Microcalcifications Detected at Screening Mammography: Synthetic Mammography and Digital Breast Tomosynthesis versus Digital Mammography

Yi-Chen Lai, MD  •  Kimberly M. Ray, MD  •  Amie Y. Lee, MD  •  Jessica H. Hayward, MD  •  Rita I. Freimanis, MD  •  Iryna V. Lobach, PhD  •  Bonnie N. Joe, MD, PhD

From the Department of Radiology, Taipei Veterans General Hospital, Taipei, Taiwan (Y.C.L.); School of Medicine, National Yang-Ming University, Taipei, Taiwan (Y.C.L.); Department of Radiology and Biomedical Imaging, University of California San Francisco, 1600 Divisadero St, Box 1667, Room C250, San Francisco, CA 94115 (A.Y.L., J.H.H., R.I.F., I.V.L., B.N.J.); and Department of Radiology, The Permanente Medical Group, 3600 Broadway, Oakland, CA 94611 (K.M.R.). Received May 16, 2018; revision requested July 2; revision received August 9; accepted August 10.

Conflicts of interest are listed at the end of this article.

May 16, 2018; revision requested July 2; revision received August 9; accepted August 10.

In breast cancer screening, digital breast tomosynthesis (DBT) in combination with full-field digital mammography (FFDM) reduces recall rates and increases invasive cancer detection rates relative to FFDM alone (1–9). However, DBT does not eliminate the need for FFDM, primarily because two-dimensional planar images facilitate the detection of grouped microcalcifications, which may be overlooked at DBT since particles are dispersed over multiple thin-section images. Two-dimensional images also facilitate comparison with prior mammograms and evaluation of symmetry between the breasts. However, performing DBT in addition to FFDM approximately doubles the radiation dose relative to FFDM or DBT alone (10,11).

Synthetic mammography (SM) refers to virtual two-dimensional mammography images that are created from the DBT acquisition. SM is generated by summing and filtering the reconstructed DBT sections. This produces an image that is more similar in appearance to the conventional two-dimensional FFDM image. A more detailed description of the methods of SM image generation is provided elsewhere (3,12–14).

SM can reduce the radiation dose at screening by approximately 45% by eliminating the need for a separate FFDM acquisition (10,11). In addition, substituting SM shortens the acquisition time, which reduces the potential for image blurring due to patient motion and reduces time in compression, thus improving the patient experience.

Although it would be desirable to substitute SM for FFDM when DBT is used at screening, it is critical to first establish that the clinical performance of SM plus DBT is noninferior to FFDM. Although SM has been reported to improve the conspicuity of noncalcified findings such as architectural distortion and mass spiculation relative to FFDM (15), it is uncertain whether SM is equivalent to FFDM for the detection of microcalcifications. There are several potential limitations of SM with respect to microcalcifications, which directly relate to the source DBT acquisition. In DBT systems that use pixel binning, the resolution of SM is lower than that of FFDM (10,14).

Purpose: To compare the performance of two-dimensional synthetic mammography (SM) plus digital breast tomosynthesis (DBT) versus conventional full-field digital mammography (FFDM) in the detection of microcalcifications on screening mammograms.

Materials and Methods: In this retrospective multireader observer study, 72 consecutive screening mammograms recalled for microcalcifications from June 2015 through August 2016 were evaluated with both FFDM and DBT. The data set included 54 mammograms with benign microcalcifications and 18 mammograms with malignant microcalcifications, and 20 additional screening mammograms without microcalcifications used as controls. FFDM alone was compared to synthetic mammography plus DBT.

Four readers independently reviewed each data set and microcalcification recalls were tabulated. Sensitivity and specificity for microcalcification detection were calculated for SM plus DBT and for FFDM alone. Interreader agreement was calculated with Fleiss kappa values.

