


Diagnostic value of the stand-alone synthetic image in digital breast tomosynthesis examinations

Julia Garayoa¹ · Margarita Chevalier²  · Maria Castillo² · Ignacio Mahillo-Fernández³ · Najim Amallal El Ouahabi⁴ · Carmen Estrada^{5,6} · Alejandro Tejerina^{5,6} · Olivia Benitez⁵ · Julio Valverde¹

Received: 4 April 2017 / Revised: 6 July 2017 / Accepted: 12 July 2017 / Published online: 15 August 2017
© European Society of Radiology 2017

Abstract

Objective To demonstrate the non-inferiority of synthetic image (SI) mammography versus full-field digital mammography (FFDM) in breast tomosynthesis (DBT) examinations.

Methods An observational, retrospective, single-centre, multireader blinded study was performed, using 2384 images to directly compare SI and FFDM based on Breast Imaging Reporting and Data System (BIRADS) categorisation and visibility of radiological findings. Readers had no access to digital breast tomosynthesis slices. Multiple reader, multiple case (MRMC) receiver operating characteristic (ROC) methodology was used to compare the diagnostic performance of SI and FFDM images. The kappa statistic was used to estimate the inter-reader and intra-reader reliability.

Results The area under the ROC curves (AUC) reveals the non-inferiority of SI versus FFDM based on BIRADS

categorisation [difference between AUC (Δ AUC), -0.014] and lesion visibility (Δ AUC, -0.001) but the differences were not statistically significant ($p=0.282$ for BIRADS; $p=0.961$ for lesion visibility). On average, 77.4% of malignant lesions were detected with SI versus 76.5% with FFDM. Sensitivity and specificity of SI are superior to FFDM for malignant lesions scored as BIRADS 5 and breasts categorised as BIRADS 1.

Conclusions SI is not inferior to FFDM when DBT slices are not available during image reading. SI can replace FFDM, reducing the dose by 45%.

Key Points

- Stand-alone SI demonstrated performance not inferior for lesion visibility as compared to FFDM.
- Stand-alone SI demonstrated performance not inferior for lesion BIRADS categorisation as compared to FFDM.

✉ Margarita Chevalier
chevalie@ucm.es

Julia Garayoa
julia.garayoa@quironosalud.es

Maria Castillo
m.castillo.g@ucm.es

Ignacio Mahillo-Fernández
imahillo@fjd.es

Najim Amallal El Ouahabi
namallal@emsor.es

Carmen Estrada
CEstrada@quironosalud.es; carmen.estrada@cpmama.es

Alejandro Tejerina
alejandrotejerina@cpmama.es; alejandro.tejerina@quironosalud.es

Olivia Benitez
OBenitez@quironosalud.es

Julio Valverde
JValverde@fjd.es

- ¹ Radiological Protection Unit, Hospital Universitario Fundación Jimenez Diaz, Avda. Reyes Católicos, 2, 28040 Madrid, Spain
- ² Medical Physics Group, Radiology Department, Facultad de Medicina, Universidad Complutense de Madrid, Pza. Ramón y Cajal s/n, 28040 Madrid, Spain
- ³ Servicio de Epidemiología y Bioestadística, Hospital Universitario Fundación Jimenez Diaz, Avda. Reyes Católicos, 2, 28040 Madrid, Spain
- ⁴ EMSOR, Isabel Colbrand 10-12, 28050 Madrid, Spain
- ⁵ Diagnostic Radiology Department, Hospital Universitario Fundación Jimenez Diaz, Avda. Reyes Católicos, 2, 28040 Madrid, Spain
- ⁶ Present address: Centro de Patología de la Mama, Fundación Tejerina, c/ Jose Abascal 40, 28003 Madrid, Spain

- *Synthetic images provide important dose savings in breast tomosynthesis examinations.*

Keywords Digital breast tomosynthesis · Digital mammography · Synthetic image · Lesion detectability · BIRADS categorisation

Abbreviations

DBT	Digital breast tomosynthesis
FFDM	Full-field digital mammography
SI	Synthetic image
C-View	Synthetic image commercial name
CC	Cranio-caudal view
MLO	Medio lateral oblique view
IDC	Invasive ductal carcinoma
DCIS	Ductal carcinoma in situ
ILC	Infiltrating lobular carcinoma
BIRADS	Breast Imaging and Reporting and Data System
MRMC	Multiple reader multiple case
ROC	Receiver operating characteristics
AUC	Area under the ROC curve

