Contrast Detail Phantom Comparison on a Commercially Available Unit. Digital Breast Tomosynthesis (DBT) versus Full-Field Digital Mammography (FFDM)

Marco Bertolini,¹ Andrea Nitrosi,¹ Giovanni Borasi,¹ Andrea Botti,¹ Davide Tassoni,¹ Roberto Sghedoni,¹ and Giulio Zuccoli²

The performance of a commercial digital mammographic system working in 2D planar versus tomosynthesis mode was evaluated in terms of the image signal difference to noise ratio (SDNR). A contrast detail phantom was obtained embedding 1 cm Plexiglas, including 49 holes of different diameter and depth, between two layers containing a breast-simulating material. The phantom was exposed with the details plane perpendicular to the X-ray beam using the manufacturer's standard clinical breast acquisition parameters. SDNR in the digital breast tomosynthesis (DBT) images was higher than that of the full-field digital mammography (FFDM) for 38 out of 49 details in complex background conditions. These differences ($p < 0.05$) are statistically significant for 19 details out of 38. The relative SDNR results for DBT and FFDM images showed a dependence on the diameter of the details considered. This paper proposes an initial framework for a global image quality evaluation for commercial systems that can operate with different image acquisition modality using the same detector.

KEY WORDS: Digital mammography, full-field digital mammography (FFDM), image quality analysis, digital breast tomosynthesis, phantoms, imaging, contrast detail phantom analysis

INTRODUCTION

In recent years, digital mammography is widely replacing screen film systems (SFS). In January 2009, the USA had 6,181 full-field digital mammography (FFDM) units out of a total of 13,241 accredited mammography units. During January 2007, the FFDM units were 2,090 out of 13,548 accredited units, an increase of about 300%.¹ The reason of this "digital revolution" can be found mostly in the intrinsic digital data nature of the images that allows image processing (e.g., contrast

enhancement), computed-aided diagnosis, and fast availability of images on the hospital PACS: the processes of image acquisition, storage, retrieval, display, and evaluation are distinct allowing each one to be independently optimized.^{2,3} Furthermore, from a physical point of view, FFDM has a higher detective quantum efficiency and a greater dynamic range with respect to SFS.⁴⁻⁸ In FFDM and SFS, benign lesions and background tissue are implicated for undetected breast lesions.^{8,9} Tissue background is probably the main factor limiting cancer detection.^{10,11} In order to give a real 3D view of the breast, while retaining as much as possible the same acquisition geometry of the FFDM, the digital breast tomosynthesis (DBT) was developed.¹²⁻¹⁴ This is a cone-beam, limited-angle (15–60°) tomographic technique able to reconstruct the whole breast volume from a sequence of projection-view mammograms. Due to the reduced acquisition angle, reconstruction artifacts are expected and many studies, both on human subject and phantom, are addressed to evaluate and improve the DBT image quality.¹⁵⁻²¹

¹From the Servizio di Fisica Medica Sanitaria, Azienda Ospedaliera S. Maria Nuova, Reggio Emilia, Italy.

²From the University of Pittsburgh School of Medicine, Pittsburgh, PA, USA.

Correspondence to: Andrea Nitrosi, Servizio di Fisica Medica Sanitaria, Azienda Ospedaliera S. Maria Nuova, Reggio Emilia, Italy; tel: +39-52-2296076; fax: +39-52-2296392; e-mail: nitrosi.andrea@asmn.re.it

Copyright © 2010 by Society for Imaging Informatics in Medicine

Online publication 4 February 2010

doi: 10.1007/s10278-009-9270-0

The goal of this paper is to compare the performance of a commercially available digital mammographic system working in 2D planar versus tomosynthesis mode. To test tomosynthesis ability of reducing the structured noise, a contrast detail (CD) phantom was embedded between two slabs of confounding “breast-like” material.

MATERIALS AND METHODS

All FFDM and DBT images were acquired on a CE-marked digital unit (Selenia Dimensions Tomosynthesis System, Hologic, Bedford, MA, USA). The system uses a selenium-based detector, enabling X-ray direct detection. The pixel pitch size was 70 μm , and the detector dimensions were 24 cm $\times$ 29 cm. The acquisition resolution for the FFDM images was equivalent to the pixel pitch size, while the DBT image reconstructed pixel size was approximately 88 μm . The DBT images were reconstructed at 1-mm in-plane spacing (between slices).

