Application of breast tomosynthesis in screening: incremental effect on mammography acquisition and reading time

[1]D BERNARDI, MD, [1]S CIATTO, MD, [1]M PELLEGRINI, MD, [1]V ANESI, TSRM, [1]S BURLON, TSRM, [1]E CAULI, TSRM, [1]M DEPAOLI, TSRM, [1]L LARENTIS, TSRM, [1]V MALESANI, TSRM, [1]L TARGA, TSRM, [1]P BALDO, TSRM and [2]N HOUSSAMI, MB BS, PhD

[1]*U O Senologia Clinica e Screening Mammografico, Dipartimento di Radiodiagnostica, APSS, Trento, Italy, and* [2]*Screening and Test Evaluation Program, School of Public Health, Sydney Medical School, University of Sydney, Sydney, Australia*

Objective: The aim of this study was to supplement the paucity of information available on logistical aspects of the application of three-dimensional (3D) mammography in breast screening.
Methods: We prospectively examined the effect on radiographers' and radiologists' workload of implementing 3D mammography in screening by comparing image acquisition time and screen-reading time for two-dimensional (2D) mammography with that of combined 2D+3D mammography. Radiologists' accuracy was also calculated.
Results: Average acquisition time (measured from start of first-view breast positioning to compression release at completion of last view) for seven radiographers, based on 20 screening examinations, was longer for 2D+3D (4 min 3 s; range 3 min 53 s–4 min 18 s) than 2D mammography (3 min 13 s; range 3 min 0 s–3 min 26 s; $p<0.01$). Average radiologists' reading time per screening examination (three radiologists reading case-mix of 100 screens: 10 cancers, 90 controls) was longer for 2D+3D (77 s; range 60–90 s) than for 2D mammography (33 s; range 25–46 s; $p<0.01$). 2D+3D screen-reading was associated with detection of more cancers and with substantially fewer recalls than 2D mammography alone.
Conclusion: Relative to standard 2D mammography, combined 2D+3D mammography prolongs image acquisition time and screen-reading time (at initial implementation), and appears to be associated with improved screening accuracy.
Advances in knowledge: These findings provide relevant information to guide larger trials of integrated 3D mammography (2D+3D) and its potential implementation into screening practice.

Received 22 December 2011
Revised 23 April 2012
Accepted 21 June 2012

DOI: 10.1259/bjr/19385909

Mammography screening has been shown to be effective in reducing breast cancer mortality [1–2], and population-based screening is currently recommended and implemented in most developed healthcare systems. Still, screening has limitations in sensitivity and specificity, with many of these dependent on the masking effect, or superimposition, of dense breasts [3–7].

New techniques less adversely affected by breast density and tissue superimposition might improve screening sensitivity and specificity, and potentially cost-effectiveness. Recent experience suggests that three-dimensional (3D) mammography with tomosynthesis may improve diagnostic accuracy compared with two-dimensional (2D) imaging [8–15], although available studies thus far have analysed relatively small and mostly heterogeneous series. Therefore, large prospective studies are needed to determine the role of 3D mammography in screening, such as the NHS ''TOMMY trial'' (A comparison of TOMosynthesis with digital MammographY in the UK NHS Breast Screening Programme; http://www.hta.ac.uk/2296). While initial studies of 3D mammography have looked at its accuracy, to our knowledge there are no data on logistical aspects of the application of 3D mammography that include its effect on acquisition and reading time. Because 3D mammography may require longer acquisition and/or reading time than 2D mammography, its impact on workload of radiographers and radiologists is a crucial issue for its potential implementation in screening.

The purpose of the present study is to define the magnitude of the impact on radiographers' and radiologists' workload (acquisition time and reading time) associated with implementing 3D mammography in population screening practice.

Methods and materials

We compared workload (measured for acquisition time and radiologists' reading time) associated with 2D mammography relative to *combined* 2D+3D mammography. Both 2D and 3D images were acquired using a Selenia Dimensions Unit (Hologic, Bedford, MA),

Address correspondence to: Associate Professor Nehmat Houssami, Screening & Test Evaluation Program, School of Public Health (A27), Sydney Medical School, University of Sydney, Australia. E-mail: nehmath@med.usyd.edu.du
Associate Professor Houssami is supported by National Health and Medical Research Council program grant 633003 to the Screening and Test Evaluation Program.

operating at the Breast Diagnosis Department of Trento, Italy. The COMBO® procedure (Hologic, Bedford, MA) was used, which acquires both 2D and 3D images with a single breast positioning and compression. Standard bilateral two-view (craniocaudal and mediolateral oblique) mammography was used. The study was approved via the ethics approval process in the participating centre.