Results: Reader agreement was kappa value of 0.66 (P < .001) for FFDM and 0.63 (P < .001) for SM plus DBT. For FFDM, the combined reader sensitivity for all microcalcifications was 80% (229 of 288; 95% confidence interval [CI]: 74%, 84%) and for malignant microcalcifications was 92% (66 of 72; 95% CI: 83%, 97%). For SM plus DBT, the combined reader sensitivity for all microcalcifications was 75% (215 of 288; 95% CI: 69%, 80%) and for malignant microcalcifications was 94% (68 of 72; 95% CI: 86%, 98%). For FFDM, the combined reader specificity for all microcalcifications was 98% (78 of 80; 95% CI: 91%, 100%) and for malignant microcalcifications was 98% (78 of 80; 95% CI: 91%, 100%). For SM plus DBT, combined reader specificity for all microcalcifications was 95% (76 of 80; 95% CI: 88%, 99%) and for malignant microcalcifications was 95% (76 of 80; 95% CI: 88%, 99%). Mixed-effects model concluded no differences between modalities (–0.03; 95% CI: −0.08, 0.01; P = .13).

Conclusion: Relative to full-field digital mammography, synthetic mammography plus digital breast tomosynthesis had similar sensitivity and specificity for the detection of microcalcifications previously identified for recall at screening mammography.

© RSNA, 2018
Continuous motion of the x-ray tube in some DBT systems results in blurring of the focal spot, which also reduces effective resolution (14). Patient motion during the DBT acquisition may reduce reconstructed image sharpness (16). Finally, DBT reconstruction artifacts may distort the appearance of microcalcifications (16–18) (Fig 1).

On the basis of these technical limitations, phantom studies show inferior sensitivity of SM for fine specks compared with FFDM (19). We hypothesized that SM with DBT would be inferior to FFDM for the detection of microcalcifications, given the technical limitations of SM. However, the impact of these technical limitations on the clinical performance of SM is less clear since the goal of screening mammography is to detect rather than to fully characterize potential abnormalities. Therefore, even if microcalcifications are less well visualized at SM compared with FFDM, the technique may nevertheless be sufficient for the detection of clinically significant malignant microcalcifications, which would then be fully characterized in form and extent at diagnostic mammography using two-dimensional spot compression magnification mammography. The purpose of this study was to compare the performance of two-dimensional SM plus DBT versus conventional FFDM in the detection of microcalcifications on screening mammograms.

**Materials and Methods**

This retrospective study was compliant with the Health Insurance Portability and Accountability Act and approved by the institutional review board with waiver of need for informed consent. No industry support was received for this study and the authors had control of all the data used for this study.

**Mammogram Selection**

From June 2015 through August 2016, a total of 3858 screening mammograms were obtained at our institution with FFDM, DBT, and two-dimensional SM. All mammograms were interpreted in the course of routine clinical care by one of 12 Mammography Quality Standards Act and Program–certified breast imaging radiologists with 2 to more than 30 years of experience. Seventy-two of the 3858 screening mammograms were recalled (ie, assessed as Breast Imaging Reporting and Data System [BI-RADS] 0: incomplete) for microcalcifications; 54 had benign microcalcifications and 18 had malignant microcalcifications. All patients with malignant microcalcifications underwent either vacuum-assisted core biopsy and/or surgical excision for diagnosis. Patients with benign lesions either underwent vacuum-assisted core biopsy or demonstrated at least 2 years of stability when comparison mammograms were subsequently reviewed. As per standard breast imaging auditing practice, atypia at pathologic examination was considered benign unless it was upgraded to cancer based on surgical excision.

Twenty randomly selected screening mammograms obtained during the study period and found to have no microcalcifications (assessed as either BI-RADS 2: benign or BI-RADS 1: negative) were added as normal controls. The choice of 20 controls was a number of convenience. A radiologist (K.R.) with 10 years of breast imaging experience who was not one of the readers confirmed the absence of microcalcifications in the mammogram control group.

A total of 92 mammograms comprised our final study data set. Mean patient age was 59 years (range, 41–75 years). Characteristics of mammography cases are summarized in the Table.

**Imaging Technique and Data Set**

All screening mammograms were acquired on a commercial system (Selenia Dimensions, Hologic, Marlborough, Mass), which included FFDM and DBT in craniocaudal and mediolateral oblique projections during a single compression of the breast for each projection (COMBO mode). SM was reconstructed from the DBT images by using Food and Drug Administration (FDA)-approved, commercially available software (C-view for Hologic, version 1, Selenia Dimensions; Hologic, Marlborough, Mass). Two anonymized data sets, each consisting of 92 mammograms, were created for blinded interpretation: one data set consisted of FFDM alone and the other consisted of SM plus DBT. Each screening mammogram was represented in both anonymized data sets.