Tomosynthesis was performed acquiring 15 low dose projections (one each degree) over a $\pm 7.5^\circ$ angular range in less than 5 s. The X-ray source rotated continuously around the breast with the center of rotation assumed to be coincident with the detector center. The acquisition scheme is reported in Figure 1a.

Phantom Design

A CD phantom was assembled beginning with a Nuclear Associates contrast detail phantom for mammography Model 18-252 (Nuclear Associates, New York, NY, USA). This contrast detail phantom contained a 7 $\times$ 7 matrix of objects. It included 49 circular holes of different diameter (ranging from 4.292 to 0.177 mm) and depth (from 0.050 to 0.853 mm). For the beam X-ray qualities employed and with respect to the mean phantom attenuation, the depth range corresponded to a contrast range between 0.3 and 5%. In terms of lesion simulations, the objects can be approximately categorized into two main groups: the first simulating the detectability of masses (object with diameter greater than 0.5 mm) and the latter simulating calcifications (object with diameter smaller than 0.5 mm).²² The CD phantom was positioned between two thick layers containing a

breast-simulating material. This confounding layer was obtained by properly treating surgical breast specimens. The phantom contained a set of radio-paque markers for image registration. The phantom set was equivalent in attenuation to 5.5 cm breast^{23–26} with about 50% fibroglandular composition at 28 kVp Mo/Mo.

Figure 1 shows: (a) the acquisition scheme; (b) a planar image of the phantom with no confounding material (an equivalent Plexiglas slab was used); (c) a planar image of the phantom with confounding material in place; (d) DBT image of the phantom with confounding material.

Acquisition Setup

Following the *European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis—Fourth Edition*, the exposure technique was chosen analogous to that used clinically.²³ In both acquisition modalities (FFDM and DBT), automatic exposure control was selected. The phantom was exposed with detailed plane, positioned at a height of approximately 20 mm above the breast support and perpendicular to the X-ray beam, using the manufacturer’s standard breast acquisition parameters and patterns. The Hologic’s standard acquisition pattern starts with acquisition of DBT images (3D) and successively, without releasing the compression paddle, a FFDM acquisition (2D). Each sequence of acquisition (3D + 2D) was repeated with the phantom rotated/translated to obtain different realizations of the noise (images were acquired by rotating to about 0°, 90°, 180°, and 270° the CD slab into the phantom and maintaining the confounding material in the same position, and each angular acquisition was repeated four times with slight translations of the entire phantom).

In the FFDM modality, the unit selected an exposure technique of Mo/Rh at 29 kVp and the use of the system’s cellular anti-scatter grid. In DBT mode, the unit selected an exposure technique of W/Al at 32 kVp without grid. The half value layer of Mo/Rh combination was 0.515 mm Al, while the W/Al was 0.554 mm Al. The entrance surface air kerma (ESAK) delivered for FFDM and DBT acquisitions were the same (7.73 ± 0.01 mGy for DBT and 7.63 ± 0.01 mGy for FFDM).²⁷ The corresponding average glandular doses were about 2.2 mGy

Fig 1. a The acquisition scheme; b FFDM image of the phantom without confounding material (an equivalent Plexiglas slab was used); c FFDM image of the phantom with confounding material in place; d DBT image of the phantom with complex background. The background mean standard deviations, evaluated in the ROIs shown in (b), are respectively: 7.7 ± 0.1 (b), 175.3 ± 33 (c), 21.9 ± 3 (d).

for FFDM and 2.3 mGy for DBT. The ESAK measurements were carried out using the “Multimeter PTW UNIDOS” within the ionization chamber PTW Freiburg TW77337-0080 placed at the “reference point” (i.e., 6 cm from the chest wall and 4.5 cm over the detector).²³ The angular independence of the ionization chamber response was verified.