Comparing acquisition time of 2D with 2D+3D mammography

Acquisition times of 2D and 2D+3D mammography were measured from the start of first view breast positioning to compression release at completion of last view acquisition. Overall ''door to door'' time was not considered, as it is quite variable, depending on local organisation and structural layout (*e.g.* double separate dressing box availability) and independent of imaging modality effect. The study involved seven mammography-dedicated radiographers, each with a minimum of 2 years' experience, who also had had a short but intensive (2–3 days) training in performing 3D mammography. Absolute (minutes, seconds) and relative (% incremental) differences in acquisition time were determined in two compared sessions (2D and 2D+3D) that included 20 women each. Women were unselected screening participants, after exclusion of those with major physical disabilities.

Comparing reading time of 2D with 2D+3D mammography

10 cancers and 90 negative controls (non-cancers) were randomly selected from the 2D+3D image archive for the study purpose. Two sets (2D and 2D+3D) of 100 cases were prepared and then divided into subsets of 25 cases. 2D and 2D+3D sets included the same screens, but cases in subsets were mixed and presented in a varied order. 2D sets were read first, and 2D+3D sets were read in a separate session, 3–7 days after 2D set reading. Subsets were displayed on a Hologic reading workstation; the hanging protocols are shown in Table 1. Conventional image-processing tools were available to the reader (black/white inversion, magnifying lens, brightness and contrast tuning).

Table 1. Hanging protocol for 2D and 2D+3D images for screen reading

Image set	Monitor right	Monitor left
2D		
1	MLO right/left, fronted	CC right/left, fronted
2	MLO right	MLO left
3	CC right	CC left
2D+3D		
1	MLO 2D right/left, fronted	CC 2D right/left, fronted
2	MLO 3D right/left, fronted	CC 3D right/left, fronted
3	MLO 2D right	MLO 2D left
4	CC 2D right	CC 2D left
5	MLO 3D right	MLO 3D left
6	CC 3D right	CC 3D left

2D, two-dimensional; 3D, three-dimensional; cc, craniocaudal; MLO, mediolateral-oblique.

Each subset of 25 cases was read in a single continuous session, and reading time was computed from the display of the first case to the last case reporting. Reporting was performed using standardised predefined forms, indicating the side and site of any abnormality warranting further assessment. Absolute (minutes, seconds) and relative (% incremental) differences in reading time were determined for each of the 2D and 2D+3D sets. Three dedicated breast radiologists with, respectively, 12, 22 and 38 years' experience in mammography were involved in screen reading, after receiving a training course in 3D mammography reading; each of these radiologists had been reporting 3D mammography for a period of 2–8 months at time of participating in the study.

Although not the primary purpose of the study, we also calculated accuracy (sensitivity and specificity) for each reader, and for 2D and 2D+3D readings, as a proxy measure of ''attention'' of the readers; this is because the reading sessions were simulated on an archival set and the risk of only ''virtual'' errors might bias readers favouring less accurate (and thus faster) reading time. It also allows an approximate estimation of relative accuracy of 2D and combined 2D+3D screen reading.

Differences between 2D and 2D+3D imaging as to acquisition (average of 20 acquisitions, for each and all radiographers) and reading time (average of 100 readings, for each and all radiologists) were calculated (Student's *t*-test), setting statistical significance at $p<0.05$.

Results

Acquisition time

Table 2 shows acquisition times recorded for the seven radiographers involved. Average acquisition time for each 2D mammographic examination was 3 min 13 s (range 3 min 0 s–3 min 26 s), whereas that of 2D+3D mammography was 4 min 3 s (range 3 min 53 s–4 min 18 s). 2D+3D mammography required an average of 49 s (range 40–60 s), representing 26% (range 19–31%) additional time to be acquired compared with 2D mammography, a statistically significant difference ($p<0.01$).