**Reader Study**

Four fellowship-trained breast imagers at an academic medical center served as readers in our study (A.L., B.J., J.H., and R.F., with 2, 10, 4, and 25 years of breast imaging experience, respectively). Readers were blinded to patient information, the clinical mammography report, and the outcomes of each mammogram. Prior comparison mammograms were not provided as most of the comparison examinations were FFDM and we wished to avoid potential bias related to providing FFDM comparisons when reading SM plus DBT data sets. Readers were instructed to identify mammograms with microcalcifications warranting recall (for multiple groups of microcalcifications, the reader identified the most suspicious group), and the reader’s decision whether to recall each mammogram with microcalcifications was recorded by a research assistant. In addition, the breast laterality and quadrant location of each microcalcification finding warranting recall was recorded. Readers were instructed to ignore nonmicrocalcification findings.

In reviewing the SM plus DBT data set, readers were instructed to first evaluate the SM images to identify mammograms...
Microcalcifications Detected at Screening Mammography

Figure 1: Images in a 66-year-old woman (not a study patient) with scattered fibroglandular breast density and grouped microcalcifications seen at (a) synthetic mammography (SM) versus (b) full-field digital mammography. Note loss of some fine microcalcifications and blurring of other microcalcifications at SM (circle). Halo artifact (arrows) adjacent to coarser microcalcifications at SM.

with microcalcifications warranting recall. Readers were allowed to view the DBT images to further evaluate microcalcifications seen at SM if they felt it would be diagnostically helpful. For example, DBT images could be used to clarify whether microcalcifications seen at SM were real or artifactual; however, viewing the DBT images was not required for this study as our intent was to compare detection of microcalcifications at SM versus FFDM.

Initially, two readers independently interpreted the FFDM data set while two other readers independently interpreted the SM plus DBT data set. After waiting a minimum 1-month interval to minimize recall bias, readers each independently interpreted the remaining data set.

A true-positive finding was defined as a match between a study reader’s recall for the microcalcifications (either benign or malignant) and a recall in the clinical interpretation documented in the medical record. The quadrant location of the microcalcifications was also required to match that described in the clinical report. A true-negative finding was defined as no recall by the study reader for microcalcifications and no recall in the clinical report documented in the medical record. Please refer to Statistical Analysis section for details regarding sensitivity and specificity calculations.

Statistical Analysis

We compared the interreader agreement by using Fleiss kappa. To determine if readings differed by the reader, modality, or truth (ie, the finding documented in the clinical report in the medical record), we performed generalized mixed-effects analyses with reader, modality, and truth being the fixed effects and mammogram modeled as the random effect. Because multiple readings on the same mammogram by various readers might be correlated due to shared characteristics of the mammogram, we chose the mixed-effects model as the primary analysis. To examine the robustness of conclusions, we also performed a logistic regression model that treats the repeated readings on the same mammogram as independent observations. The two models were compared based on the likelihood ratio test.

Sensitivities and specificities were calculated for the two reading arms: for all microcalcifications and for subgroups of benign and malignant microcalcifications. Uncertainty in the estimates of sensitivity and specificity was examined based on 95% binomial confidence intervals (CI). We interpreted $P$ values less than .001 as indicating a statistically significant finding. Post hoc power was applied to analysis to determine minimum detectable differences in sensitivity and specificity. Statistical analyses were performed by using RStudio version 1.0.136 (20).

Post hoc power analyses indicated that our study had 80% power to detect differences of 7% or greater in sensitivity and specificity between the FFDM and SM plus DBT arms at the two-sided .05 significance level.

Results

Reader Agreement

There was moderate agreement between the readers based on Fleiss kappa. For the FFDM arm, Fleiss kappa was 0.67 ($P < .001$); for the SM plus DBT arm, Fleiss kappa was 0.63 ($P < .001$).

