Accuracy of GE digital breast tomosynthesis vs supplementary mammographic views for diagnosis of screen-detected soft-tissue breast lesions

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: CORNFORD, E. ... et al., 2016. Accuracy of GE digital breast tomosynthesis vs supplementary mammographic views for diagnosis of screen-detected soft-tissue breast lesions. British Journal of Radiology, 89:1058.

Additional Information:

- This paper was accepted for publication in the journal British Journal of Radiology and the definitive published version is available at http://dx.doi.org/10.1259/bjr.20150735

Metadata Record: https://dspace.lboro.ac.uk/2134/21476

Version: Published

Publisher: British Institute of Radiology (© the authors)

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
FULL PAPER

Accuracy of GE digital breast tomosynthesis vs supplementary mammographic views for diagnosis of screen-detected soft-tissue breast lesions

ELEANOR J CORNFORD, BM BS, ANNE E TURNBULL, BM BCh, JONATHAN J JAMES, BM BS, RACHEL TSANG, BSc, TAYEBA AKRAM, BSc, HELEN C BURRELL, BM BS, LISA J HAMILTON, BM BS, SARAH L TENNANT, BM BS, MARK J BAGNALL, MB BS, SHAMA PURI, MB BS, GRAHAM R BALL, BSc, PhD, YAN CHEN, PhD and VIVIENNE JONES BSc, MBA

Objectives: To compare the accuracy of standard supplementary views and GE digital breast tomosynthesis (DBT) for assessment of soft-tissue mammographic abnormalities.

Methods: Women recalled for further assessment of soft-tissue abnormalities were recruited and received standard supplementary views (typically spot compression views) and two-view GE DBT. The added value of DBT in the assessment process was determined by analysing data collected prospectively by radiologists working up the cases. Following anonymization of cases, there was also a retrospective multireader review. The readers first read bilateral standard two-view digital mammography (DM) together with the supplementary mammographic views and gave a combined score for suspicion of malignancy on a five-point scale. The same readers then read bilateral standard two-view DM together with two-view DBT. Pathology data were obtained. Differences were assessed using receiver operating characteristic analysis.

Results: The study population was 342 lesions in 322 patients. The final diagnosis was malignant in 113 cases (33%) and benign/normal in 229 cases (67%). In the prospective analysis, the performance of two-view DM plus DBT was at least equivalent to the performance of two-view DM and standard mammographic supplementary views—the area under the curve (AUC) was 0.946 and 0.922, respectively, which did not reach statistical significance. Similar results were obtained for the retrospective review—AUC was 0.900 (DBT) and 0.873 (supplementary views), which did not reach statistical significance.

Conclusion: The accuracy of GE DBT in the assessment of screen-detected soft-tissue abnormalities is equivalent to the use of standard supplementary mammographic views.

Advances in knowledge: The vast majority of evidence relating to the use of DBT has been gathered from research using Hologic equipment. This study provides evidence for the use of the commercially available GE DBT system demonstrating that it is at least equivalent to supplementary mammographic views in the assessment of soft-tissue screen-detected abnormalities.

INTRODUCTION

Supplementary mammographic views have traditionally been a crucial part of the assessment process in women recalled following an abnormal screening mammogram. These might include spot compression views (SCVs), lateral projections and extended or exaggerated cranio-caudal (CC) views. Their purpose is to confirm the presence and characterize a true breast lesion rather than a pseudoabnormality simulated by summation of normal breast tissue.

Digital breast tomosynthesis (DBT) is an advanced form of digital mammography (DM), providing three-dimensional-like reconstructed images of the breast tissue. This has the potential to reduce the tissue overlap effect thereby improving the sensitivity and specificity of a screening programme, by improving margin characterization and eliminating summation effects. The improved specificity offered by DBT can also be utilized in the assessment of screen-detected abnormalities. Several studies have already suggested that DBT can replace the traditional
supplementary mammographic views in the assessment of soft-tissue breast lesions, with performance at least equivalent.\textsuperscript{2–7}

DBT is currently available from multiple vendors. Each system differs in design, with significant variations in parameters such as tube motion, angle range, number of projections and processing algorithms. Each of these features has the potential to affect clinical performance.\textsuperscript{8} The majority of published DBT research has been conducted on Hologic equipment (Bedford, MA). Concern has been raised that evidence produced from studies using one system cannot necessarily be extrapolated to other manufacturers’ devices and that further research using other vendors’ devices has been suggested before other manufacturer DBT systems are introduced into routine clinical practice replacing established assessment protocols.