Reading time

Table 3 shows reading times for three radiologists (2D and 2D+3D), and the average screen-reading time (data are shown both per screen read and for total screens in the study test set). Average 2D mammography reading time *per screening examination* was 33 s (range for three radiologists 25–46 s), whereas that of 2D+3D mammography was 77 s *per screening examination* (range for three radiologists 60–90 s). 2D+3D mammography required 44 s more reading time *per screening examination* (range for three radiologists 32–54 s), representing an average +135% screen-reading time compared with 2D mammography alone, a statistically significant difference ($p<0.01$). The results for the total test set of 100 screens were very similar (Table 3).

Table 2. Acquisition time of 2D or combined 2D+3D mammography by seven radiographers

Acquisition time	Radiographer							Average time
	A	B	C	D	E	F	G	
2D, average time per screen (range)	3 min 28 s (3 min 1 s–3 min 49 s)	3 min 19 s (2 min 52 s–3 min 54 s)	3 min 5 s (2 min 33 s–3 min 30 s)	3 min 13 s (2 min 33 s–4 min 30 s)	3 min 1 s (2 min 37 s–3 min 28 s)	3 min 14 s (2 min 47 s–3 min 57 s)	3 min 13 s (2 min 47 s–3 min 46 s)	3 min 13 s (3 min 0 s–3 min 26 s)
3D, average time per screen (range)	4 min 8 s (3 min 46 s–4 min 39 s)	4 min 2 s (3 min 43 s–4 min 49 s)	3 min 51 s (3 min 20 s–4 min 32 s)	4 min 13 s (3 min 42 s–5 min 37 s)	3 min 48 s (3 min 21 s–4 min 22 s)	4 min 5 s (3 min 35 s–4 min 40 s)	4 min 13 s (3 min 51 s–4 min 47 s)	4 min 3 s (3 min 53 s–4 min 18 s)
Absolute difference in time (2D+3D *vs* 2D)	40 s	42 s	45 s	60 s	47 s	51 s	60 s	49 s
Difference (2D+3D *vs* 2D)	+19%	+21%	+25%	+31%	+26%	+26%	+31%	+26%

2D, two-dimensional; 3D, three-dimensional.

Accuracy

At 2D reading each of three radiologists identified 8 of 10 cancers in the set (total 24 cancers detected by all reads from three radiologists), and recalled for abnormalities 10, 16 or 7 negative controls, respectively. Using 2D+3D reading, there were 10, 10 or 9 true-positive cancers detected, respectively (total of 29 cancers detected by all reads from three radiologists), an overall incremental detection of +20.8%. Recall for an abnormality in negative controls using 2D+3D was 3, 9 or 2, respectively, with a recall rate reduction (all control readings) of −57.5%.

Discussion

We report an evaluation of the effect of introducing 3D mammography (breast tomosynthesis) on acquisition time and reading time for screening mammography, finding that both measures were prolonged with integration of 3D imaging (adding 3D to conventional mammography). The study is based on a sufficient number of screening examinations to allow a reliable estimate of differences in (radiographer) acquisition time and in radiologists' reading time of combined 2D+3D compared with 2D mammography. It is reasonable to expect that adding (rather than substituting) further image acquisition and interpretation would increase workload (based on time measures). The primary intention of the study was therefore to establish the magnitude of effect on both acquisition and reading time, especially in the relatively early phase of implementation of 3D mammography in established screening services.

It could be argued that 3D mammography training of radiographers and radiologists, prior to conducting the study, was relatively brief. There is not an issue for radiographers' training, as the only difference between 2D and 2D+3D mammography is related to 3D image acquisition and reconstruction, which does not directly involve the radiographer (and hence does not require special training); radiographers' training essentially related to switching the machine from the 2D to the COMBO modality. Radiologists' training, however, was based on a minimum review of 100 3D images, so we cannot exclude that the reading time may improve (reduce) with longer experience in 3D interpretation. However, it has been our experience that long training is not needed for expert breast radiologists dedicated to mammography reading, and the learning curve may be expected to be modest for combined 2D/3D as it is a *mammography-based* technology. We also examined (in a related evaluation of tomosynthesis) radiologists' reading time at 6 months further experience with 3D images and found minor changes (essentially slight reductions in average reading time for 2D+3D). Furthermore, the workstation software is simple and easy to manage: in the USA the manufacturer recommends 8 h of training before managing and reporting 3D images in practice, while in the present experience, training in managing and reporting 3D images was at least 10 times longer (based on training and several months application). Even if a learning curve in interpreting 3D reconstructed images to achieve optimal accuracy is assumed, this is likely to have less of an effect on reading time, as for 2D

