Eur Radiol
DOI 10.1007/s00330-013-2831-0

BREAST

Combination of one-view digital breast tomosynthesis with one-view digital mammography versus standard two-view digital mammography: per lesion analysis

Gisella Gennaro · R. Edward Hendrick · Alicia Toledano · Jean R. Paquelet · Elisabetta Bezzon · Roberta Chersevani · Cosimo di Maggio · Manuela La Grassa · Luigi Pescarini · Ilaria Polico · Alessandro Proietti · Enrica Baldan · Fabio Pomerri · Pier Carlo Muzzio

Received: 21 December 2012 / Revised: 7 February 2013 / Accepted: 21 February 2013

Abstract

Objective To evaluate the clinical value of combining one-view mammography (cranio-caudal, CC) with the complementary view tomosynthesis (mediolateral-oblique, MLO) in comparison to standard two-view mammography (MX) in terms of both lesion detection and characterization.

Methods A free-response receiver operating characteristic (FROC) experiment was conducted independently by six breast radiologists, obtaining data from 463 breasts of 250 patients. Differences in mean lesion detection fraction (LDF) and mean lesion characterization fraction (LCF) were analysed by analysis of variance (ANOVA) to compare clinical performance of the combination of techniques to standard two-view digital mammography.

Results The 463 cases (breasts) reviewed included 258 with one to three lesions each, and 205 with no lesions. The 258 cases with lesions included 77 cancers in 68 breasts and 271 benign lesions to give a total of 348 proven lesions. The combination, $DBT_{(MLO)}+MX_{(CC)}$, was superior to MX (CC+MLO) in both lesion detection (LDF) and lesion characterization (LCF) overall and for benign lesions. $DBT_{(MLO)}+MX_{(CC)}$ was non-inferior to two-view MX for malignant lesions.

Conclusions This study shows that readers' capabilities in detecting and characterizing breast lesions are improved by combining single-view digital breast tomosynthesis and single-view mammography compared to two-view digital mammography.

Key Points

- *Digital breast tomosynthesis is becoming adopted as an adjunct to mammography (MX)*
- *$DBT_{(MLO)}+MX_{(CC)}$ is superior to $MX_{(CC+MLO)}$ in lesion detection (overall and benign lesions)*
- *$DBT_{(MLO)}+MX_{(CC)}$ is non-inferior to $MX_{(CC+MLO)}$ in cancer detection*
- *$DBT_{(MLO)}+MX_{(CC)}$ is superior to $MX_{(CC+MLO)}$ in lesion characterization (overall and benign lesions)*
- *$DBT_{(MLO)}+MX_{(CC)}$ is non-inferior to $MX_{(CC+MLO)}$ in characterization of malignant lesions*

G. Gennaro (✉) · E. Bezzon · L. Pescarini · I. Polico · A. Proietti · E. Baldan · F. Pomerri · P. C. Muzzio
Veneto Institute of Oncology (IRCCS), via Gattamelata, 64, 35128 Padua, Italy
e-mail: gisella.gennaro@ioveneto.it

R. E. Hendrick
Department of Radiology, School of Medicine, University of Colorado-Denver, 12700 E. 19th Avenue, Mail Stop C278, Aurora, CO 80045, USA

A. Toledano
Biostatistics Consulting, LLC, Kensington, MD 20895, USA

J. R. Paquelet
Advanced Medical Imaging Consultants, 2008 Caribou Drive, Fort Collins, CO 80525, USA

J. R. Paquelet
Breast Imaging, McKee Medical Center, 2000 Boise Ave., Loveland, CO 80538, USA

R. Chersevani
Private Medical Practice, 34010 Gorizia, Italy

C. di Maggio
Private Medical Practice, 35126 Padua, Italy

M. La Grassa
Department of Radiology, Oncological Reference Center (IRCCS), 33081 Aviano, Italy

Published online: 26 April 2013

Keywords Breast tomosynthesis · Mammography · Tomography · Per lesion analysis · Clinical performance

Abbreviations

ANOVA	analysis of variance
CC	cranio-caudal
DBT	digital breast tomosynthesis
$DBT_{(MLO)}+MX_{(CC)}$	combination of one-view (MLO) tomosynthesis and one-view (CC) mammography
FROC	free-response receiver operating characteristics
LCF	lesion characterization fraction
LDF	lesion detection fraction
MLO	medio-lateral oblique
MX	standard mammography
$MX_{(CC+MLO)}$	standard mammography in two views

Introduction

Mammography is an effective imaging technique for the detection of early-stage breast cancer [1, 2]. Mammography has significant limitations, however, due to masking of suspicious findings by fibroglandular densities, reducing its diagnostic performance, particularly in denser breasts. Another limitation of mammography is that some features interpreted as lesions are merely summations of superimposed tissues, leading to false positive interpretations [3–5].

Although several papers have demonstrated the multiple potential benefits of digital breast tomosynthesis (DBT) [6–14], several questions remain regarding its clinical application, particularly whether it should be used in screening, diagnosis, or both; and whether DBT should be used alone, replacing 2D mammography, or in combination with 2D mammography [10–14].