Image Quality Evaluation

Two different representations of the same object were obtained with the two techniques. The image quality comparison was conducted on the 2D images obtained directly by FFDM versus the

DBT image of the slice containing the holes. According to the scope of this paper (to compare the performance of a commercially available digital mammographic system working in 2D versus 3D mode) and because of the DBT’s image-intrinsic processed nature, only DICOM “for presentation images” were analyzed. These conditions included, for 2D and 3D, proprietary denoising and histogram equalization algorithms. 3D images were obtained with a filtered back-projection algorithm. The reported results then refer to the real working condition of the system in the clinical setting. The use of different diameters allowed to tentatively consider the resolution of the systems tested. Images were evaluated using an

automatic software analysis tool (CD-Eval). The implemented algorithm^{28–30} takes advantage of the mathematical description of the phantom. The known geometry of the phantom (and the phantom holes' relative position) allowed the correct registration also for the smaller details and permitted to determine the signal and noise with adequate precision. For each disk, a square region centered on its center is selected; its side is fixed for all detail size (5.5 mm). In a “perfect” imager, the signal corresponding to a disk would be represented by the function $H(|\bar{r} - \bar{r}_0|)$, defined as:

$$H(|\bar{r} - \bar{r}_0|) = \begin{cases} h & \text{if } |\bar{r} - \bar{r}_0| \leq R_0 \\ 0 & \text{if } |\bar{r} - \bar{r}_0| > R_0 \end{cases} \quad (1)$$

Where $\bar{r}$ is the 2D coordinate vector, $\bar{r}_0$ is the vector pointing to the disk center, R_0 is the disk radius, and h is the signal value. A digital version of this function is generated, taking into account the partial volume effect. This latter function is fitted to the 2D experimental data of the selected region by minimizing the χ^2 function. The searching is done using a non-linear least square routine based on the conjugate gradient algorithm. The starting point for the searching is the central pixel of the region where, in principle, the disk center should be located. Registration inaccuracies move, in practice, the disk center a few pixels away from this position. When the minimum is reached, the fitted h value represents the disk average signal level. In the real systems, where blurring occurs, this value is somewhat lower than the maximum value. The fit accuracy is controlled by the χ^2 value and by the residual analysis. The signal difference to noise ratio (SDNR) is defined by^{14,20–32} where S_t is the average pixel intensity of the target, S_{bg} and σ_{bg} are, respectively,

$$SDNR_{d,c} = \frac{S_{bg} - S_t}{\sigma_{bg}} \quad (2)$$

the average pixel intensity and the standard deviation of the background pixel intensity and subscript d and c are referred to target diameter and contrast. The background standard deviation was evaluated in the regions surrounding the evaluated detail. The area of this region was about 25 mm² for all of the details. A sample of the

results obtained with the CD-Eval algorithm is reported in Figure 2, where, for the largest detail, manually and automatically calculated contrasts (SDNR) are compared. Displayed results refer to the planar acquisition with the homogeneous phantom (Fig. 1b). The whole results (Table 1) were summarized by averaging all the SDNRs for a fixed diameter, where d and c are referred to target diameter and contrast and

$$ISDNR_d = \frac{\sum_{c=1}^n SDNR_{d,c}}{n} \quad (3)$$

is the number of details per diameter (7).

For DBT, to evaluate the image resolution in the Z plane (i.e., in the direction perpendicular to the detector plane) and to select the central DBT slice for SDNR evaluation, we calculated the artifact spread function (ASF) as proposed by Zhang et al.^{20,14} where z_0 is the location of the in-focus plane of the target and z is the location of a plane of interest (off-focus plane).

$$ASF(z) = \frac{SDNR(z)}{SDNR(z_0)} \quad (4)$$

Statistical Analysis

Distribution characteristics of SDNR and ISDNR measures were evaluated by using the Kolmogorov–Smirnov one-sample test for normality.³³ SDNRs of the same detail (corresponding to a definite diameter and contrast for 16 different realizations) obtained with DBT was compared with SDNRs obtained with FFDM using Student's two-sided paired t (using the paired samples t test with a two-sided $P < 0.05$ being statistically significant). Two different t tests were performed. The first was applied to the SDNR values for both modalities and each diameter/contrast. The second one was applied to ISDNR values. The statistical computations were carried out using SPSS 15.0 (SPSS, Chicago, IL, USA).