Sensitivity and Specificity for Detection of Microcalcifications

The combined sensitivity of all four readers for all microcalcifications (benign and malignant) was 80% (229 of 288; 95% CI: 74%, 84%) for FFDM and 75% (215 of 288; 95% CI: 69%, 80%) for SM plus DBT. Combined specificity of all readers for all microcalcifications (benign and malignant) was 97% (78 of 80; 95% CI: 91%, 100%) for FFDM and 95% (76 of 80; 95% CI: 88%, 99%) for SM plus DBT. Data for individual readers are shown in Figure 2.

Combined sensitivity of all readers for malignant microcalcifications alone was 92% (66 of 72; 95% CI: 83%, 97%) for FFDM and 94% (68 of 72; 95% CI: 86%, 98%) for SM plus DBT. Combined specificity of all readers for malignant microcalcifications (benign and malignant) was 98% (78 of 80; 95% CI: 91%, 100%) for FFDM and 95% (76 of 80; 95% CI: 88%, 99%) for SM plus DBT. Data for individual readers are shown in Figure 3.

The mixed-effects model concluded that there were no differences between readers ($-0.02; 95\% CI: -0.03, 0.004; P = .09$) or between modalities ($-0.03; 95\% CI: -0.08, 0.01; P = .13$).
Lai et al

Radiology: Volume 289: Number 3—December 2018 • radiology.rsna.org

Table 1: Characteristics of Mammography Examinations and Pathology Results

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All Mammograms (n = 92)</th>
<th>Mammograms with Malignant Microcalcifications (n = 18)</th>
<th>Mammograms with Benign Microcalcifications (n = 54)</th>
<th>Normal Controls* (n = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast density</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatty</td>
<td>13</td>
<td>2</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Scattered</td>
<td>36</td>
<td>8</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>Heterogeneously dense</td>
<td>34</td>
<td>5</td>
<td>23</td>
<td>6</td>
</tr>
<tr>
<td>Extremely dense</td>
<td>9</td>
<td>3</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Size (mm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>7</td>
<td>11</td>
<td>9</td>
<td>...</td>
</tr>
<tr>
<td>Min</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>...</td>
</tr>
<tr>
<td>Max</td>
<td>20</td>
<td>22</td>
<td>21</td>
<td>...</td>
</tr>
<tr>
<td>Morphology †</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear branching</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>...</td>
</tr>
<tr>
<td>Fine pleomorphic</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>...</td>
</tr>
<tr>
<td>Amorphous</td>
<td>29</td>
<td>3</td>
<td>26</td>
<td>...</td>
</tr>
<tr>
<td>Coarse heterogeneous</td>
<td>24</td>
<td>7</td>
<td>17</td>
<td>...</td>
</tr>
<tr>
<td>Round and punctate</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>Vascular</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>...</td>
</tr>
<tr>
<td>Rim</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>Pathology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malignant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ductal carcinoma in situ</td>
<td>...</td>
<td>12</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Invasive ductal carcinoma</td>
<td>...</td>
<td>2</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>with ductal carcinoma in situ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invasive ductal carcinoma</td>
<td>...</td>
<td>4</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Benign</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atypical ductal hyperplasia</td>
<td>...</td>
<td>...</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>Papilloma</td>
<td>...</td>
<td>...</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>Hyalinizing fibroadenoma</td>
<td>...</td>
<td>...</td>
<td>12</td>
<td>...</td>
</tr>
<tr>
<td>Fat necrosis</td>
<td>...</td>
<td>...</td>
<td>2</td>
<td>...</td>
</tr>
<tr>
<td>Other ‡</td>
<td>...</td>
<td>...</td>
<td>20</td>
<td>...</td>
</tr>
<tr>
<td>Benign at diagnostic imaging</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vascular calcification</td>
<td>...</td>
<td>...</td>
<td>3</td>
<td>...</td>
</tr>
<tr>
<td>Benign based on long-term stability</td>
<td>...</td>
<td>...</td>
<td>15</td>
<td>...</td>
</tr>
</tbody>
</table>

* Randomly selected screening mammograms without microcalcifications (assessed as either Breast Imaging Reporting and Data System [BI-RADS] 2: benign or BI-RADS 1: negative).

