualised on BT [19]. However, further development in the processing may change this; in a recent study BT has been shown to have equal or greater clarity regarding microcalcifications than DM using iterative reconstruction algorithms [45].

The readers had extensive experience with DM, but limited experience with BT. This is a problem when evaluating any new technology. The in-house developed user interface is different from what the radiologists would expect in a mature commercial product, which could also cause BT performance to be non-optimal. It would be highly desirable for equipment manufacturers to implement a data collection tool in their products that would allow performance to be measured under more realistic conditions. Owing to the design of the study, including the use of an enriched population, the results cannot necessarily be extrapolated to a general female population. It is therefore desirable to assess the impact of BT and its cost–benefit aspects in routine clinical conditions, as well as in the screening setting, but this would require a large-scale clinical trial. Although technically the results of this study were generalised to the population of readers and cases, such generalisation may not be valid. The cases were selected at a single institution, which may not be typical of other centres. Less variability is expected between readers from the same institution than readers from different institutions. A large population-based study is needed, using readers and cases from different institutions, with readers who are more experienced in BT.

Conclusions

In this study, BT in one view (MLO view) was compared with DM in two views (MLO and CC views) using a population enriched with difficult cases. The results of the study showed superior diagnostic accuracy in BT than in DM. This suggests that breast cancer detection can be improved with BT, but the result needs to be confirmed in large population-based studies. The cost–benefit aspects of screening with BT compared with DM also need to be assessed.
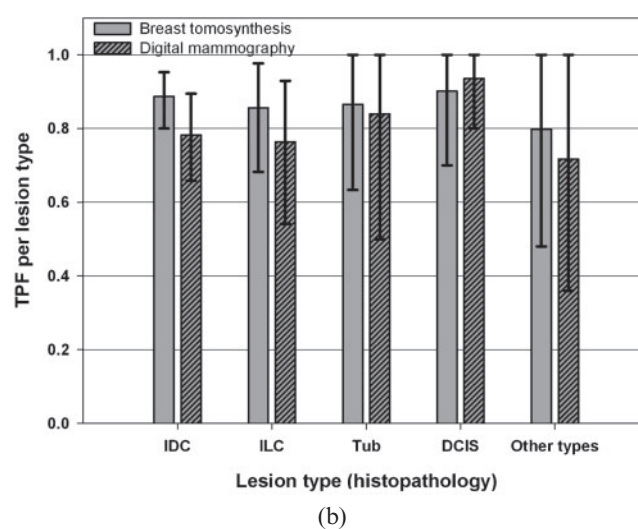
Acknowledgments

(a)

(b)