Table 3. Reading time for 2D and 2D+3D mammography by three radiologists and the average screen-reading time (data are shown both per screen read and for total screens)

Time based on (by modality and whether individual or total screens)	Radiologist			Average reading time
	A	B	C	
Results for time per individual screen-read (per test)				
2D mammography	25 s	27 s	46 s	33 s
2D+3D mammography	80 s	60 s	90 s	77 s
Absolute difference	54 s	33 s	44 s	44 s
Difference (2D+3D *vs* 2D)	+220%	+122%	+96%	+133%
Results for time per total screens in test set (100 screens)				
2D mammography	41 min 28 s	45 min 32 s	76 min 10 s	54 min 23 s
2D+3D mammography	133 min 6 s	100 min 11 s	150 min 16 s	127 min 51 s
Absolute difference	91 min 38 s	54 min 39 s	74 min 6 s	73 min 28 s
Difference (2D+3D *vs* 2D)	+220%	+120%	+97%	+135%

2D, two-dimensional; 3D, three-dimensional.

mammography, where reading time stabilises after a relatively short period, while individual differences in accuracy persist over a longer time [16].

2D+3D mammography was associated with a modest (though statistically significant) increase in acquisition time. The study design considered only the crude patient positioning and image acquisition time, and the observed +26% excess for 2D+3D might be further reduced with respect to overall ''door to door'' time, including time for undressing/dressing, which is variable. Overall, time to perform 2D+3D mammography would be expected to exceed that of 2D mammography, and will vary with the mammography unit used—we had a relatively short acquisition time for 3D images (4 s per view, over a 15° angle). Acquisition time may be higher with other machines that use longer scanning time (up to 15 s per view).

When measuring radiologists' reading time we considered combined 2D+3D image reading relative to standard 2D mammography reading. We intentionally did not use 3D image reading alone because (based on current knowledge and screening practice) panoramic analysis of 2D images is essential prior to 3D image analysis. In our setting, radiologists are as yet developing expertise in 3D diagnostic pattern definition, and the superiority of 2D+3D imaging compared with 3D imaging alone has been reported [17–18]; hence, we maintained the standard screen reading of 2D and evaluated the addition of 3D to 2D reading. It may be, nevertheless, that 3D-only mammography has a future role, possibly using a ''synthetic'' reconstruction as a 2D imaging surrogate, but this needs to be evaluated in appropriately designed trials, and needs to be supported by enhanced knowledge and consolidation of information from 3D diagnostic patterns. Deliberation over the almost double exposure dose delivered in the COMBO procedure (though still within acceptable exposure values), while not the purpose of the present study, is nonetheless a relevant issue that might prompt exploration of the use of 3D-only imaging in breast screening in future studies.

Our findings indicate that reading time is substantially different for 2D *vs* 2D+3D mammography. As pointed out earlier, this is not entirely surprising, as the time measured was for *combined* 2D+3D mammography, and because 3D images are reviewed by scrolling through different breast planes (similar to looking at a movie). This does not allow a panoramic view of the whole breast as with 2D mammography: 3D imaging requires separate analysis of different parts of each breast view (at least two segments), and this must be repeated for all four views. An increase in reading time for 2D+3D mammography has also been noted in other studies [19–20]: based on the data reported in each of these studies, for Good et al [19] we estimated a +172% incremental reading time, and for Gur et al [20] we estimated +96% incremental reading time, and our data showed a relatively intermediate value (+133%) for incremental reading time for combined 2D+3D mammography. This aspect of 3D imaging is particularly unique to 3D mammography reading in screening practice, where the whole breast must be examined in search of abnormalities that may have been missed at 2D mammography. 3D mammography use for targeted analysis (for example, as a triage to further assessment where abnormalities have been identified at 2D mammography) is obviously less demanding in terms of reading time.