Introduction

Digital breast tomosynthesis (DBT) is a relatively new imaging technique that is expanding widely in breast diagnosis centres. DBT uses a series of individual low-dose projections acquired while the x-ray tube is rotating over a limited arc above the compressed breast. Using mathematical algorithms, data from these multiple low-dose projections are reconstructed into a quasi-3D breast volume of thin slices parallel to the detector plane. Thus, DBT potentially facilitates diagnosis of breast lesions by reducing tissue overlap. Published clinical studies show that the accuracy of one- or two-view DBT is equal or better than that of conventional full-field digital mammography (FFDM). These studies also show superior lesion detection and lower recall rate when DBT is used in combination with FFDM [1–10]. The combination of DBT plus FFDM yields mean glandular doses that may double the doses delivered in FFDM examinations [11]. According to our previous studies [12–15] the mean glandular dose due to DBT compared to the FFDM is 50% higher for 5–6-cm breast thickness (most common), 40% for thicknesses of 3 to 4 cm, and 30% for 7 to 8 cm. Dose values are of great concern, especially for the purposes of incorporating this technology in breast-screening programs. This has led most manufacturers to develop a 2D synthetic image (SI) from the reconstructed tomographic slices with the aim of substituting the FFDM images.

Some studies have addressed the clinical performance of SIs. Skaane et al. [16] conducted a prospective study over a screening

population (12,270 people) with an arm aiming to compare SI with FFDM. The results show comparable performance of SI + DBT and FFDM + DBT in terms of cancer detection rates and false positive scores. The TOMMY trial [17] is a retrospective reading study with three arms (FFDM vs. DBT + FFDM vs. DBT + SI) in which 7060 cases were blindly reviewed. This study concluded that DBT + SI showed a similar performance to that of DBT + FFDM. In a retrospective study of 214 cases, Choi et al. [18] concluded that SI and FFDM show comparable detection rates for T1-stage breast cancers.

Most published studies compare the sensitivity of SI versus FFDM when used in combination with DBT slices. Direct comparison between FFDM and SI with no access to DBT slices may be interesting for avoiding the influence of DBT reading in the diagnosis. The aim of the present work is to evaluate the clinical performance of the SI alone, compared with FFDM alone in terms of lesion detectability and BIRADS lesion categorisation [19]. In [15], we published a preliminary study based on phantom images and a reduced patient sample (50 patients). We found that the visibility of radiological findings in the clinical images (grouped as architectural distortions, micro-calcifications, and nodules) was similar for both types of images except for distortions, which were better visualised in SI ($p < 0.01$). However, this was a preliminary study with some limitations in that it lacked a sufficiently large patient sample, readers had access to the corroborated patient diagnosis, and intra-observer variability was not evaluated. Therefore, we have now developed a more conclusive study based on a sufficiently large enough image sample where intra-observer variability is included as one of the sources of uncertainty.

We compare the sensitivity and specificity of the FFDM and the SI in order to prove the non-inferiority of the SI. A positive result would allow the replacement of clinical protocol based on DBT + FFDM in favour of DBT + SI, with the subsequent dose savings.

Material and methods

An observational, retrospective, single-centre, multireader blinded study was performed following approval of the institutional ethics committee.

Study design and patient sample

The sample size was calculated to provide a statistical power of at least 80% when establishing the non-inferiority of SI compared to FFDM regarding the diagnostic capability of the two image types [20]. A set of 244 patients who underwent a 4-projection (2 breasts x 2 projections: CC and MLO) COMBO exam (exam routinely performed in our institution at the time of the study) in a Selenia Dimensions DBT unit

(Hologic Inc., Bedford, MA, USA) between May 2013 and July 2014 were included in the sample. As is well known, the COMBO modality performs FFDM acquisition followed by a DBT acquisition with the breast under the same compression force. For all patients and all acquired projections, the SI (version 1.0.0.1) was obtained (C-View for Hologic).

All recruited patients arrived at our breast unit for a screening or diagnostic appointment. A radiologist who did not participate in the reading study selected the patients based on their final diagnosis as determined by final interpretation with complementary ultrasound or magnetic resonance examinations, or biopsy with histological studies. Inclusion criteria were: a) breasts with no mammographic findings randomly selected from all available cases and b) breasts with mammographic benign and malignant findings representing the typical range of lesions found in the clinical practice. Patients with breast prosthesis were not enrolled. A subset of 54 patients was included twice in order to evaluate intra-observer variability. Thus, the effective sample size was 298 patients. The sample included 119 biopsy-proven cancers, 15 high-risk lesions, 110 benign lesions, and 350 breasts with no lesions. 26 breasts in the sample had two lesions. The ground truth in this study was defined in terms of BIRADS categorisation and radiological findings reported during the routine diagnosis, complementary examinations, or histological studies.