Digital breast tomosynthesis is a 3D technique, and, as such, could allow full coverage of breast tissue with a single view. This is why some earlier studies of DBT compared medio-lateral oblique (MLO)-only DBT views with two-view mammography. Another reason was the desire to keep DBT radiation doses at levels delivered by standard two-view mammography. Earlier studies using prototype equipment were designed to acquire DBT with the breast compressed in the MLO position at the same dose as two-view mammography [6–9]. As manufacturers move from prototype systems to clinical products and exposure control is optimized, it is becoming possible to acquire a DBT view at the same dose as a single mammography view. This opens the possibility of improving clinical performance by adding a second view of each breast, without an increase in radiation dose compared to two-view mammography [10–14].

The best way to demonstrate the clinical utility of a new imaging technique is by a properly designed reader study that independently compares radiologists' performance with the new technique to their performance with an accepted technique, in this case two-view mammography. Two main statistical methods are used to compare imaging investigations: the most common is receiver operating characteristic (ROC) curve analysis [15, 16], which considers interpretations per case (where a case can be a patient or a breast) and focuses on determining whether a case has a target condition (e.g. breast cancer); this type of analysis conventionally does not require spatial localization of findings and analysis is done on the basis of the most suspicious finding per case. Another method is per lesion analysis, which takes into account proper localization of findings and allows comparison of techniques in terms of individual lesion detection and characterization, including possible multiple lesions per case. Per lesion analysis requires image interpretation using a free-response approach, where readers are asked to mark all detected findings and subsequent analysis by a "scoring panel" verifies the correctness of each finding marked in terms of location and possibly also lesion type. Compared with ROC methods, the free-response paradigm has the advantage of closely resembling clinical practice in terms of basing accuracy on all proven lesions, not just the most suspicious lesion per case, and accounting for the correct localization of each identified lesion [17, 18]. Another advantage of per lesion analysis is that it permits separation of the two steps in the diagnostic process: (1) lesion detection, correctly noting the location of any breast finding, and (2) lesion characterization, a diagnostic decision concerning the degree of suspicion of each finding [19].

In this paper, the combination of a cranio-caudal (CC) digital mammography (MX) view with a single-view DBT (MLO) is compared to two-view digital mammography in terms of both lesion detection and characterization, using a per lesion approach. This new analysis technique is applied to a dataset obtained from a clinical performance study, the results of which were already published using standard ROC methods [14] that did not consider individual lesions.

Materials and methods

Study population and reference standard

This study was approved by the institutional ethics review board and by the Italian Ministry of Health. All participant patients gave written informed consent. From April 2007 to July 2008, we enrolled 250 diagnostic patients aged 40–83 years (mean 58 years) with at least one breast finding found by mammography and/or ultrasound and classified as probably benign, suspicious or highly

suspicious for malignancy. Both symptomatic and asymptomatic women with a breast finding were eligible for the study. Patients with prior mastectomy, breast size exceeding the digital detector size, breast implants, or high genetic risk were excluded. Besides the standard protocol applied to diagnostic patients (clinical breast examination, bilateral two-view digital mammography and bilateral breast ultrasound [US]), subjects enrolled in the study underwent an additional single-view (MLO) DBT of both breasts. DBT images were not included in the standard of care; they were used only in the retrospectively performed reader study. Truth was established using fine-needle aspiration cytology or histology of core or surgical biopsy for findings classified as suspicious or probably malignant, and a minimum of 1-year follow-up for unbiopsied (doubtful) findings and normal breasts.

Image dataset

Digital mammography (Senographe 2000D; GE Healthcare, Chalfont St Giles, UK) included at least the two standard views (CC and MLO) of both breasts acquired in a fully automatic exposure mode.

Single-view DBT examinations were performed using an investigational device developed by GE Healthcare on a Senographe DS platform. Fifteen low-dose projection images were acquired in a step-and-shoot mode, rotating the x-ray source from −20° to +20° around the axis perpendicular to the detector while the breast was compressed in the fixed MLO position. Exposure was controlled by manual selection of technique factors on the basis of compressed breast thickness, such that the total radiation dose for the MLO DBT sequence was approximately equal to the dose delivered in standard two-view mammography [20]. As mentioned in the "Introduction", the condition of dose equivalence between one-view DBT and two-view mammography used in this paper has moved toward dose equivalence of one-view DBT to one-view MX as manufacturers optimize tomosynthesis data acquisition. DBT projection views were reconstructed into parallel, contiguous, 0.5-mm-interval planar images using an iterative algorithm [21]. One-centimetre-thick slabs were also reconstructed from sets of adjacent planes to allow fast volume scroll and to aid detection of calcification clusters.

Reading protocol

Six radiologists with between 5 and 30 years of experience in breast imaging participated in the multireader study. Each had an initial training session that included 25 DBT cases to establish familiarity with DBT images and display software. Training in DBT interpretation was done for each reader independently by the principal investigator, who was not a study reader. Finally, all training cases were discussed in consensus, including all readers, to agree on the general methods of DBT evaluation.