† Predominant morphology.

‡ Other benign pathology: sclerosing adenosis, apocrine hyperplasia, fibrocystic change, heterotopic bone calcifications.

Reader interpretations of recall versus no recall for microcalcifications, however, were associated with true recall results (0.73; 95% CI: 0.60, 0.87; \( P < .001 \)).

The likelihood ratio test comparing the mixed-effects model and the logistic regression inferred no statistically significant difference between the mixed-effects and logistic regression models (\( P = .99 \)). Both mammogram and reader’s interpretation were statistically significantly associated with the truth in the logistic regression model, with coefficients 0.04 (95% CI: 0.03, 0.05; \( P = .001 \)) for mammogram and 5.1 (95% CI: 4.2, 6.1; \( P = .001 \)) for reader interpretation. Reader and modality were not statistically significant, with coefficients 0.15 (95% CI: −0.08, 0.37; \( P = .21 \)) and 0.22 (95% CI: −0.27, 0.71; \( P = .38 \)), respectively. We note that the mixed-effects model and the logistic regression model reached similar conclusions.

Malignant Microcalcifications Missed by Readers

No cancers were missed by a majority (three or four) of readers. However, there were instances where one or two readers missed cancer on a mammography examination. For both SM and FFDM, there were three missed cancers in common: one 9-mm group of fine pleomorphic microcalcifications (ductal carcinoma in situ [DCIS] missed by two of four readers at SM and by one reader at FFDM), one 6-mm group of fine pleomorphic microcalcifications (DCIS missed by one reader at SM and by one reader at FFDM), and one 8-mm group of coarse heterogeneous and linear microcalcifications (DCIS missed by one reader at SM and by two readers at FFDM). For FFDM, there were two additional missed cancers: one 20-mm group of round and punctate microcalcifications (invasive ductal carcinoma missed by one reader) and one 7-mm group of
Microcalcifications Detected at Screening Mammography

634

Discussion

Advantages of reduced radiation dose and shorter imaging times make SM a desirable substitute for conventional digital mammography when DBT is used at screening. However, due to technical limitations inherent in SM, it is uncertain whether SM is equivalent to FFDM for the detection of microcalcifications. In this retrospective reader study, we compared FFDM alone versus SM plus DBT for the detection of microcalcifications. The combined sensitivity of all four readers for all microcalcifications (benign and malignant) was 80% (229 of 288; 95% CI: 74%, 84%) for FFDM and 75% (215 of 288; 95% CI: 69%, 80%) for SM plus DBT. Combined specificity of all readers for all microcalcifications (benign and malignant) was 97% (78 of 80; 95% CI: 91%, 100%) for FFDM and 95% (76 of 80; 95% CI: 88%, 99%) for SM plus DBT. Relative to FFDM, SM plus DBT had similar sensitivity and specificity for the detection of microcalcifications previously identified for recall at screening mammography.

Our standard practice is to focus on detection at time of screening and reserve characterization for the diagnostic mammography examination, which typically includes spot magnification compression views of the microcalcifications in two projections to more clearly portray the shapes of the microcalcifications. However, DBT images can be useful to confirm the presence of grouped microcalcifications (ie, not superimposition), as well as approximate location within the breast. Kopans et al has previously demonstrated the potential for DBT to detect microcalcifications, although synthetic views were not available at that time (21). Our practice routinely compares with multiple prior mammograms when available, as this has been shown to reduce recall rate and increase cancer detection rate (22).

A recent reader study by Zuley et al (10), also using FDA-approved SM software (C-View; version 1, Hologic, Bedford, Conn), demonstrated that SM alone or in combination with DBT had comparable performance to FFDM alone or in combination with DBT; however, only 19 mammograms with microcalcifications were included in the study data set. Our larger study data set of 72 mammograms with microcalcifications allowed us to more definitively compare the performance of SM plus DBT to FFDM with respect to microcalcifications.