In order to guarantee blind evaluation, the FFDM and SI of each breast were separately anonymised so that the two types of images and the images corresponding to contralateral breasts were de-coupled. DBT slices were discarded since they were not used in the study. For each patient included in the sample, four anonymised studies were generated containing two images corresponding to the FFDM (SI) CC and MLO projections of the left (right) breast (see Fig. 1). All the anonymised studies were randomly ordered. In total, 1192 anonymised studies (2384 images) were included (596 FFDM and 596 SI).

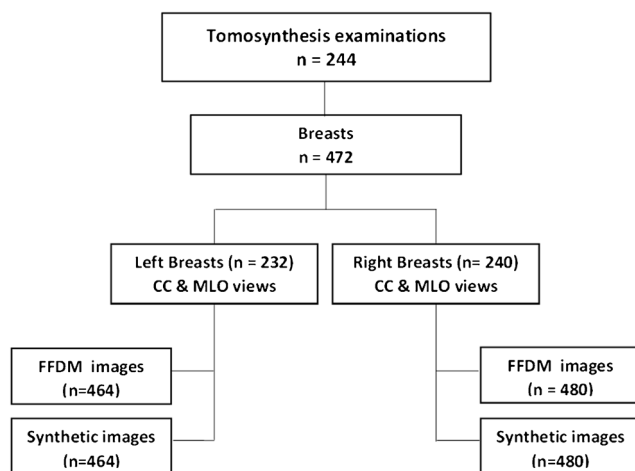


Fig. 1 Study flowchart

Reader study

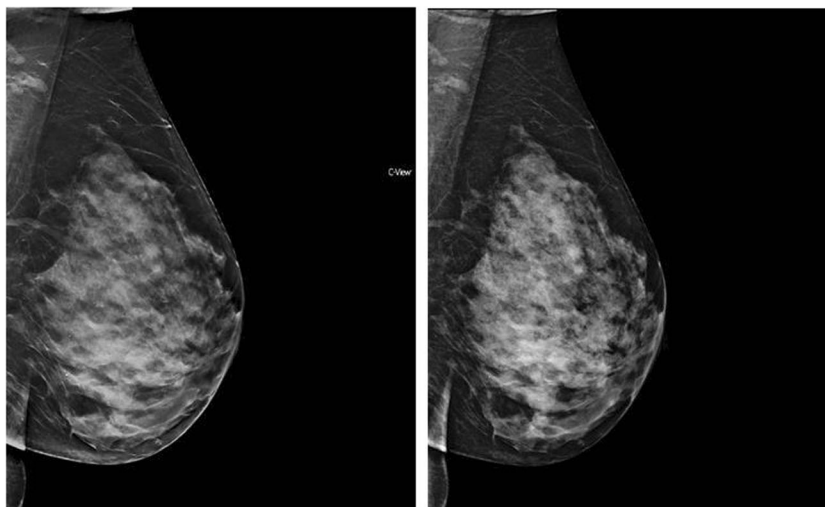
The images were read by three experienced radiologists in digital mammography (over 5000 mammograms per year [21]) and DBT (over 7000 studies per year). The SI reading experience of the radiologists was 1 year on average (first SI version was installed in late 2012). The reading sessions started 4 months after patient recruitment and were performed in multiple sessions in an independent 5Mp Hologic workstation courtesy of Emsor (distributor of Hologic systems in Spain) without the ability to recall DBT slices or previous studies of the patients. The CC and MLO projections of a single breast were presented to the reader, who had to detect mammographic findings, score their visibility, and classify the breast according to the BIRADS categorisation. Readers were blinded to patient clinical history or images of the contralateral breast. It is important to point out that SI images are easily traceable due to their characteristic texture, which is easily identifiable for experienced readers, as well as to the C-View tag present in the images (Fig. 2). The randomly ordered image set guaranteed that images corresponding to the same patient were separated. Image readings were performed over 8 months in order to prevent memory effects.

Readers were allowed to score a maximum of three mammographic findings at each image that had to be classified in five categories: micro-calcification, nodule, nodule with micro-calcifications, architectural distortion, and focal density. The visibility was rated on a scale of 0 to 3 (0: no finding detected; 1: subtle visibility and very difficult characterisation; 2: medium visibility and difficult characterisation; 3: clearly visibility and characterisation). Finally, the BIRADS categorisation (1–5) was used to classify each breast according to the probability of malignancy of the more suspicious finding. Readers were provided with a data sheet designed using database software (Microsoft Access) and were asked to assign a BIRADS category and to select the type of finding, scoring its visibility. A case number that matched with the one viewed on the workstation was provided on each data sheet.