After the training period, readings were organised in multiple sessions, and images were evaluated per breast (not per patient), to include normal cases (breasts) in the study population. Each reading session contained 50–70 cases, half mammography ($MX_{(CC+MLO)}$) and half combinations of MLO DBT and CC mammography ($DBT_{(MLO)}$+ $MX_{(CC)}$), randomised and presented in alternating order, using a dedicated workstation equipped with two high-resolution monitors. All cases were anonymized and no clinical information about the patients or the true diagnosis was made available to readers. A washout period of 2–3 months was maintained between readings of $MX_{(CC+MLO)}$ and $DBT_{(MLO)}$+$MX_{(CC)}$ for the same case.

Each observer's task was to detect and classify up to three findings per breast, benign or suspicious of malignancy, for each imaging technique. Each finding had to be specified by finding type and largest finding dimension, by location in both CC and MLO MX views or by location and preferred plane in MLO DBT views, and rated according to BIRADS [22]. BIRADS scores from 1 (negative) through 5 (highly suggestive of malignancy) were permitted, with BIRADS 4 subscored as 4A (low-suspicion abnormality), 4B (moderate-suspicion abnormality), or 4C (high-suspicion abnormality). A scoring panel comprising the principal investigator and two radiologists not involved in the readings (E.Ba., E.Be.) determined in consensus whether each finding was a true positive (TP) or false positive (FP) by cross-checking findings and the location descriptions (quadrant in 2D MX images, depth in DBT images) marked by each radiologist during the reading study with other documents used for truth establishment, such as image interpretation and biopsy reports, which included both histology information in cases of biopsied lesions, and follow-up imaging reports for unbiopsed lesions and normal breasts.

Statistical analysis

As mentioned above, per lesion analysis considers the scores for individual reader findings in a breast and allows separate analyses of the two clinical tasks involved in the diagnostic process: lesion detection and lesion characterization. For the task of lesion detection, each finding scored BIRADS 2 or higher and matching the location of a proven lesion (malignant or benign) was classified by the scoring panel as a "detected lesion". For each reader in each technique, we estimated the lesion detection fraction (LDF) as the number of detected lesions divided by the number of proven lesions.

In addition to detecting lesions, proper characterization is necessary to decide which lesions should be biopsied. For this task, each reader finding was scored as "correctly

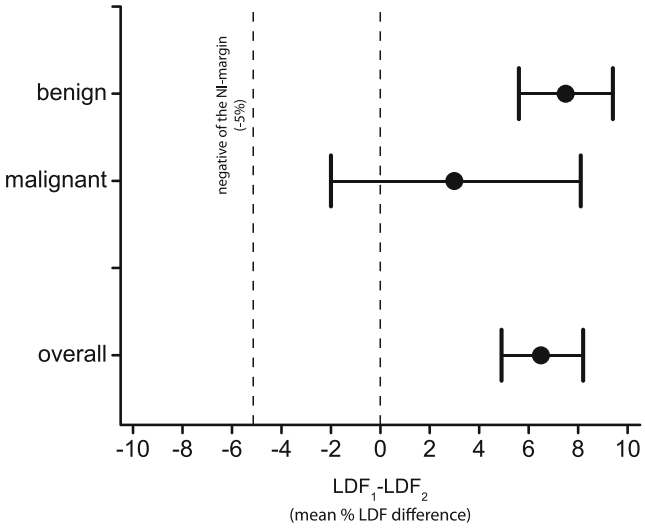

Fig. 1 Lesion detection. Difference between mean lesion detection fractions (LDF) with technique 1, the combination $DBT_{(MLO)}+MX_{(CC)}$, and technique 2, $MX_{(CC+MLO)}$, for all lesions correctly localized, cancers and benign lesions, and related 95 % confidence intervals. The zero difference line and negative of the non-inferiority margin (−5 %) are shown as dashed vertical lines. The mean LDF difference between the two techniques was +6.5 % (95 % CI +4.9 %, +8.2 %) overall, +3.1 % (95 % CI −2.0 %, 8.1 %) for malignant lesions, and +7.5 % (95 % CI +5.6 %, 9.4 %) for benign lesions, demonstrating the superiority of the combined technique for all lesions and all benign lesions, and the non-inferiority of the combined technique for cancers

classified" if it was detected, and if it was rated with an appropriate BIRADS score: 2 or 3 for benign lesions, or 4 or 5 for malignant lesions. Benign lesions rated BIRADS 4 or 5 were treated as false positives (FPs), whereas malignant lesions classified as BIRADS 2 or 3 were processed as false negatives (FNs). For each reader and for each imaging technique we estimated the lesion characterization fraction (LCF) as the total number of lesions correctly classified divided by the number of proven lesions. These same definitions (LDF and LCF) also were applied separately to the subsets of malignant and benign lesions.