Figure 2. (a) The distribution of breast cancers by histopathology of 89 diseased cases with a total of 95 breast cancers. (b) The reader-averaged true-positive fraction (TPF; sensitivity) by histopathology with 95% confidence intervals for breast tomosynthesis and digital mammography. DCIS, ductal carcinoma *in situ*; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; Tub, tubular carcinoma.
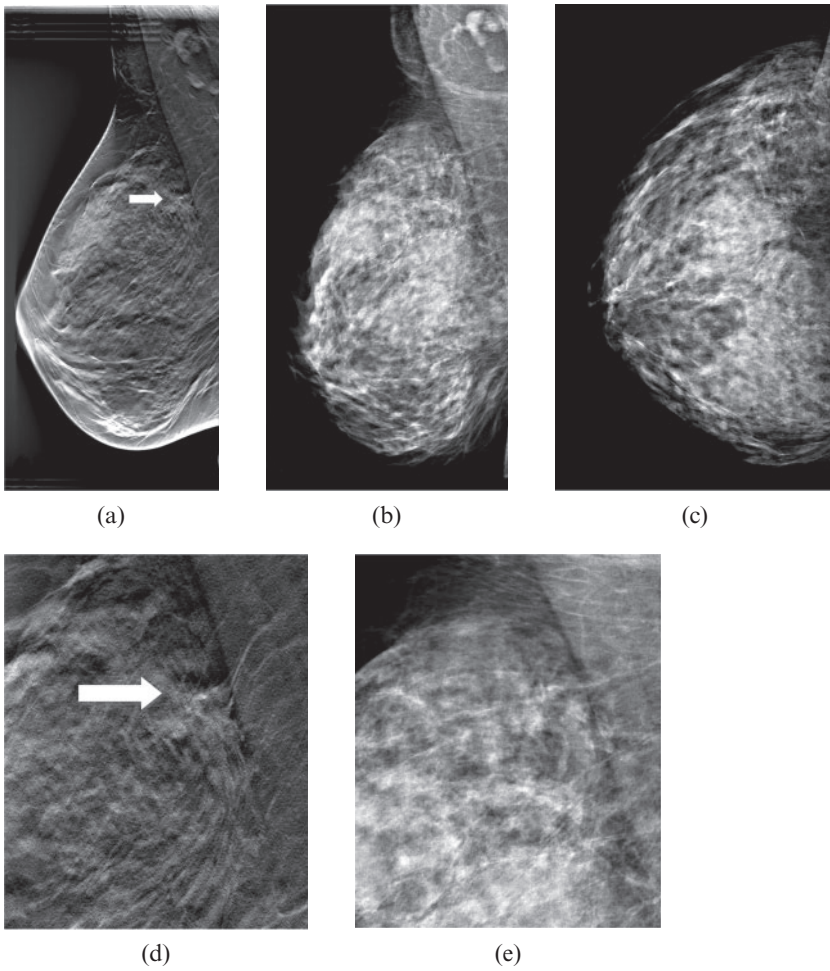
Figure 3. (a) Breast tomosynthesis (BT) mediolateral oblique (MLO) view (one-view BT), (b) digital mammography (DM) MLO view and (c) DM craniocaudal (CC) view (two-view DM). (d, e) Close-ups of (a, b). A 69-year-old female with a 15-mm (diameter) spiculated tumour, invasive ductal carcinoma grade 2. Owing to its juxtathoracic position, the tumour was not included in the CC view (c). The tumour, indicated by the arrow in (a, d), was detected on one-view BT by all readers, while on two-view DM (b, c) it was not detected by any of the readers.
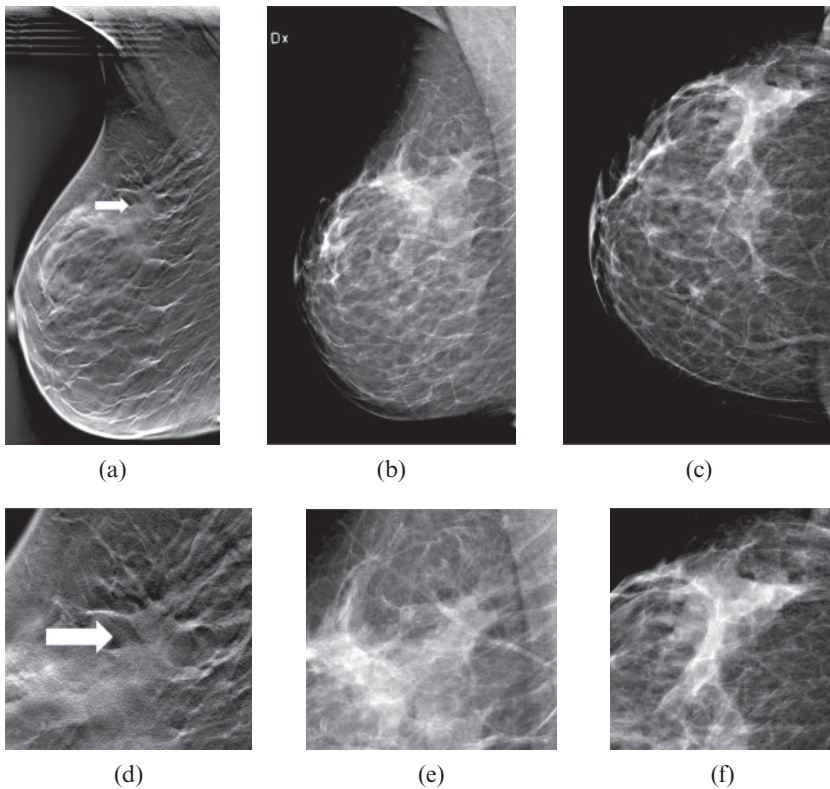

Figure 4. (a) Breast tomosynthesis (BT) mediolateral oblique (MLO) view (one-view BT), (b) digital mammography (DM) MLO view and (c) DM craniocaudal (CC) view (two-view DM). (d–f) Close-ups of (a–c). A 79-year-old female with a 10-mm (diameter) spiculated tumour (arrow in a and d), invasive ductal carcinoma grade 3. The tumour was detected on one-view BT by 4/5 readers, while on two-view DM (b, c) the tumour was undetected by all readers.