The aim of this study was to assess whether the performance of GE Healthcare (Buc, France) DBT (GE DBT) is equivalent to standard supplementary mammographic views and can potentially replace their use in the assessment of screen-detected soft-tissue abnormalities.

METHODS AND MATERIALS
This was a dual-centre paired comparison study involving a prospective component assessing the clinical utility of GE DBT as an adjunct tool in screening assessment and a retrospective multireader component comparing diagnostic performance of GE DBT vs standard supplementary views in the assessment of soft-tissue screen-detected abnormalities. The study was approved by the local National Health Service (NHS) Research Ethics Committee, received support from the National Cancer Research Network and was funded by GE Healthcare. GE Healthcare had no control over study design, data collection, analysis or presentation of the results.

The study group comprised women requiring further mammographic work-up of screen-detected soft-tissue lesions at two NHS breast screening sites between 1 February and 30 September 2014. Prior to participation, written consent was obtained from all women. The standard supplementary views obtained were usually SCVs in standard mediolateral oblique (MLO) and CC planes. Occasionally, extended CC views, lateral projections or magnification views were obtained if felt clinically indicated. DBT views were obtained in MLO and CC projections of the breast being assessed. DBT was performed on a commercially available GE DBT system (SenoClaire®). The GE system uses a “step and shoot” acquisition with nine exposures over a 25° gantry angle and uses iterative reconstruction techniques. Supplementary views were performed on either a GE Essential or GE DS full-field DM machine. Images were reviewed and interpreted on a GE Image Diagnostic International workstation.

For the prospective part of the study, standard supplementary views and DBT images were reviewed by two consultant radiologists conducting the clinic. It is our standard practice for two radiologists to review supplementary imaging performed in the work-up of screen-detected abnormalities to reduce the likelihood of false-negative assessment. A consensus imaging score

Figure 1. Recruitment and assessment process. ALH, atypical lobular hyperplasia; DBT, digital breast tomosynthesis; DCIS, ductal carcinoma in situ.
was prospectively assigned for both modalities prior to ultrasound being performed. Typically, the standard supplementary views were reviewed and scored first followed by the DBT images. The imaging score used was the UK Royal College of Radiologists 1–5 score.9 Decisions on the need for biopsy were made on a case-by-case basis, but if the supplementary mammographic views were deemed abnormal and DBT normal, then biopsy was still attempted as the supplementary mammographic views remained the standard of care. However, if supplementary mammographic views were deemed normal but DBT showed a persistent or additional abnormality, then management was changed accordingly. Data were also prospectively recorded on size and multifocality for both modalities. The radiologists working up the cases were also asked to rate the usefulness of the DBT images in their decision-making.

All images were then anonymized and retrospectively reviewed by one of eight specialist breast radiologists not involved in the initial assessment. The radiologists’ breast imaging experience ranged from 7 to 24 years (median 15 years). All radiologists fulfilled the NHS Breast Screening Programme (NHSBSP) Quality Assurance Criteria for screening mammography film reading and assessment and have undergone specialist training in interpretation of DBT images.10 As a minimum, this involved a 1-day training course involving review of 80 cases with both two-dimensional (2D) full-field DM and DBT. The first review involved analysis of the two-view digital screening mammogram with the supplementary assessment views (usually SCV). The second review took place at least 3 weeks later and involved analysis of the two-view DM with DBT. The radiologist undertaking the retrospective review was blinded to the assessment outcome. For both of the retrospective reviews, an imaging score was assigned using the UK Royal College of Radiologists 1–5 score.9 Histopathology provided the ground truth outcome for those subject to biopsy.

For both components of the study, performance of the supplementary mammographic views and DBT was determined by receiver operating characteristic (ROC) curves. Areas under the curve (AUCs) were compared using the method of DeLong et al.11 A χ² test was used to determine significance in sensitivity, specificity and positive-predictive value (PPV) between the two modalities. A p-value of <0.05 was regarded as statistically significant.

RESULTS

324 women were recruited to take part in the study. 344 lesions were assessed in 324 patients. Two patients were later excluded because they were diagnosed with non-breast malignancies (one aggressive fibromatosis and one primary lymphoma). The study population therefore consisted of 342 lesions in 322 patients. The predominant mammographic abnormalities undergoing assessment were masses (50%), distortions (13.5%) and asymmetric densities (36.5%). Microcalcification was an associated feature in 3.8%, but patients had not been invited to participate when microcalcification was the predominant mammographic feature. Additional mammographic views were performed in all cases. Occasionally, SCVs needed to be repeated because of the lesion being missed. Repeat SCVs were requested by the assessing radiologists in 49 of the 322 patients (15.2%). There were no instances where DBT was