The clinical performance of the combination of one-view DBT and one-view mammography ($DBT_{(MLO)}+MX_{(CC)}$) versus two-view mammography ($MX_{(CC+MLO)}$) was compared in a per lesion analysis, using both LDF and LCF.

Non-inferiority analysis was applied to LDF and LCF, using an Obuchowski–Rockette-type model [23], after adjusting the degrees of freedom according to Hillis's formula, to increase the accuracy of the lower 95 % confidence limit estimation [24]. The non-inferiority margin was set at 5 %. To conclude that technique 1 ($DBT_{(MLO)}+MX_{(CC)}$) was non-inferior to technique 2 ($MX_{(CC+MLO)}$), two conditions were necessary: (1) the difference between the means of LDFs (or LCFs) with technique 1 and technique 2 must have been above zero and (2) the lower limit of the confidence interval must have been above the negative of the non-inferiority margin [23, 25]. The null hypothesis (mean values of LDF or LCF for $DBT_{(MLO)}+MX_{(CC)}$ and $MX_{(CC+MLO)}$ are equal) was tested using analysis of variance according to the Obuchowski–Rockette model [23], with a Matlab routine (Matlab R2008a). One technique was regarded as superior to the other when the mean difference, as well as the entire 95 % confidence interval, was above zero. A P value less than 0.05 was considered statistically significant. The same type of analysis was repeated separately for the subsets of malignant and benign lesions.

Results

The reader study was performed on 463 breasts of 250 subjects. Four patients (8 breasts) were excluded from the study because MX images were missing, 10 patients (20 breasts) were excluded because MX examinations were acquired with MLO views only, 1 patient (2 breasts) was excluded because MX images were obtained by a computed radiography (CR) system, and 1 breast could not be used in the study because of a technical issue during DBT acquisition; finally, 6 breasts were retrospectively excluded because some finding localizations were uncertain. The 463 cases (breasts) reviewed included 258 cases with one to three lesions each, and 205 cases without any lesions. The 258 cases with lesions included 77 cancers (7 of them multifocal) in 68 breasts and 271 benign lesions to give a total of 348 proven lesions.

Figure 1 shows the difference between mean LDFs obtained with the two investigations for all 348 lesions, and separately for malignant and benign lesions. For the entire dataset (malignant and benign lesions combined), 66.5 % of lesions were detected by $DBT_{(MLO)}+MX_{(CC)}$ and 60.0 % by $MX_{(CC+MLO)}$, giving a +6.5 % difference (95 % CI +4.9 %, +8.2 %) that was statistically significant ($P<0.0001$). Cancer detection was comparable with the two protocols ($DBT_{(MLO)}+MX_{(CC)}$ = 78.6 %, $MX_{(CC+MLO)}$ =

Table 1 Lesion detection fraction for individual readers, including their individual experience with MX (in years)

Reader ID	Experience with MX (years)	Lesion detection fraction (LDF) (%)	
		$MX_{(CC+MLO)}$	$DBT_{(MLO)}+MX_{(CC)}$
1	32	52.9	59.8
2	36	56.0	61.8
3	26	53.7	60.9
4	32	69.3	73.6
5	7	70.4	77.0
6	8	57.8	66.1
Mean	23.5	60.0	66.5

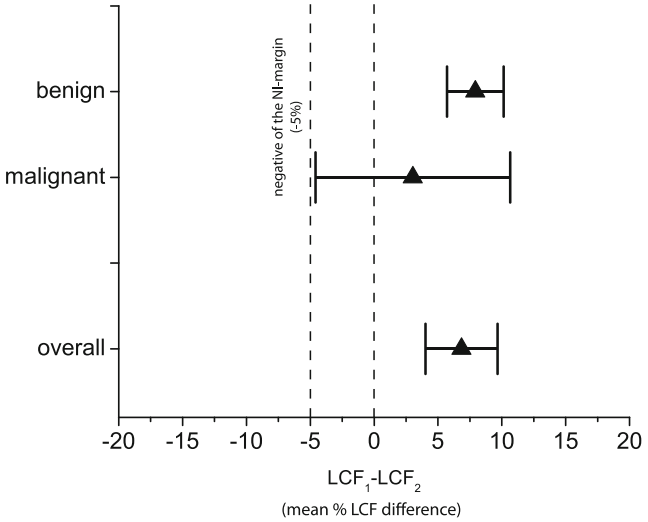

Fig. 2 Lesion characterization. Difference between mean lesion characterization fractions (LCF) with technique 1, the combination $DBT_{(MLO)}+MX_{(CC)}$, and technique 2, $MX_{(CC+MLO)}$, for all lesions correctly localized, cancers and benign lesions, and related 95 % confidence intervals. The zero difference line and negative of the non-inferiority margin (−5 %) are shown as dashed vertical lines. The mean LCF difference between the two techniques was +6.8 % (95 % CI +4.0 %, +9.7 %) overall, +3.1 % (95 % CI −4.6 %, 10.7 %) for malignant lesions, and +8.0 % (95 % CI +5.7 %, +10.2 %) for benign lesions, demonstrating the superiority of the combined technique for all lesions and all benign lesions, and the non-inferiority of the combined technique for cancers