Statistical Analysis

Inter- and intra-reader agreements for both BIRADS categorisation and lesion visibility were separately evaluated for SI and FFDM. The agreement level was assessed by calculating the kappa coefficient with a 95% confidence interval (CI). Conventionally, kappa values of 0.00–0.20, 0.21–0.40, 0.41–0.60, 0.61–0.80, and 0.81–1.00 indicate minimal, fair, moderate, substantial, and near-perfect agreement, respectively [22]. Multiple-reader, multiple-case (MRMC) ROC methodology was used to compare the diagnostic capabilities of SI and FFDM. ROC curves were determined for each of the three readers and overall, using the BIRADS categorisation and lesion visibility. Smooth ROC curves were calculated using

Fig. 2 (Left) Synthetic and (right) FFDM mediolateral oblique images of a 51-year-old woman



the bi-normal model [23], and this was performed using the ROC function from the pROC package of the R programme. The diagnostic capability of SI and FFDM was defined in terms of the area under the ROC curve (AUC). The overall comparison of diagnostic capabilities of SI and FFDM were obtained from the difference in AUC. Non-inferiority of SI against FFDM was evaluated using the 95% CI for the difference between mean AUCs. Confidence limits were obtained using an Obuchowski–Rockette model with Hillis improvements to calculate degrees of freedom [24]. This was performed using the ORH analysis function from the RJafroc package of R programme. To conclude that SI was non-inferior to FFDM, it was required that the lower limit of the CI be above the non-inferiority margin.

Results

The radiological findings were: 121 nodules, 42 microcalcifications, 19 nodules with microcalcifications, 24 distortions, and 13 focal densities (the values correspond to the number of confirmed findings present in the breast sample). The cancers in the effective sample were 77 invasive ductal carcinoma (IDC), 7 ductal carcinoma in situ (DCIS), 11 infiltrating lobular carcinoma (ILC), 6 IDC + DCIS, and 18 other cancers.

Table 1 shows kappa coefficients (95% CI) for agreement between all three readers based on the BIRADS categorisation and lesion detectability (detected/non-detected) for both SI and FFDM.

BIRADS were grouped into two categories: 1–3 and 4–5, which separates healthy breasts or breasts with benign findings (BIRADS 1–3) from breasts with malignant lesions (BIRADS 4–5). Results show substantial agreement between readers for BIRADS categorisation in both image modalities, with a slightly higher kappa for SI. The results of the analysis

performed over the 5-step BIRADS categorisation reveal a slightly poorer agreement (results not shown).

Substantial inter-reader agreement was also found for nodule and micro-calcifications detectability in both FFDM and SI, while fair to moderate agreement was found for densities, distortions, and nodule+micro findings. Here, it is important to consider that this result is based on a poorer statistical sample: only 13 densities, 24 distortions, and 20 nodules+micos were available in the patient sample as compared to 121 nodules and 42 micro-calcifications.

Intra-reader agreement for BIRADS categorisation (1–3, 4–5) and lesion detectability shows almost perfect agreement for all the readers and both image modalities (Table 2). Some exceptions were for reader 1 that showed a fair agreement with himself for densities in the SI, and for reader 3 in the case of architectural distortions in the FFDM. A high 95% CI must be noted for densities, distortions, and nodules+micos due to the lower number of cases in the sample.

A substantial agreement between SI and FFDM was found for BIRADS categorisation and nodule and micro-calcification detectability for all three readers (Table 3). Moderate agreement was found for all other radiological findings.

Table 1 Inter-reader concordance for BIRADS categorisation (1–3, 4–5) and lesion detectability (detected/non-detected) for each reader and for digital mammography (FFDM) and synthetic image (SI)

	FFDM kappa (95% CI)	SI kappa (95% CI)
BIRADS	0.70 (0.65–0.75)	0.76 (0.71–0.80)
Lesion detectability		
Nodule	0.63 (0.58–0.67)	0.67 (0.62–0.72)
Micro-calcification	0.75 (0.71–0.80)	0.70 (0.65–0.75)
Nodule+micro	0.49 (0.44–0.53)	0.36 (0.31–0.41)
Distortion	0.45 (0.40–0.50)	0.53 (0.48–0.58)
Density	0.38 (0.33–0.43)	0.29 (0.24–0.33)

Table 2 Intra-reader agreement for BIRADS categorisation (1–3, 4–5) and lesion detectability (detected/non-detected) for each reader and for digital mammography (FFDM) and synthetic image (SI)