Table 1. Final pathology and imaging scores prospectively assigned to each soft tissue abnormality during the assessment work-up following screening recall

<table>
<thead>
<tr>
<th>Final pathology after full assessment</th>
<th>Images read</th>
<th>Imaging score (n = 342)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2D plus SCV</td>
<td>M1/2  M3  M4  M5</td>
</tr>
<tr>
<td>Normal/Benign</td>
<td>2D plus DBT</td>
<td>180  42  6  1</td>
</tr>
<tr>
<td>Malignant</td>
<td>2D plus SCV</td>
<td>11   17  35  50</td>
</tr>
<tr>
<td></td>
<td>2D plus DBT</td>
<td>7    16  25  65</td>
</tr>
</tbody>
</table>

2D, standard two-dimensional digital mammography; DBT, digital breast tomosynthesis; SCV, spot compression views and other supplementary mammographic views.

Table 2. Accuracy of digital breast tomosynthesis (DBT) and supplementary mammographic views in the assessment of soft-tissue abnormalities from the imaging scores prospectively assigned to each lesion during the assessment work-up following screening recall

<table>
<thead>
<tr>
<th>Measured statistic</th>
<th>2D and SCV (%)</th>
<th>2D and DBT (%)</th>
<th>χ² significance (two-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute sensitivity (M5)</td>
<td>44.24</td>
<td>57.52</td>
<td>0.045</td>
</tr>
<tr>
<td>Complete sensitivity (M3,4,5)</td>
<td>90.26</td>
<td>93.80</td>
<td>0.326</td>
</tr>
<tr>
<td>Specificity</td>
<td>78.60</td>
<td>82.53</td>
<td>0.288</td>
</tr>
<tr>
<td>M5 PPV</td>
<td>98.03</td>
<td>97.01</td>
<td>0.726</td>
</tr>
<tr>
<td>M1/2 NPV</td>
<td>94.24</td>
<td>96.42</td>
<td>0.307</td>
</tr>
</tbody>
</table>

2D, standard two-dimensional digital mammography; NPV, negative-predictive value; PPV, positive-predictive value; SCV, spot compression views and other supplementary mammographic views.
repeated. Consequently, the mean number of additional views was 2.3 (range 1–5). For DBT, the average mean glandular dose for an MLO of a 50–60 mm breast was 1.50 mGy for the two centres.

All lesions underwent an ultrasound scan, with any biopsies performed under ultrasound or stereotactic guidance as appropriate.

Figure 1 shows a flow diagram of the trial. The final diagnosis was malignant in 113 lesions (33%) and benign or normal in 229 lesions (67%). Of the malignant lesions, 107 (94.7%) were invasive cancers and 6 (5.3%) were ductal carcinoma in situ. Of the 229 lesions with a final outcome of benign or normal, 54 patients (23.5%) had a biopsy at assessment. 44 of these had a benign biopsy result and were discharged back to normal screening. The other 10 patients required a diagnostic excision of the lesion—7 of these patients had a final pathological diagnosis of radial scar. There has been 1 interval cancer in the group of 229 patients with a normal/benign assessment outcome which was diagnosed 12 months after the assessment process.

Prospective analysis
There were three patients where DBT demonstrated multifocal disease thought to be unifocal on the screening mammogram and supplementary mammographic views. These additional lesions are included in the total of 342 lesions assessed in the 322 patients. There were no cases where the supplementary mammographic views demonstrated multifocal disease thought to be unifocal on DBT.

The imaging scores assigned to each lesion prospectively following supplementary views and DBT during the assessment work-up are shown in Table 1, with sensitivities, specificity, PPV and negative-predictive value (NPV) for each technique shown in Table 2. DBT demonstrated significantly higher absolute sensitivity than standard supplementary mammographic views \((p = 0.045)\). There was no significant difference between the overall performance of standard supplementary views and DBT as determined by ROC analysis with the AUC being 0.922 [95% confidence interval (CI) 0.889–0.948] and 0.946 (95% CI 0.917–0.968), respectively (Figure 2).

The radiologists working up the cases found DBT to be useful in 56.9% of cases, the main reason quoted was improved margin delineation (54.7%) with the second most common reason being that it helped confirm the benign nature of lesions (35.3%).

Retrospective analysis
During the retrospective review, an additional nine lesions were documented by the readers, but all were categorized as benign. The retrospective review generated no additional malignant lesions. A total of 351 lesions were documented in the 322 patients, of which the final diagnosis was malignant in 113 (32.2%) lesions and benign or normal in 238 (67.8%) lesions. The imaging scores assigned to each lesion during the retrospective reader review of the supplementary views and DBT are shown in Table 3 with sensitivities, specificity, and NPV and PPV shown in Table 4. There was no significant difference between the overall performance of the standard supplementary views and DBT as determined by ROC analysis with the AUC being 0.873 (95% CI 0.834–0.906) and 0.900 (95% CI 0.864–0.929), respectively (Figure 3).