75.5 %; difference = +3.1 %; 95 % CI −2.0 %, +8.1 %; P=0.198), whereas detection of benign lesions by the combined technique $DBT_{(MLO)}+MX_{(CC)}$ was significantly better ($DBT_{(MLO)}+MX_{(CC)}$ = 63.1 %, $MX_{(CC+MLO)}$ = 55.6 %; difference = +7.5 %; 95 % CI +5.6 %, +9.4 %; P<0.0001). Figure 1 shows the superiority of the combination of single-view DBT plus single-view MX versus two-view MX in lesion detection overall and for benign lesions, the 95 % confidence intervals for the differences being entirely above zero; and the non-inferiority of single-view DBT plus single-view MX versus two-view MX for malignant lesions, the 95 % confidence interval for the difference being entirely above the non-inferiority margin.

Table 1 provides LDF values for each reader, along with each reader's experience with interpreting mammograms.

Figure 2 illustrates the results for lesion characterization. Overall, the mean fraction of lesions correctly localized and rated (LCF) was 48.9 % with $DBT_{(MLO)}+MX_{(CC)}$ and 42.1 % with $MX_{(CC+MLO)}$, leading to a statistically significant +6.8 % difference (CI +4.0 %, +9.7 %; p=0.0007) in favour of the combined technique. The same analysis applied to malignant lesions showed that cancer classification was comparable with $DBT_{(MLO)}+MX_{(CC)}$ and $MX_{(CC+MLO)}$ ($DBT_{(MLO)}+MX_{(CC)}$ LCF = 71.9 %, $MX_{(CC+MLO)}$ LCF = 68.8 %; difference + 3.1 %; 95 % CI −4.6 %, +10.7 %; P=0.369). For benign lesions, mean LCF was found to be significantly higher with $DBT_{(MLO)}+MX_{(CC)}$, 42.4 % versus 34.4 % with $MX_{(CC+MLO)}$ (difference +8.0 %; 95 % CI +5.7 %, +10.2 %; P<0.0001). Table 2 provides LCF values for each reader.

These results show that the superiority of the combined technique in benign lesion detection (LDF) does not increase the false positive rate, because the lesion characterization performance (LCF) of $DBT_{(MLO)}+MX_{(CC)}$ is superior for benign lesions. The LCF mean differences demonstrate the superiority of DBT for lesion characterization overall and for benign lesions, along with the non-inferiority of $DBT_{(MLO)}+MX_{(CC)}$ compared to $MX_{(CC+MLO)}$ for malignant lesions.

Discussion

A new imaging technique provides true clinical benefit if it (1) shows more lesions than the imaging technique in routine clinical use, and (2) allows radiologists to make the "right decision" concerning malignancy of detected lesions. As discussed by Kopans, lesion detection is the first step in early breast cancer diagnosis based on imaging [19]; it is the recognition of a finding as potentially abnormal in a specific image location, before any clinical decision concerning probability of malignancy is made. It is useful to compare a new imaging technique to an imaging technique currently in routine clinical use in terms of lesion detection, because the lesion detection task is less influenced by individual readers' experience, capability, and confidence than lesion characterization. This study found that the combined technique, $DBT_{(MLO)}+MX_{(CC)}$, allowed the detection of more lesions than two-view mammography, $MX_{(CC+MLO)}$, as demonstrated by the significantly higher LDF of $DBT_{(MLO)}+MX_{(CC)}$. $DBT_{(MLO)}+MX_{(CC)}$ did not increase cancer detection significantly, but the proportion of benign lesions detected was significantly higher. A sample case is shown in Fig. 3, where a heterogeneously dense breast is shown imaged by MX (CC and MLO) in the left and middle images, and by

Table 2 Lesion characterization fraction for individual readers, including their individual experience with MX (in years)

Reader ID	Experience with MX (years)	Lesion characterization fraction (LCF) (%)	
		$MX_{(CC+MLO)}$	$DBT_{(MLO)}+MX_{(CC)}$
1	32	37.9	44.5
2	36	48.6	51.1
3	26	28.7	38.5
4	32	55.5	61.5
5	7	46.0	52.3
6	8	35.6	45.4
Mean	23.5	42.1	48.9