In the TOMMY (comparison of Tomosynthesis with Digital Mammography in the United Kingdom National Health Service Breast Screening Programme) trial (23,24), a large multicenter United Kingdom study of cancer detection using DBT, a subgroup analysis showed that SM plus DBT had amorphous microcalcifications (DCIS missed by one reader). Examples of malignant microcalcifications missed by one of the readers at FFDM and missed by two of the readers at SM plus DBT are shown in Figures 4 and 5, respectively. An example of benign microcalcifications missed by two readers at FFDM is shown in Figure 6.

Figure 2: (a) Sensitivity and (b) specificity of full-field digital mammography (FFDM) versus synthetic mammography plus digital breast tomosynthesis (SM+DBT) for microcalcifications prompting recalls. For FFDM, sensitivity by reader is as follows: reader 1 (R1), 65 of 72 (90%; 95% confidence interval [CI]: 81%, 96%); reader 2 (R2), 52 of 72 (72%; 95% CI: 60%, 82%); reader 3 (R3), 59 of 72 (82%; 95% CI: 71%, 90%); and reader 4 (R4), 53 of 72 (74%; 95% CI: 62%, 83%). For FFDM, specificity by reader is as follows: reader 1 (R1), 20 of 20 (100%; 95% CI: 76%, 100%); reader 2 (R2), 20 of 20 (100%; 95% CI: 76%, 100%); reader 3 (R3), 18 of 20 (90%; 95% CI: 68%, 99%); and reader 4 (R4), 20 of 20 (100%; 95% CI: 76%, 100%). For SM+DBT, sensitivity by reader is as follows: reader 1 (R1), 58 of 72 (81%; 95% CI: 70%, 89%); reader 2 (R2), 51 of 72 (71%; 95% CI: 59%, 81%); reader 3 (R3), 46 of 72 (64%; 95% CI: 52%, 75%); and reader 4 (R4), 60 of 72 (83%; 95% CI: 73%, 91%). For SM+DBT, specificity by reader is as follows: reader 1 (R1), 18 of 20 (90%; 95% CI: 68%, 99%); reader 2 (R2), 19 of 20 (95%; 95% CI: 75%, 100%); reader 3 (R3), 20 of 20 (100%; 95% CI: 76%, 100%); and reader 4 (R4), 19 of 20 (95%; 95% CI: 75%, 100%).
in performance with respect to microcalcifications. Due to sample size limitations, our study could not exclude a difference in sensitivity for microcalcification detection of 7% or less between SM plus DBT and FFDM alone. In addition, our study results may differ due to differences in patient populations, as well as interpretive standards between the United States and United Kingdom. It is also important to note that our study was specifically designed to focus on detection of microcalcifications rather than cancer detection in general.

Our results are concordant with those of prior prospective screening studies. In the Oslo Tomosynthesis Screening Trial, Skaane et al found that the FDA-approved version of SM in combination with DBT yielded comparable performance to FFDM plus DBT, with no significant differences in the cancer detection rate or false-positive rate (8). Zuckerman et al showed that screening with SM plus DBT in a large urban U.S. practice resulted in similar performance outcomes of cancer detection rate, recall rate, and biopsy rate when compared with their institution’s historical performance using FFDM plus DBT (25). Although results of these prospective studies suggest that substituting SM for FFDM when screening with DBT would likely achieve similar outcomes, differences in sensitivity for cancers manifesting as microcalcifications cannot be excluded on the basis of these studies. This is because reduced sensitivity of SM plus DBT for microcalcifications could have been masked by improved sensitivity for other types of findings, such as masses and distortions, when overall performance is compared. Zuckerman et al did report a reduction in the recall rate for microcalcification findings with SM plus DBT, indicating the need for further study to determine whether clinically significant microcalcifications are missed with this technique (25).

More recently, Aujero et al found no significant difference in DCIS detection rates in their community practice when switching from FFDM plus DBT to SM plus DBT (26). Population screening results comparing SM plus DBT to conventional digital mammography found an increase in DCIS detection with SM plus DBT versus conventional digital mammography (27). These more recent study results suggest that SM may be sufficient to detect clinically significant microcalcifications.