References

1. Sasieni PD, Shelton J, Ormiston-Smith N, Thomson CS, Silcocks PB. What is the lifetime risk of developing cancer? The effect of adjusting for multiple primaries. Br J Cancer 2011;105:406–65.

2. Socialstyrelsen.se [homepage on the Internet]. Stockholm, Sweden: Swedish National Board of Health and Welfare; Cancer Incidence in Sweden 2008 [cited 9 July 2011]. Available from: www.socialstyrelsen.se/publikationer2009/2009-12-1

3. Laming D, Warren R. Improving the detection of cancer in the screening of mammograms. J Med Screen 2000;7:24–30.

4. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27 825 patient evaluations. Radiol 2002;225:165–75.

5. Skaane P. Studies comparing screen-film mammography and full-field digital mammography in breast cancer screening: updated review. Acta Radiol 2009;50:503–14.

6. Bochud FO, Valley JF, Verdun FR, Hessler C, Schnyder P. Estimation of the noisy component of anatomical backgrounds. Med Phys 1999;26:1365–70.

7. Burgess AE, Jacobson FL, Judy PF. Human observer detection experiments with mammograms and power-law noise. Med Phys 2001;28:419–37.

8. Boone JM, Kwan ALC, Yang K, Burkett, GW, Lindfors KK, Nelson TR. Computed tomography for imaging the breast. J Mammary Gland Biol Neoplasia 2006;11:103–11.

9. Lindfors KK, Boone JM, Nelson TR, Yang K, Kwan ALC, Miller DF. Dedicated breast CT: initial clinical experience. Radiology 2008;246:725–33.

10. Kalender WA. Concepts for high-resolution CT of the breasts. LNCS 2010;6136:421–7.

11. Smith AP, Rafferty EA, Niklason L. Clinical performance of breast tomosynthesis as a function of radiologist experience level. LNCS 2008;5116:61–6.

12. Svahn T, Andersson I, Chakraborty D, Svensson S, Ikeda D, Förnvik D, et al. The diagnostic accuracy of dual-view digital mammography, single-view breast tomosynthesis and a dual-view combination of breast tomosynthesis and digital mammography in a free-response observer performance study. Radiat Prot Dosim 2010;139:113–17.

13. Michell MJ, Wasan RK, Iqbal A, Peacock C, Evans DR, Morel JC. Two-view 2D digital mammography versus one-view digital breast tomosynthesis. Breast Cancer Res 2010;12(Suppl. 3):3.

14. Poplack SP, Tosteson TD, Kogel CH, Nygy HM. Digital breast tomosynthesis: initial experience in 98 women with abnormal digital screening mammography. AJR Am J Roentgenol 2007;189:616–23.

15. Good WF, Abrams GS, Catullo VJ, Chough DM, Ganott MA, Hakim CM, et al. Digital breast tomosynthesis: a pilot observer study. AJR Am J Roentgenol 2008;190:865–9.

16. Gur D, Abrams GS, Chough DM, Ganott MA, Hakim CM, Perrin RL, et al. Digital breast tomosynthesis—an observer performance study. AJR Am J Roentgenol 2009;193:586–91.

17. Teertstra HJ. Breast tomosynthesis in clinical practice: initial results. Eur Radiol 2010;20:16–24.

18. 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 Jul;20:1545–53.

19. Andersson I, Ikeda DM, Zackrisson S, Ruschin M, Svahn T, Timberg P, et al. Breast tomosynthesis and digital mammography: a comparison of breast cancer visibility and BI-RADS classification in a population of cancers with subtle mammographic findings. Eur Radiol 2008;18:2817–25.

20. Förnvik D, Zackrisson S, Ljungberg O, Svahn T, Timberg P, Tingberg A, et al. Breast tomosynthesis: accuracy of tumor measurement compared with digital mammography. Acta Radiol 2010;51:240–7.

21. Pisano ED, Hendrick RE, Yaffe MJ, Baum JK, Acharyya S, Cormack JB, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. Radiol 2008;246:376–83.

22. Metz C. Some practical issues of experimental design and data analysis in radiological ROC studies. Invest Radiol 1989;24:234–45.

23. Zhou XH, Obuchowski NA, McClish DK. Statistical methods in diagnostic medicine. New York, NY: Wiley & Sons; 2002.

24. Bissonnette M, Hansroul M, Masson E, Savard S, Cadieux S, Warmoes P, et al. Digital breast tomosynthesis using an amorphous selenium flat panel detector. Proc SPIE 2005;5745:529–40.