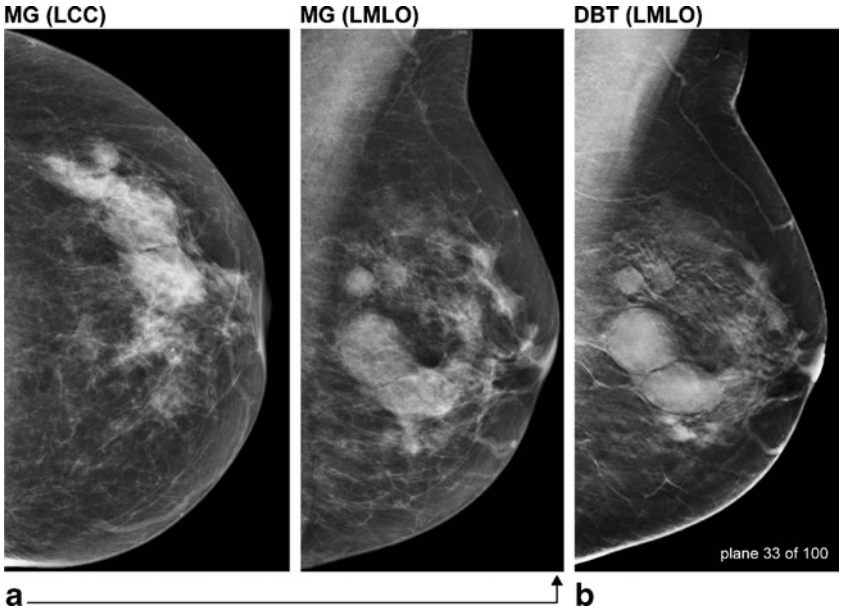

Fig. 3 Sample case showing DBT superiority in detection of benign lesions. **a** Left breast: two-view MX (CC and MLO). One small lesion is demonstrated, while two larger lesions are likely. **b** MLO DBT of the same breast, plane 33 of 100. At least four different benign lesions are demonstrated

DBT on the right. In MX, multiple lesions can be perceived, whereas in DBT, even in a single plane, four likely benign lesions can be clearly distinguished.

Whereas lesion detection is the first step of the diagnostic process, lesion characterization is a more important next step. As noted in the "Materials and methods" section, per lesion analysis has the benefit compared to standard ROC analysis of decomposing the diagnostic process into lesion detection and lesion characterization. Our findings indicate that single-view DBT plus single-view MX is superior to two-view MX in terms of correctly localizing and characterizing lesions, as demonstrated by the significantly higher overall LDF and LCF for $DBT_{(MLO)}+MX_{(CC)}$. This overall benefit is due to the superiority of $DBT_{(MLO)}+MX_{(CC)}$ in correctly localizing and characterizing benign lesions. This result is consistent with previous results obtained using per breast analysis, where the combined technique had higher specificity for benign lesions, whereas the fraction of cancers correctly rated with $DBT_{(MLO)}+MX_{(CC)}$ and $MX_{(CC+MLO)}$ was comparable. Our results also are consistent with what was found for sensitivity by ROC methods [14].

Two examples of this increase in diagnostic information are shown in Figs. 4 and 5, where cancers were better depicted by DBT (Fig. 4) and MX (Fig. 5). Figure 4 shows a bifocal cancer case where both cancers were seen (detected) in both MX and DBT, but spicules and tumour structure were better depicted (characterized) by DBT.

In Fig. 5, a malignant calcification cluster is shown. Margins and shapes of microcalcifications are better defined in MX. Microcalcifications appear brighter with DBT than with MX.

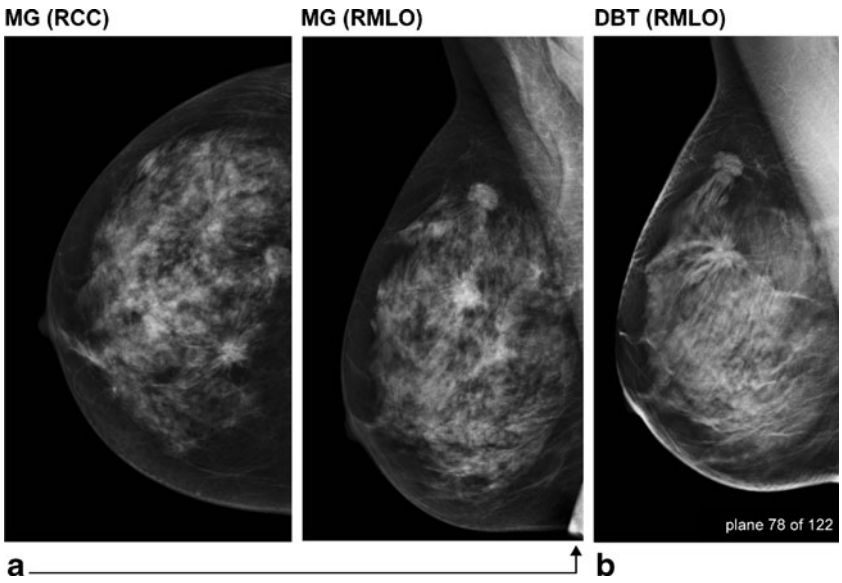

Fig. 4 Case showing better capability of DBT in depiction of spiculated masses. **a** Right breast: two-view MX (CC and MLO) shows two suspicious masses, one of them spiculated. **b** MLO DBT of the same breast (plane 78 of 122) shows both masses being spiculated, highly suggestive of malignancy