BIRADS	Kappa (95% CI)		
	Reader 1	Reader 2	Reader 3
FFDM	0.70 (0.54–0.85)	0.75 (0.59–0.90)	0.79 (0.64–0.93)
SI	0.76 (0.62–0.90)	0.85 (0.73–0.98)	0.89 (0.78–0.99)
Lesion detectability	FFDM kappa (95% CI)		
Nodule	0.80 (0.66–0.94)	0.77 (0.58–0.96)	0.74 (0.59–0.89)
Micro-calcification	0.78 (0.57–0.99)	0.87 (0.68–1.00)	0.81 (0.60–1.00)
Nodule+micro	0.42 (0.01–0.83)	0.79 (0.40–1.00)	0.82 (0.58–1.00)
Distortion	0.59 (0.27–0.92)	0.76 (0.53–0.98)	0.38 (0.00–0.77)
Density	0.78 (0.55–1.00)	0.77 (0.58–0.96)	0.47 (0.16–0.77)
	SI Kappa (95% CI)		
Nodule	0.72 (0.54–0.90)	0.87 (0.73–1.00)	0.72 (0.52–0.92)
Micro-calcification	0.90 (0.76–1.00)	0.88 (0.71–1.00)	0.81 (0.60–1.00)
Nodule+micro	0.42 (0.01–0.84)	0.79 (0.40–1.00)	0.88 (0.71–1.00)
Distortion	0.60 (0.32–0.87)	0.76 (0.53–0.98)	0.52 (0.16–0.89)
Density	0.31 (0.18–0.81)	0.68 (0.42–0.94)	0.65 (0.28–1.00)

AUC for each reader and mean AUC for the three readers for both SI and FFDM were obtained by combining the visibility scores for all the radiological findings (Table 4).

The 5-step BIRADS categorisation was used to compute the ROC curve [Fig. 3a]. The difference between the AUC of SI and FFDM across the three readers (Fig. 4) is -0.014 (95% CI: -0.042–0.016), which is not statistically significant ($p = 0.282$). Therefore, SI proved to be non-inferior to FFDM based on BIRADS categorisation.

The difference between the computed AUC for lesion visibility (Fig. 4) in SI and FFDM across the three readers is -0.001 (95% CI: -0.035–0.037), which is not statistically significant ($p = 0.9607$). Therefore, SI proved to be non-inferior to FFDM based on lesion visibility.

Regarding the sensitivity of both image modalities, the rate of correct detection of malignant lesions (ground truth = BIRADS 5) was computed, assuming that the lesion would have been detected if a BIRADS 5 or 4 had been assigned

during image reading (Table 5). On average, FFDM images had a higher sensitivity (79%) than SI (75%) although this difference was not statistically significant (95%CI: -0.15–0.16). The sensitivity was also calculated by considering only those malignant lesions scored as BIRADS 5 (Table 5). The results showed that on average, SI images had higher sensitivity (63%) than FFDM (58%) and the differences were statistically significant (Mean difference = 0.046, $p = 0.001$).

Table 5 also shows the specificity for each reader and the mean for each imaging type. The specificity was calculated by dividing the total number of breasts scored as benign or without lesion (BIRADS 1–3) by the total number of breasts in the sample that were benign and without lesion [3]. SI and FFDM presented similar specificity, and the differences in the mean values are not statistically significant. In a similar way, the specificity was recalculated by only taking into account the number of lesions scored as BIRADS 1 (breasts without lesions). SI presented a higher specificity (86%) than FFDM

Table 3 Agreement between synthetic image and digital mammography for BIRADS categorisation (1–3, 4–5) and lesion detectability (detected/non-detected) for the three readers

Kappa (95% CI)	Kappa (95% CI)		
	Reader 1	Reader 2	Reader 3
BIRADS	0.74 (0.68–0.80)	0.73 (0.67–0.80)	0.73 (0.67–0.80)
Nodule	0.73 (0.66–0.79)	0.79 (0.73–0.85)	0.64 (0.57–0.71)
Micro-calcification	0.79 (0.71–0.87)	0.76 (0.67–0.86)	0.72 (0.62–0.82)
Nodule+micro	0.61 (0.40–0.82)	0.80 (0.60–0.99)	0.73 (0.59–0.86)
Distortion	0.51 (0.37–0.66)	0.55 (0.39–0.67)	0.67 (0.53–0.81)
Density	0.42 (0.27–0.58)	0.52 (0.38–0.65)	0.40 (0.25–0.54)

Table 4 Area under the ROC curves for each reader and for synthetic image and digital mammography based on BIRADS categorisation and lesion visibility