25. Sechopoulos I, Suryanaranayan S, Vedantham S, D'Orsi C, Karellas A. Computation of the glandular radiation dose in digital tomosynthesis of the breast. Med Phys 2007;34:221–32.

26. Mertelmeier T, Orman J, Haerer W, Kumar MK. Optimizing filtered backprojection reconstruction for a breast tomosynthesis prototype device. Proc SPIE 2006;6142:131–42.

27. Zhao B, Zhou J, Hu YH, Mertelmeier T, Ludwig J, Zhao W. Experimental validation of a three-dimensional linear system model for breast tomosynthesis. Med Phys 2009;36:240–51.

28. American College of Radiology. Breast imaging reporting and data system (BI-RADS). 3rd edn. Reston, VA: American College of Radiology; 1998.

29. Håkansson M, Svensson S, Zachrisson S, Svalkvist A, Båth M, Månsson LG. VIEWDEX: an efficient and easy-to-use software for observer performance studies. Radiat Prot Dosimetry 2010;139:42–51.

30. DICOM [homepage on the Internet]. DICOM Part 14 Grayscale Standard Display Function PS 3.14-2011. Rossly, VA: National Electrical Manufacturers Association (NEMA), Digital Imaging and Communications in Medicine (DICOM) [cited 24 May 2012]. Available from: http://medical.nema.org/standard.html

31. Samei E, Badano A, Chakraborty DP, Compton K, Cornelius C, Corrigan K, et al. Assessment of display performance for medical imaging systems: report of the American Association of Physicists in Medicine (AAPM). AAPM On-Line Report no. 03. Madison, WI, Task Group 18, Medical Physics Publishing; 2005.

32. Van Engen R, Young KC, Bosmans H, Thijssen M. The European protocol for the quality control of the physical and technical aspects of mammography screening. Part B: Digital mammography. In: European Guidelines for Breast Cancer Screening. 4th edn. Luxembourg: European Commission; 2006.

33. Chakraborty DP, Berbaum KS. Radiologist studies involving detection and localization: modeling, analysis and validation. Med Phys 2004;31:2313–30.

34. Chakraborty DP. Recent advances in radiologist performance methodology: jackknife free-response receiver operating characteristic (JAFROC) methodology. Radiat Prot Dosim 2005;114:45–52.

35. ACR.org [homepage on the Internet]. BI-RADS ATLAS and MQSA: frequently asked questions. Reston, VA: The American College of Radiology [updated: 8 November 2011; cited 24 June 2010]. Available from: www.acr.org/SecondaryMainMenuCategories/quality_safety/BI-RADSAtlas/BI-RADSFAQs.aspx

36. Chakraborty DP. A search model and figure of merit for observer data acquired according to the free-response paradigm. Phys Med Biol 2006;51:3449–62.

37. Chakraborty DP. New developments in observer performance methodology in medical imaging. Semin Nucl Med 2011;41:401–18.
38. Dorfman DD, Berbaum KS, Metz CE. ROC characteristic rating analysis: generalization to the population of readers and patients with the jackknife method. Invest Radiol 1992;27:723–31.
39. Metz CE, Pan X. ''Proper'' binormal ROC curves: theory and maximum-likelihood estimation. J Math Psychol 1999;43:1–33.
40. Gur D, Rockette HE, Armfield DR, Blachar A, Bogan JK, Brancatelli G, et al. Prevalence effect in a laboratory environment. Radiol 2003;228:10–14.
41. Hardesty LA, Ganott MA, Hakim CM, Cohen CS, Clearfield RJ, Gur D. ''Memory effect'' in observer performance studies of mammograms. Acad Radiol 2005;12:286–90.
42. Hilleren DJ, Andersson I, Lindholm K, Linnell FS. Invasive lobular carcinoma: mammographic findings in a 10-year experience. Radiology 1991;178:149–54.
43. Gathani T, Bull D, Green J, Reeves G, Beral V; Million Women Study Collaborators. Breast cancer histological classification: agreement between the Office for National Statistics and the National Health Service Breast Screening Programme. Breast Cancer Res 2005;7:R1090–6.
44. Klemi PJ, Parvinen L, Pylkkänen L, Kauhava L, Immonen-Räihä P, Räsänen O, et al. Significant improvement in breast cancer survival through population-based mammography screening. Breast 2003;12:308–13.
45. Kopans D, Gavenonis S, Halpern E, Moore R. Calcifications in the Breast and Digital Breast Tomosynthesis. Breast J 2011;17:638–44.