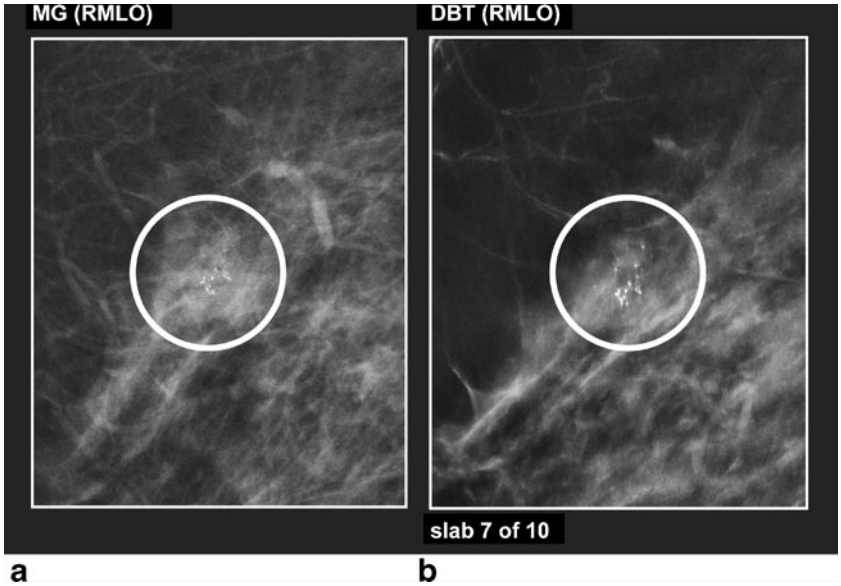

Fig. 5 Case showing a malignant calcification cluster in the upper-outer quadrant of the right breast. **a** Electronic magnification of the cluster in MLO MX view; the cluster was rated as suspicious by 6 out of 6 readers. **b** Electronically magnified view of a DBT slab shows the same cluster with high conspicuity; however, as calcifications appear coarser and denser, the lesion was rated as BIRADS 3 by 3 out of 6 readers, while being correctly classified as BIRADS 4 by the other 3 readers

In mammography, denser (or brighter) microcalcifications are more often interpreted as benign. In this study, some malignant lesions presenting as microcalcifications considered suspicious with MX (BIRADS 4A or higher) were scored BIRADS 3 with DBT because of their increased brightness, reducing the sensitivity of DBT relative to that of MX.

The ability of the combined technique, $DBT_{(MLO)}$+ $MX_{(CC)}$, to detect more lesions overall, and, in particular, benign lesions, might lead the reader to wonder if this means a potential increase in the false positive rate. The ability of $DBT_{(MLO)}$+$MX_{(CC)}$ to better characterize benign lesions, however, offers the potential to reduce unnecessary biopsies and better manage follow-up. Realization of this potential will depend on radiologists learning which image characteristics of cancers are different using DBT, requiring time and training. Once radiologists gain sufficient experience with DBT to appreciate differences in lesion appearance compared to conventional MX, the clinical performance of DBT could be further improved [26, 27].

There are some limitations to this study. Firstly, negative cases were taken from contralateral breasts of patients belonging to a diagnostic population and the reader study was conducted per breast rather than per patient; this prevented recognition of possible asymmetries between the two breasts of the same patient. This same limitation, however, was present for both techniques compared. Secondly, there was a selection bias in favour of mammography, because patients were recruited to the study on the basis of mammographically detected, not DBT-detected, lesions. Thirdly, readers' experience with mammography was extensive, but their DBT experience was limited to a previous study that compared clinical performance of single-view DBT versus standard mammography. Finally, the so-called laboratory effect, which affects any retrospective reader study, related to readers' awareness of participating in an experiment rather than clinical practice, might lead to results different from those obtained in a clinical context [28].

In conclusion, this study showed that readers' capabilities in detecting and characterizing breast lesions are improved with single-view digital breast tomosynthesis combined with single-view mammography compared to two-view digital mammography. On average, the combined technique is superior for both detection and characterization of all lesions and particularly of benign lesions, with comparable detection and characterization of cancers. The combined technique of MLO DBT and CC mammography might permit radiologists to transition more gradually from mammography to digital breast tomosynthesis, allowing time to get used to the differences and potential advantages of the new technique.

Acknowledgements The authors would like to thank Luc Katz, Aurora Talaverano, Francesca Braga, Henri Souchay, Razvan Iordache, Sylvain Bernard and Laura Hernandez from GE Healthcare for scientific discussions and technical support. They are also grateful to Andrea Azzalini for his help in preparation of manuscript illustrations.

R Edward Hendrick and Alicia Toledano are consultants to GE Healthcare.

This paper uses the same diagnostic subjects as another paper previously published in European Radiology, but applies per-lesion analysis rather than the more standard per-case analysis. Perlesion analysis, less popular than conventional ROC analysis, allows for the possibility of multiple lesions per case, and considers both lesion detection and lesion characterization. The per-lesion approach increases the statistical power of the analysis and allows us to better determine the potential diagnostic role of DBT in clinical application.