RESULTS

All the SDNR distributions passed the Kolmogorov–Smirnov normality test. The ISDNR values passed the normality test, too. A sample of the

Fig 2. Comparison between the SDNR measures obtained with the CD-Eval algorithm and manual calculation for the largest detail. These results are referred to the planar acquisition with the homogeneous phantom. The error bars correspond to ± 1 standard error.

results is shown in Figure 3, where FFDM and DBT SDNR are compared. The whole results are reported in Table 1. Negative t values mean that DBT was better than FFDM. The significance levels are also reported. In order to make more intelligible results presented in Table 1, the ISDNR

values in function of target diameters were reported in Figure 4 with their relative P values. Higher ISDNRs mean that DBT performed better than FFDM and these differences were statistically significant for all the details' diameters with the exception of the smallest one (0.177 mm).

Table 1. Results of Paired Samples t Test for the 49 Targets

DBT vs FFDM							
Diameter range	4.292 mm	2.524 mm	1.485 mm	0.873 mm	0.513 mm	0.302 mm	0.177 mm
Radiological contrast range							
4.5%							
t	-3.793	-2.008	-4.073	-4.815	-2.511	-2.043	-1.379
p	0.002	0.063	0.001	0.0002	0.024	0.059	0.188
2.8%							
t	-5.114	-4.073	-4.608	-3.733	-3.733	-2.9	-2
p	0.0001	0.001	0.0003	0.002	0.002	0.011	0.064
2.0%							
t	-1.476	-1.424	-2.857	-2.333	-1.644	-3.286	-2.397
p	0.161	0.175	0.012	0.034	0.121	0.005	0.03
1.2%							
t	-0.138	-0.257	0.5	-1.644	0.272	1.899	2.101
p	0.892	0.801	0.624	0.121	0.789	0.077	0.053
0.8%							
t	-1.457	-1.067	-0.851	-2.38	-1.178	-2.999	-0.82
p	0.166	0.303	0.408	0.031	0.257	0.009	0.425
0.5%							
t	-2.495	1.271	-0.891	1.331	-0.891	1.189	0.097
p	0.025	0.223	0.387	0.203	0.387	0.253	0.924
0.3%							
t	-2.15	-1.525	-2.78	1.72	-0.77	0.338	0.95
p	0.048	0.148	0.014	0.106	0.453	0.74	0.357

Negative t values mean that DBT was better than FFDM. The significance levels are also reported (italic values are statistically significant).

Fig 3. Comparison of FFDM versus DBT results in terms of SDNRs for three different target diameters (4.292, 0.873, and 0.177 mm). The error bars correspond to ± 1 standard error.

DISCUSSION

In Table 1, it is worth noting that the SDNR values for the DBT images were higher than the FFDM's for 38 out of 49 details. The DBT's SDNR values were higher for details with diameters greater than 1.485 mm over the entire contrast range covered by the phantom. For details with diameters smaller than 1 mm, the DBT SDNR values remained higher than the FFDM's only for higher contrasts. These differences were statistically significant for 19 details out of 38, probably due to the complex paradigm chosen (breast-simulating material superimposed to signal). For the smallest detail, the presence of negative SDNR values probably depended on the mixing effect of the complex background (breast tissue) and the resolution of the system (with DBT reconstructed pixel size very close to the detail diameter). The ISDNR results showed a dependence on the size of the objects. The ISDNR values in function of target diameters showed higher values for DBT

with respect to FFDM. The differences for ISDNR are statistically significant for all the details' diameters with the exception of the smallest one (0.177 mm).

The results discussed above are somewhat depending on the proprietary processing algorithms: this is, however, an unavoidable consequence when dealing with commercial systems and "for presentation" images. The planar configuration of the contrast detail slab did not perfectly mimic a real breast. An actual 3D detail distribution into the phantom would be of great interest and the presence of structured noise, also in the detail plane, could make more realistic the DBT versus FFDM comparison.