Table 3. Final pathology and imaging scores assigned to each soft-tissue abnormality during the retrospective reader review

<table>
<thead>
<tr>
<th>Final pathology after full assessment</th>
<th>Images read</th>
<th>Imaging score ((n = 351))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2D plus SCV</td>
<td>M1/2 M3 M4 M5</td>
</tr>
<tr>
<td>Normal/Benign</td>
<td>2D plus SCV</td>
<td>168 60 8 2</td>
</tr>
<tr>
<td></td>
<td>2D plus DBT</td>
<td>172 54 8 4</td>
</tr>
<tr>
<td>Malignant</td>
<td>2D plus SCV</td>
<td>9 23 41 40</td>
</tr>
<tr>
<td></td>
<td>2D plus DBT</td>
<td>7 22 30 54</td>
</tr>
</tbody>
</table>

2D, standard two-dimensional digital mammography; DBT, digital breast tomosynthesis; SCV, spot compression views and other supplementary mammographic views.
DISCUSSION
X-ray mammography is our current primary diagnostic tool for detection of breast cancer, however, it has limitations in terms of both sensitivity and specificity. Its poor specificity means that approximately 75% of women recalled following screening mammography in the UK will not have cancer but will be subjected to further tests including biopsy. Further imaging investigation has traditionally involved supplementary mammographic views to investigate potential masses, asymmetries and distortions, most often SCVs performed in MLO and CC projections. The results of this study show that GE DBT is at least equivalent to the use of these supplementary mammographic views in the assessment of screen-detected soft-tissue abnormalities with improved absolute sensitivity in the prospective evaluation.

Others have also demonstrated equivalent or improved accuracy of DBT compared with supplementary mammographic views in the assessment process. It has recently been shown that when DBT is introduced as the primary screening test, significantly fewer recalled women required supplementary mammographic views as part of the assessment work-up. There have also been two recently published overviews of current evidence on DBT. Houssami and Skaane concluded that the addition of DBT to standard mammography for mammographic interpretation or for assessment or triage of screen-called abnormalities increases accuracy. Alakhras et al concluded that DBT has shown promise in comparison with additional 2D views, demonstrating the same or higher diagnostic accuracy than digital SCVs. This body of evidence has led to changes in protocols with a switch from SCVs to DBT in the further evaluation of mammographic abnormalities found at screening and in symptomatic diagnostic practice.

The vast majority of the research to date on DBT has been performed on Hologic DBT equipment and very little literature exists on other commercially available systems. Concern has been raised that evidence acquired from one vendor’s device cannot necessarily be extrapolated to other manufacturers, and more information is needed on the effectiveness of non-Hologic DBT systems. In the UK breast-screening programme, the recommendation that DBT can be used to replace SCVs only extends to Hologic equipment.

The current commercially available DBT systems differ in design with significant variations in parameters such as tube motion, angle range, number of projections and processing algorithms. Each of these features has the potential to affect clinical performance. Hologic DBT uses a pulsed acquisition during a continuous sweep of the detector with 15 exposures over a 15° angle range and uses a filtered back projection algorithm for image processing. The GE system uses a “step and shoot” acquisition with 9 exposures over a 25° gantry angle and uses iterative reconstruction techniques. The advantages of these different approaches are open to debate but can affect image quality. A wider tube angle provides better depth resolution, whereas a narrow one improves in-plane resolution. Iterative reconstruction is said to improve visualization of masses and soft-tissue margins, whereas filtered back projection can enhance the conspicuity of microcalcifications.

Table 4. Accuracy of digital breast tomosynthesis (DBT) and supplementary mammographic views in the assessment of soft-tissue abnormalities from the imaging scores assigned during the retrospective reader review

<table>
<thead>
<tr>
<th>Measured statistic</th>
<th>2D and SCV (%)</th>
<th>2D and DBT (%)</th>
<th>$\chi^2$ significance (two-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute sensitivity (M5)</td>
<td>35.39</td>
<td>47.78</td>
<td>0.059</td>
</tr>
<tr>
<td>Complete sensitivity (M3,4,5)</td>
<td>92.03</td>
<td>93.80</td>
<td>0.604</td>
</tr>
<tr>
<td>Specificity</td>
<td>70.58</td>
<td>72.26</td>
<td>0.685</td>
</tr>
<tr>
<td>M5 PPV</td>
<td>95.23</td>
<td>93.10</td>
<td>0.657</td>
</tr>
<tr>
<td>M1/2 NPV</td>
<td>96.90</td>
<td>96.08</td>
<td>0.708</td>
</tr>
</tbody>
</table>