	FFDM AUC (95% CI)	SI AUC (95% CI)
BIRADS		
Reader 1	0.878 (0.849–0.907)	0.851 (0.818–0.883)
Reader 2	0.849 (0.816–0.881)	0.852 (0.819–0.885)
Reader 3	0.875 (0.845–0.904)	0.857 (0.825–0.889)
Mean	0.867 (0.835–0.899)	0.853 (0.826–0.881)
Lesion visibility		
Reader 1	0.817 (0.783–0.851)	0.822 (0.788–0.855)
Reader 2	0.798 (0.762–0.833)	0.813 (0.779–0.848)
Reader 3	0.822 (0.790–0.854)	0.800 (0.765–0.834)
Mean	0.811 (0.800–0.823)	0.812 (0.800–0.823)

(81%), and the differences were statistically significant (mean difference = 0.049; 95% CI: -0.072–0.015; $p = 0.007$).

Discussion

In this work, we demonstrate that the clinical performance of SI is not inferior to that of FFDM images for lesion visibility or BIRADS categorisation. Other published studies [16–18] compared the clinical performance of SI + DBT and FFDM + DBT. At present, the use of SI as a valid image for replacing FFDM in DBT examinations is under debate. The good results obtained with DBT in the screening programs reinforce this debate. The inclusion of DBT in these programs entails overcoming important challenges such as the time it takes to interpret the DBT exams and the dose of radiation. Therefore, we consider that direct comparison of both images can inform this discussion. Zuley et al. [25] directly compared SI and FFDM in terms of the malignancy probability assigned to various radiological findings, and they found that both image types

Fig. 3 ROC curve for synthetic image (SI) and digital mammography (FFDM) based on BIRADS categorisation (a) and lesion visibility (b)

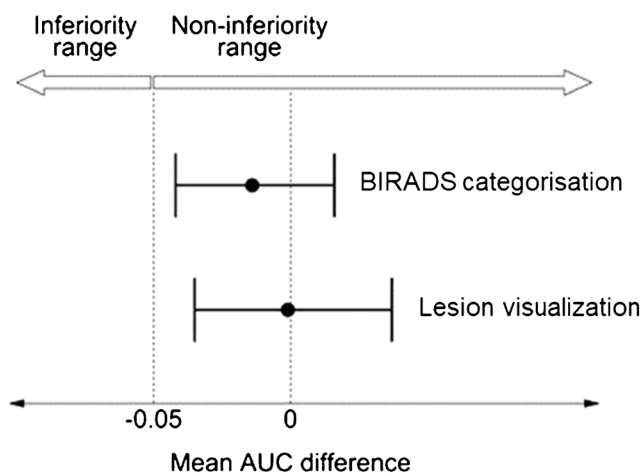
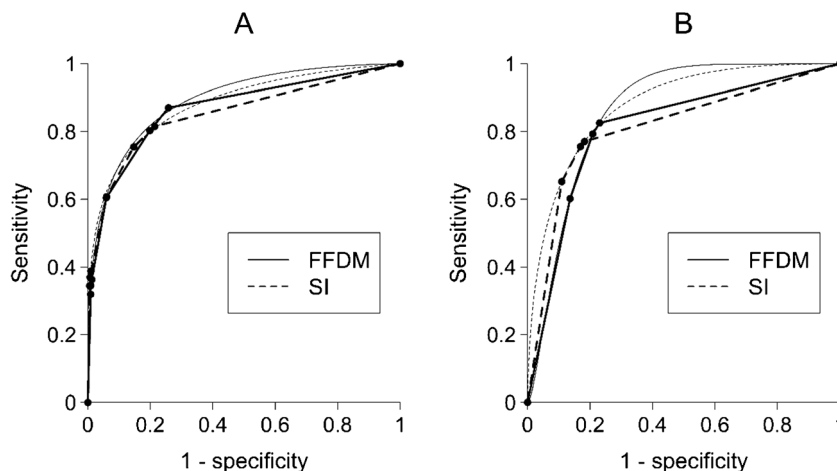


Fig. 4 Mean AUC difference between synthetic image and digital mammography (SI-FFDM) based on BIRADS categorisation and on lesion visibility, and 95% confidence interval for each difference. The zero difference line and the -0.05 non-inferiority margin are shown as dashed vertical lines

were comparable in performance. In our work, SI and FFDM were compared in terms of lesion visibility, while malignancy probability was evaluated through BIRADS categorisation. Both studies conclude the validity of SI for replacing FFDM images in DBT examinations causing substantial dose savings. According to our results in previous studies, dose values are reduced by 40–45% when using the SI instead of FFDM [13–15].