CONCLUSIONS

These results suggested that the complex background plays a less important role in lesion conspicuity for DBT than for FFDM. This paper should offer an initial framework for a global image quality evaluation, which takes into account the mixing effects of noise, contrast, and spatial system resolution, typical of a real imaging system. To apply the conclusions of this study to clinical practice, several issues should be addressed, such as further perception studies that are needed to point out any conclusion on tomosynthesis benefits in the visualization of simulated pathologic lesions with respect to FFDM.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Andy Smith, Dr. Ren Baorui, and Eng. Enrico Tedesco for the cooperation in the experiment and helpful scientific discussion. They also thank Mrs. Roseanne Smith for English revision.

Fig 4. Comparison of FFDM versus DBT results in terms of ISDNRs. Significance levels of paired samples *t* test are reported for each target diameter (**bold values** are statistically significant). The error bars correspond to ± 1 standard error.

REFERENCES

1. U.S. Food and Drug Administration: MQSA National Statistics Archive. <http://www.fda.gov/cdrh/mammography/archives/0108>
2. Yaffe MJ, Mainprize JG, Jong RA: Technical developments in mammography. *Health Phys* 95:599–611, 2008
3. Nitrosi A, Borasi G, Nicoli F, Modigliani G, Botti A, Bertolini M, Notari P: A filmless radiology department in a full digital regional hospital: quantitative evaluation of the increased quality and efficiency. *J Digit Imaging* 20:140–148, 2007
4. Saunders Jr, RS, Samei E, Jesneck JL, Lo JY: Physical characterization of a prototype selenium-based full field digital mammography detector. *Med Phys* 32:588–599, 2005
5. Vedantham S, Karellas A, Suryanarayanan S, Albagli D, Han S, Tkaczyk EJ, Landberg CE, Opsahl-Ong B, Granfors PR, Levis I, D'Orsi CJ, Hendrick RE: Full breast digital mammography with an amorphous silicon-based flat panel detector: physical characteristics of a clinical prototype. *Med Phys* 27:558–567, 2000
6. Ghetti C, Borriani A, Ortenzia O, Rossi R, Ordóñez PL: Physical characteristics of GE Senographe Essential and DS digital mammography detectors. *Med Phys* 35:456–462, 2008
7. Suryanarayanan S, Karellas A, Vedantham S: Physical characteristics of a full-field digital mammography system. *Nucl Instrum Methods Phys Res A* 533:560–570, 2004
8. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, Conant EF, Fajardo LL, Bassett L, D'Orsi C, Jong R, Rebner M: Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 353:1773–1783, 2005
9. Skaane P, Hofvind S, Skjennald A: Randomized trial of screenfilm versus full-field digital mammography with soft-copy reading in population-based screening program: follow-up and final results of Oslo, II study. *Radiology* 244:708–717, 2007
10. Ruschin M, Timberg P, Båth M, Hemdal B, Svahn T, Saunders RS, Samei E, Andersson I, Mattsson S, Chakrabort DP, Tingber AM: Dose dependence of mass and microcalcification detection in digital mammography: free response human observer studies. *Med Phys* 34:400–407, 2007
11. Burgess AE, Jacobson FL, Judy PF: Human observer detection experiments with mammograms and power-law noise. *Med Phys* 28:419–437, 2001
12. Niklason LT, Christian BT, Niklason LE, Kopans DB, Castleberry DE, Opsahl-Ong BH, Landberg CE, Slanetz PJ, Giardino AA, Moore R, Albagli D, DeJule MC, Fitzgerald PF, Fobare DF, Giambattista BW, Kwasnick RF, Liu J, Lubowski SJ, Possin GE, Richotte JF, Wei CY, Wirth RF: Digital tomosynthesis in breast imaging. *Radiology* 205:399–406, 1997