Our study had several limitations. This was a retrospective reader study with a limited sample size, likely introducing some selection bias. In addition, as discussed previously, our sample size limited our ability to detect a difference in sensitivity or specificity of 7% or less between the study groups. Our study was conducted at a single academic center with fellowship-trained radiologists exclusively practicing breast imaging, and therefore our results may not be

3% lower sensitivity for the depiction of microcalcifications and 7% lower sensitivity for depiction of 11–20-mm DCIS relative to FFDM (23,24). Because of the large sample size of 1288 mammograms with microcalcifications, the TOMMY trial had greater statistical power than our study to detect small differences

Figure 3: (a) Sensitivity and (b) specificity of full-field digital mammography (FFDM) versus synthetic mammography plus digital breast tomosynthesis (SM+DBT) for detection of malignant microcalcifications. For FFDM, sensitivity by reader is as follows: reader 1 (R1), 18 of 18 [100%; 95% confidence interval (CI): 81%, 100%]; reader 2 (R2), 16 of 18 [89%; 95% CI: 65%, 99%]; reader 3 (R3), 16 of 18 [89%; 95% CI: 65%, 99%]; and reader 4 (R4), 16 of 18 [89%; 95% CI: 65%, 99%]. For SM+DBT, sensitivity by reader is as follows: reader 1 (R1), 18 of 18 [100%; 95% CI: 81%, 100%]; reader 2 (R2), 16 of 18 [89%; 95% CI: 65%, 99%]; reader 3 (R3), 15 of 18 [83%; 95% CI: 59%, 96%]; and reader 4 (R4), 16 of 18 [89%; 95% CI: 65%, 99%].
Microcalcifications Detected at Screening Mammography

As a consequence, we were not able to generate receiver operating characteristic curves to analyze reader performance.

Another limitation is that we did not test all the permutations of SM alone or FFDM plus SM plus DBT as a matter of practicality. We did not expect any practice to interpret only SM images without the DBT images. Given the desire to minimize radiation dose and the continued improvements in SM image quality, we expect that practices would move toward SM plus DBT without need for separate FFDM acquisition. Also, there was concern about potential reader recall bias with a third or fourth review of the data set.

We evaluated synthetic views only from one vendor (Hologic C-View) as this was all we had available at the time of this study. The synthetic view is a maximum intensity projection of the tomosynthesis images, which is produced by vendor computer software (29). The clinical image quality of this view varies by vendor based on the same factors that affect the three-dimensional tomosynthesis image quality, namely image acquisition and reconstruction algorithm (30). Spatial resolution of the image depends on the number of pixels and is an important factor in calcification detection. The number of effective pixels is affected by a process called pixel binning, which depends on the type of detector and therefore varies by manufacturer. For example, the GE image maintains full spatial resolution because each pixel is read by the detector, improving calcification detection on this unit (21,31). In contrast, the Hologic detector does not read each pixel (pixel binning), which decreases spatial resolution.
(21,31). As there are now many tomosynthesis vendors, it is important to review the manufacturer specifications to inform clinical image quality of the particular unit.

In conclusion, in our retrospective reader study, SM plus DBT was comparable to FFDM alone in the identification of microcalcifications previously warranting recall at screening based on FFDM plus SM plus DBT findings during routine clinical care. Further prospective reader studies with larger data sets are warranted to exclude small differences in performance.

**Acknowledgment:** The authors wish to thank Stanley Lee, PACS Administrator in the UCSF Department of Radiology and Biomedical Imaging, for assistance in anonymizing studies and preparing the worklists for this study.

**Author contributions:** Guarantors of integrity of entire study, Y.C.L., B.N.J.; study concepts/ study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, Y.C.L., K.M.R., A.Y.L., B.N.J.; clinical studies, Y.C.L., K.M.R., A.Y.L., R.I.F., B.N.J.; experimental studies, Y.C.L., K.M.R., B.N.J.; statistical analysis, Y.C.L., I.V.L.; and manuscript editing, all authors.

**Disclosures of Conflicts of Interest:** Y.C.L. disclosed no relevant relationships. K.M.R. disclosed no relevant relationships. A.Y.L. disclosed no relevant relationships. J.H.H. disclosed no relevant relationships. R.I.F. disclosed no relevant relationships. B.N.J. disclosed no relevant relationships.

**References**