It is important to note that SI is the result of computational algorithms that evolve over time and differ between manufacturers. Skaane et al. [16] and Gur et al. [26] analysed the performance of one of the first versions of the Hologic C-View SI. They demonstrated worse performance of SI when comparing with FFDM. Locatelli et al. [27] reported low sensitivity and reduced conspicuity when using the SI generated from a DBT system of a different manufacturer. Thus, the conclusions of this study are only valid for the SI used in this research.

Table 5 Sensitivity (S) and specificity (Sp) for synthetic image (SI) and digital mammography (FFDM) of each individual reader and the means across the three readers. Sensitivity values are for malignant lesions categorised as both BIRADS 4 and 5, and BIRADS 5. Specificity values are for breasts categorised as BIRADS 1 and 3, and for BIRADS 1 breasts

	Malignant lesions N = 109 BIRADS = 4–5		Benign lesions + no lesions N = 459 BIRADS = 1–3		Malignant lesions N = 109 BIRADS = 5		No lesions N = 350 BIRADS = 1	
	Sensitivity		Specificity		Sensitivity		Specificity	
	S _{SI}	S _{FFDM}	Sp _{SI}	Sp _{FFDM}	S _{SI}	S _{FFDM}	Sp _{SI}	Sp _{FFDM}
Reader1	0.79	0.86	0.84	0.87	0.66	0.61	0.81	0.78
Reader2	0.76	0.79	0.90	0.88	0.56	0.52	0.90	0.85
Reader3	0.82	0.84	0.89	0.88	0.67	0.62	0.87	0.81
Mean	0.79	0.83	0.88	0.88	0.63	0.58	0.86	0.81

The clinical protocol followed in our institution prior to this study included two-view DBT + FFDM acquisitions per breast. The results obtained in this study encourage avoidance of FFDM, and, currently, only DBT acquisitions with SI are performed, with the subsequent dose savings. Other clinical protocols as one-view DBT + two-view 2D can provide also important dose savings [28, 29]. As with SI, the option of performing one-view DBT versus two-view DBT needs to be supported in studies that demonstrate they have a similar diagnostic capability. The results of these studies will also be dependent on the specific characteristics of the different DBT systems and can not be easily generalised.

Inter- and intra-reader agreement was performed by grouping BIRADS categories: 1–3 and 4–5, to separate healthy breasts and breasts with benign lesions from breasts with malignant lesions. This may have introduced a limitation in the study as less conspicuous lesions with BIRADS assignments of 2 or 3 become indistinguishable from un-detected lesions, where a BIRADS 1 would be assigned. To overcome this limitation, the specificity was computed considering only breasts categorised as BIRADS 1. Furthermore, the sensitivity was estimated taking into account only the breasts in the BIRADS 5 category. In both cases, this caused high reliability for the SI image. Another potential limitation is the diminished statistical power obtained for lesion visibility due to the smaller sample available for each type of lesion.

In conclusion, this study proves that the clinical performance of SI is not inferior to that of FFDM even when DBT planes are not present during image reading.

Acknowledgements The authors would like to thank Arturo Carreto for his helpful collaboration in collecting data and the important contribution and assistance of the Radiology Protection Unit and Radiology Department of the Hospital Universitario Fundación Jiménez Díaz (Madrid, Spain). They would also like to thank the representative of Hologic for the collaboration of EMSOR S.A. (Madrid, Spain).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is: Margarita Chevalier.

Conflict of interest One of the co-authors of this manuscript (Najim Amallal) declares a relationships with the company EMSOR, representative of Hologic Inc. in Spain.

The rest of the authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Funding The authors state that this work has not received any funding.

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Written informed consent was waived by the Comité Ético de Investigación Clínica del Hospital Universitario Fundación Jiménez Díaz.

Ethical approval Institutional review board (Comité Ético de Investigación Clínica del Hospital Universitario Fundación Jiménez Díaz) approval was obtained.