13. Wu T, Stewart A, Stanton M, McCauley T, Phillips W, Kopans DB, Moore RH, Eberhard JW, Opsahl-Ong B, Niklason L, Williams MB: Tomographic mammography using a limited number of low-dose cone-beam projection images. *Med Phys* 30:365–380, 2003
14. Wu T, Moore RH, Rafferty EA, Kopans DB: A comparison of reconstruction algorithms for breast tomosynthesis. *Med Phys* 31:2636–2647, 2004
15. Wu T, Moore RH, Kopans DB: Voting strategy for artifact reduction in digital breast tomosynthesis. *Med Phys* 33:2461–2471, 2006
16. Andersson I, Ikeda DM, Zackrisson S, Ruschin M, Svahn T, Timberg P, Tingberg A: Breast tomosynthesis and digital mammography: a comparison of breast cancer visibility and BIRADS classification in a population of cancers with subtle mammographic findings. *Eur Radiol* 18:2817–2825, 2008
17. Suryanarayanan S, Karellas A, Vedantham S, Glick SJ, D'Orsi CJ, Baker SP, Webber RL: Comparison of tomosynthesis methods used with digital mammography. *Acad Radiol* 7:1085–1097, 2000
18. Diekmann F, Meyer H, Diekmann S, Puong S, Muller S, Bick U, and Rogalla P: Thick slices from tomosynthesis data sets: phantom study for the evaluation of different algorithms. *J Digit Imaging* 22:519–526, 2009
19. Diekmann F, Bick U: Tomosynthesis and contrast-enhanced digital mammography: recent advances in digital mammography. *Eur Radiol* 17:3086–3092, 2007
20. Zhang Y, Chan HP, Sahiner B, Wei J, Goodsitt MM, Hadjiiski LM, Ge J, Zhou C: A comparative study of limited-angle cone-beam reconstruction methods for breast tomosynthesis. *Med Phys* 33:3781–3795, 2006
21. Rakowski JT, Dennis MJ: A comparison of reconstruction algorithms for C-arm mammography tomosynthesis. *Med Phys* 33:3018–3032, 2006
22. American College of Radiology. <http://www.acr.org/SecondaryMainMenuCategories/quality_safety/BIRADSAtlas/BIRADSAtlasexcerptedtext/BIRADSMammographyFourthEdition/GuidanceChapterDoc6.aspx>
23. European Guidelines for Quality Assurance in Breast Cancer screening and Diagnosis, 4th edition. Luxembourg: CEC, 2006
24. Park JM, Franken Jr, EA, Garg M, Fajardo LL, Niklason LT: Breast tomosynthesis: present considerations and future applications. *RadioGraphics* 27:S231–S240, 2007
25. Saunders Jr, RS, Samei E: The effect of breast compression on mass conspicuity in digital mammography. *Med Phys* 35:4464–4473, 2008
26. Chida K, Komatsu Y, Sai M, Nakagami A, Yamada T, Yamashita T, Mori I, Ishibashi T, Maruoka S, Zuguchi M: Reduced compression mammography to reduce breast pain. *Clin Imaging* 33:7–10, 2009
27. Toroi P, Zanca F, Young KC, van Ongeval C, Marchal G, Bosmans H: Experimental investigation on the choice of the tungsten/rhodium anode/filter combination for an amorphous selenium-based digital mammography system. *Eur Radiol* 17:2368–2375, 2007
28. Borasi G, Nitrosi A, Ferrari P, Tassoni D: On site evaluation of three flat panel detectors for digital radiography. *Med Phys* 30:1719–1731, 2003
29. Borasi G, Samei E, Bertolini M, Nitrosi A, Tassoni D: Contrast-detail analysis of three flat panel detectors for digital radiography. *Med Phys* 33:1707–1719, 2006
30. Rivetti S, Lanconelli N, Campanini R, Bertolini M, Borasi G, Nitrosi A, Danielli C, Angelini L, Maggi S: Comparison of different commercial FFD units by means of

physical characterization and contrast-detail analysis. *Med Phys* 33:4198–4209, 2006

31. Samei E, Dobbins III, JT, Lo JY, Tornai MP: A framework for optimising the radiographic technique in digital x-ray imaging. *Radiat Prot Dosim* 114:220–229, 2005

32. Tapiovaara MJ, Wagner RF: SNR and noise measurements for medical imaging: I. A practical approach based on statistical decision theory. *Phys Med Biol* 38:71–92, 1993

33. Siegel S: *Nonparametric Statistical Methods for the Behavioral Sciences*, New York: McGraw-Hill, 1956