References

1. Elmore JG, Armstrong K, Lehman CD, Fletcher SW (2005) Screening for breast cancer. JAMA 293:1245–1256

2. Pisano ED, Gatsonis C, Hendrick E et al (2005) Digital Mammography Imaging Screening Trial (DMIST) Investigators Group: diagnostic performance of digital versus film mammography for breast-cancer screening. N Engl J Med 353:1773–1783
3. Carney PA, Miglioretti DL, Yankaskas BC et al (2003) Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med 138:168–175
4. Harvey JA, Bovbjerg VE (2004) Quantitative assessment of mammographic breast density: relationship with breast cancer risk. Radiology 230:29–41
5. Bird RE, Wallace TW, Yankaskas BC (1992) Analysis of cancer missed at screening mammography. Radiology 184:613–617
6. Andersson I, Ikeda DM, Zackrisson S et al (2008) Breast tomosynthesis and digital mammography: a comparison of breast cancer visibility and BIRADS classification in a population of cancers with subtle mammographic findings. Eur Radiol 18:2817–2825
7. Gur D, Abrams GS, Chough DM et al (2009) Digital breast tomosynthesis: observer performance study. AJR Am J Roentgenol 193:586–591
8. Teertstra HJ, Loo CE, van den Bosch MA et al (2012) Breast tomosynthesis in clinical practice: initial results. Eur Radiol 2010:16–24
9. Gennaro G, Toledano A, di Maggio C et al (2012) Digital breast tomosynthesis versus digital mammography: a clinical performance study. Eur Radiol 20:1545–1553
10. Wallis MG, Moa E, Zanca F, Leifland K, Danielsson M (2012) Two-view and single-view tomosynthesis versus full-field digital mammography: high resolution x-ray imaging observer study. Radiology 262:78–796
11. Skaane P, Gullien R, Bjørndal H et al (2012) Digital breast tomosynthesis (DBT): initial experience in a clinical setting. Acta Radiol 53:524–529
12. Michell MJ, Iqbal A, Wasan RK et al (2012) A comparison of the accuracy of film-screen mammography, full-field digital mammography, and digital breast tomosynthesis. Clin Radiol 67:976–981
13. Svahn TM, Chakraborty DP, Ikeda D, Zackrisson S, Do Y, Mattsson S, Andersson I (2012) Breast tomosynthesis and digital mammography: a comparison of diagnostic accuracy. Br J Radiol 85:e1074–82
14. Gennaro G, Hendrick RE, Ruppel P et al (2012) Performance comparison of single-view digital breast tomosynthesis plus single-view digital mammography with two-view digital mammography. Eur Radiol 23:664–72
15. Obuchowski NA (2005) Fundamentals of clinical research for radiologists: ROC analysis. AJR Am J Roentgenol 184:364–372
16. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29–36
17. Chakraborty DP, Berbaum KS (2004) Observer studies involving detection and localization: modelling, analysis and validation. Med Phys 31:2313–2330
18. Gur D, Bandos AI, Rockette HE et al (2011) Localized detection and classification of abnormalities on FFDM and tomosynthesis examinations rated under an FROC paradigm. AJR Am J Roentgenol 196:737–741
19. Kopans DB (2007) Breast imaging, 3rd edn. Lippincott Williams & Wilkins, Philadelphia, p 323–342
20. Wu T, Liu B, Moore R, Kopans D (2006) Optimal acquisition techniques for digital breast tomosynthesis screening. In: Flynn MJ, Hsieh J (eds) Medical imaging 2006: physics of medical imaging. Proc SPIE 6142:61425E
21. Wu T, Moore RH, Rafferty EA, Kopans DB (2004) A comparison of reconstruction algorithms for breast tomosynthesis. Med Phys 31:2636–2647
22. American College of Radiology (ACR) (2003) Breast Imaging Reporting and Data System (BIRADS), 4th edn. American College of Radiology, Reston
23. Obuchowski NA (1997) Testing for equivalence of diagnostic tests. AJR Am J Roentgenol 168:13–17
24. Hillis SL (2007) A comparison of denominator degrees of freedom methods for multiple observer ROC analysis. Stat Med 26:596–619
25. Chen W, Petrick NA, Sahiner B (2012) Hypothesis testing in noninferiority and equivalence MRMC ROC studies. Acad Radiol 19:1158–1165
26. Spangler ML, Zuley ML, Sumkin JH et al (2011) Detection and classification of calcifications on digital breast tomosynthesis and 2D digital mammography: a comparison. AJR Am J Roentgenol 196:320–324
27. Kopans DB, Gavenonis S, Halpern E, Moore R (2011) Calcifications in the breast and digital breast tomosynthesis. Breast J 17:638–664
28. Gur D, Bandos AI, Cohen CS et al (2008) The "laboratory" effect: comparing radiologists' performance and variability during prospective clinical and laboratory mammography interpretations. Radiology 249:47–53