2D, standard two-dimensional digital mammography; NPV, negative-predictive value; PPV, positive-predictive value; SCV, spot compression views and other supplementary mammographic views.
has produced detailed publications on the technical aspects of DBT and some of the main differences between different manufacturers’ equipment, but detailed discussion is beyond the scope of this article. This study contributes to the evidence that the performance of the commercially available GE system performance is comparable to that seen with Hologic equipment in the evaluation of screen-detected soft-tissue abnormalities.

We did not use DBT to assess any patients with microlcalfication because previous studies have found that the improved accuracy for DBT in the work-up of screen-detected abnormalities only applied to soft-tissue abnormalities. The detection and characterization of microlcalfications with DBT is a more controversial area. Kopans et al reported equal or better visualization of calcifications with DBT, whereas others have reported improved sensitivity of standard 2D DM compared with DBT. Morel et al found no difference in the performance of DBT compared with magnification mammography in the work-up of screen-detected microcalcifications. In retrospect, it may have been useful to include some calcified lesions in this study, as there may be a manufacturer-specific difference.

The dose of GE DBT is only marginally more than standard 2D DM. For DBT, the average mean glandular dose for an MLO of a 50–60 mm breast was 1.50 mGy for the two centres. The corresponding doses for standard 2D mammography on these systems are 1.43 mGy (Derby) and 1.14 mGy (Nottingham), which fall well below the national diagnostic reference level for mammography of 3.5 mGy mean glandular dose to a lateral oblique view of a 55-mm compressed breast. Performing SCVs can be technically challenging for the radiographer, and repeats are not an uncommon occurrence, whereas DBT uses standard patient positioning and so is technically easier to perform. During the study, repeat SCVs were requested in 15.2% of assessment work-ups because of the lesion having been missed, adding to the radiation dose, and potentially increasing assessment time and patient anxiety. The advantage of DBT is that it provides additional information of the whole of the imaged breast, avoiding repeat exposures. When the repeat views are taken into account, there is a potential dose saving by replacing supplementary mammographic views with DBT. The ability to image the whole of the breast with DBT also has the advantage of identifying unsuspected multifocal disease as occurred in three patients in this study.

The utility of DBT is not just down to its accuracy. The radiologists working up the cases found the additional information provided by DBT in lesion work-up helpful in the decision-making process over and above that obtained from the supplementary mammographic views. The assessing radiologists found DBT to be useful in 56.9% of assessment cases with the main reason quoted being improved margin delineation (54.7%) and the second most common reason being that it helped confirm the benign nature of lesions (35.3%). Other potential advantages of DBT over supplementary mammographic views are improved lesion localization in three dimensions, aiding ultrasound and biopsy targeting.

There are some potential problems with the study. The prospective review can be criticized for bias as the assessing radiologists were not blinded to the results of DBT and supplementary mammographic views during lesion work-up. By contrast, the readers in the retrospective review had no knowledge of the assessment outcome, and reading of DBT and supplementary mammographic views was separated by at least 3 weeks. Another problem of the study is the relatively short follow-up time for the benign and normal cases, but because of the paired nature of the study, meaningful conclusions can still be drawn. The study was designed to compare the performance of GE DBT to supplementary mammographic views in the assessment of screen-detected soft-tissue abnormalities, and we have shown equivalence in the characterization and diagnosis of soft-tissue abnormalities. Further research is still needed, with GE DBT to confirm the improved sensitivity and specificity shown with Hologic DBT in the detection of malignancy.

In conclusion, GE DBT is at least equivalent to supplementary mammographic views in the assessment of soft-tissue screen-detected abnormalities. The vast majority of evidence relating to the use of DBT has been gathered from research using Hologic equipment. A major strength of this study is that it provides evidence for the use of the commercially available GE DBT system. The findings can be used to support a change in the assessment protocols allowing the replacement of supplementary mammographic views with GE DBT for the assessment of screen detected soft-tissue abnormalities.

ACKNOWLEDGMENTS
Nottingham University Hospital acknowledges the support of the National Institute for Health Research, through the CRN East Midlands. We acknowledge Dr M Michell and team at Kings Hospital, London for sharing their protocol. Nottingham University Hospitals sponsored this study.

FUNDING
Funding was provided by GE Health Care, however, GE had no involvement in study design, data analysis or manuscript preparation.

REFERENCES