Methodology

- Retrospective
- Observational

References

1. 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* 18:2817–25
2. Gur D, Abrams GS, Chough DM et al (2009) Digital breast tomosynthesis: observer performance study. *Am J Roentgenol* 193:586–591
3. Gennaro G, Toledano A, di Maggio C et al (2010) Digital breast tomosynthesis versus digital mammography: a clinical performance study. *Eur Radiol* 20:1545–1553
4. Gennaro G, Hendrick RE, Ruppel P et al (2013) Performance comparison of single-view digital breast tomosynthesis plus single-view digital mammography with two-view digital mammography. *Eur Radiol* 23:664–72
5. Svahn T, Andersson I, Chakraborty D et al (2010) The diagnostic accuracy of dual-view digital mammography, single-view breast tomosynthesis and a dual-view combination of breast tomosynthesis and digital mammography in a free-response observer performance study. *Radiat Prot Dosimetry* 139:113–7
6. 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* 262:788–796
7. 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* 67:976–81
 8. Skaane P, Bandos AI, Gullien R et al (2013) Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population based screening programme using independent double reading with arbitration. *Eur Radiol* 23:2061–2071
 9. Bernardi D, Ciatto S, Pellegrini M et al (2012) Prospective study of breast tomosynthesis as a triage to assessment in screening. *Breast Cancer Res Treat* 133:267–71
 10. Gilbert FJ, Tucker L, Gillan MG et al (2015) Accuracy of Digital Breast Tomosynthesis for Depicting Breast Cancer Subgroups in a UK Retrospective Reading Study (TOMMY Trial). *Radiology* 77:697–706
 11. Svahn TM, Houssami N, Sechopoulos I et al (2015) Review of radiation dose estimates in digital breast tomosynthesis relative to those in two-view full-field digital mammography. *Breast* 24:93–99
 12. Chevalier M, Castillo M, Calzado A et al. (2012) Breast doses for tomography examinations: a pilot study. Proc. International Conference on Radiation Protection in Medicine - Setting the Scene for the Next Decade. STI/PUB/1663 (International Atomic Energy Agency, Vienna, Austria). ISBN 978–92–0–103914–9
 13. Garayoa Roca J, Castillo García M, Valverde Morán J et al. (2014) Breast tomosynthesis: dose saving and image quality of the synthesized image. Poster No.:C-0990. <http://dx.doi.org/10.1594/ecr2014/C-0990>.
 14. Garayoa J, Hernández-Girón I, Castillo M et al (2014) Digital Breast Tomosynthesis: Image Quality and Dose. *Breast Imaging: 12th International Workshop, IWDM 2014*, Fujita H, Takeshi H, Chisako M Eds. Gifu City, Japan
 15. Castillo M, Garayoa J, Estrada C et al (2015) Breast tomosynthesis: Synthesized versus digital mammography. Impact on dose. *Rev Senol Patol Mamar* 28:3–10
 16. Skaane P, Bandos A, Eben E et al (2014) Two-View digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. *Radiology* 271:655–663
 17. Gilbert F, Tucker L, Gillan M et al (2015) Accuracy of digital breast tomosynthesis for depicting breast cancer subgroups in a UK retrospective reading study (TOMMY Trial). *Radiology* 277:697–706
 18. Choi J, Han B, Ko EY et al (2016) Comparison between two-dimensional synthetic mammography reconstructed from digital breast tomosynthesis and full-field digital mammography for the detection of T1 breast cancer. *Eur Radiol* 26:2538–46
 19. American College of Radiology (2013) BIRADS Atlas — Mammography 4th. American College of Radiology, Reston. Available via: <http://www.acr.org>. Accessed 11 July 2014
 20. Obuchowski NA (2000) Sample size tables for receiver operating characteristic studies. *Am J Roentgenol* 175:603–8
 21. European Commission (2006) European guidelines for quality assurance in breast cancer screening and diagnosis 4th Ed Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L Eds European Communities (Luxemburgo) Available via: bookshop.europa.eu. Accessed October 2013
 22. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–74
 23. Hanley JA (1988) The robustness of the "binormal" assumptions used in fitting ROC curves. *Med Decis Making* 8:197–203
 24. Hillis SL (2007) A comparison of denominator degrees of freedom methods for multiple observer ROC analysis. *Stat Med* 26:596–619
 25. Zuley M, Guo B, Catullo V, Chough DM, Kelly AE, Lu AH (2014) Comparison of two-dimensional synthesized mammograms versus original digital mammograms alone and in combination with tomosynthesis images. *Radiology* 271:664–671
 26. Gur D, Zuley ML, Anello MI et al (2012) Dose reduction in digital breast tomosynthesis (DBT) screening using synthetically reconstructed projection images: an observer performance study. *Acad Radiol* 19:166–171
 27. Locatelli M, Tonutti M, Trianni A. First experience with the new generation low-dose digital breast tomosynthesis: can 2D synthetic image replace digital mammography in combination with digital breast tomosynthesis? In European Congress of Radiology 2014, 4e8 March, Vienna, Austria. Abstract B-0333.
 28. Lång K, Andersson I, Rosso A et al (2016) Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study. *Eur Radiol* 26:184–90
 29. Shin SU, Chang JM, Bae MS et al (2015) Comparative evaluation of average glandular dose and breast cancer detection between single-view digital breast tomosynthesis (DBT) plus single-view digital mammography (DM) and two-view DM: correlation with breast thickness and density. *Eur Radiol* 25:1–8