Based on present study findings, mammography screen reading using 2D+3D mammography more than doubled (on average) the reading time (see Table 3), and the magnitude of the effect was similar for participating radiologists. In a 3D mammography screening scenario, even if only used for a subgroup of subjects (*e.g.* with dense breast, therefore at higher risk of false-positives due to superimposition or masking), the radiologist's workload will be substantially increased with prolonged reading time. For the current population screening scenario (age 50–69 years), 2D+3D mammography screening of subjects with Breast Imaging-Reporting and Data System (BI-RADS) D3–D4 dense breast (around 25% of this population [21]), assuming an incremental reading time of +100–200%, radiologists' reading time would be expected to increase by approximately 25–50%. Alternatively, extending 3D mammography screening to BI-RADS D2–4 subjects (around 70% of screening participants in targeted age group) will increase radiologists' reading time in the range of 70–140%. Although combined 2D+3D breast screening is not currently proposed, in the absence of high-level scientific evidence supporting its efficacy in screening, we recommend that future screening evaluations of 3D mammography consider the additional time demands in both feasibility and cost-effectiveness studies. In breast screening, computer-aided detection (CAD) may prove a useful

complementary tool in the future; however, at this stage, development and application of 3D mammography CAD is still in its infancy.

We have pointed out that reader accuracy was *not* a primary objective of this evaluation, but was integrated into the study as a measure of reader attention, and also provides an indication of *relative* accuracy. It is therefore worth noting (but with cautious interpretation of these data, which have not excluded a learning effect from repeat reads) that an increased sensitivity was observed using 2D+3D mammography (relative to 2D alone), and this was associated with a substantial increase in specificity (reduction in recalls), also shown in other studies [8–15]. The striking reduction in the numbers recalled for assessment based on 2D+3D, in particular, should be ''tested'' in large-scale studies, as this has the potential to reduce the burden of false-positive screens—an issue of considerable relevance to both screening participants and screening services.

We conclude that integrating 3D with 2D mammography in breast screening prolongs image acquisition and screen-reading times, relative to 2D mammography alone, and the effect is more marked for radiologists' reading time (at least in the implementation phase of this new technology). The prolonged reading time might reduce with further experience, and may be acceptable if associated with enhanced screening accuracy (yet to be demonstrated in large trials), particularly reduced recalls, and warrants evaluation in prospective trials of population breast screening.

References

1. Glasziou P, Houssami N. The evidence base for breast cancer screening. Prev Med 2011;53:100–2.
2. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammography. A meta-analysis. J Am Med Ass 1995;273:149–54.
3. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 7825 patient evaluations. Radiology 2002;225:165–75.
4. Ciatto S, Visioli C, Paci E, Zappa M. Breast density as a determinant of interval cancer at mamographic screening. Br J Cancer 2004;90:393–6.
5. Ciatto S, Brancato B. Role of multimedial diagnosis of breast cancer in women below 36 years of age. Radiol Med 2005;109:321–9.
6. Laya MB, Larson EB, Taplin SH, White E. Effect of estrogen replacement therapy on the specificity and sensitivity of screening mammography. J Natl Cancer Inst 1996;88: 643–9.
7. Lehman CD, White E, Peacock S, Drucker MJ, Urban N. Effect of age and breast density on screening mammograms with false positive findings. AJR Am J Roentgenol 1999;173:1651–5.
8. Park JM, Franken EA Jr, Garg M, Fajardo LL, Niklason LT. Breast tomosynthesis: present considerations and future applications. Radiographics 2007;27:S231–40.
9. Niklason LT, Christian BT, Niklason LE, Kopans DB, Castleberry DE, Opsahl-Ong BH, et al. Digital breast tomosynthesis in breast imaging. Radiology 1997;205:25–31.
10. Rafferty EA, Georgian-Smith D, Kopans DB, Hall DA, Moore R, Wu T. Comparison of full-field digital tomosynthesis with two view conventional film screen mammography in the prediction of lesion malignancy. Radiology 2002;225:268.
11. Teertstra HJ, Loo CE, van den Bosch MA, van Tinteren H, Rutgers EJ, Muller SH, et al. Breast tomosynthesis in clinical practice: initial results. Eur Radiol 2010;20:16–24.
12. Mitka M. New screening methods offer hope for more accurate breast cancer detection. J Am Med Ass 2008; 299:397–8.
13. Gur D. Tomosynthesis: potential clinical role in breast imaging. AJR Am J Roentgenol 2007;189:614–15.
14. Poplack SP, Tosteson TD, Kogel CA, Nagy HM. Digital breast tomosynthesis: initial experience in 98 women with abnormal digital screening mammography. AJR Am J Roentgenol 2007;189:616–23.
15. Gur D, Zuley ML, Anello MI, Rathfon GY, Chough DM, Ganott MA, et al. Digital breast tomosynthesis: observer performance study. AJR Am J Roentgenol 2009;193:586–91.
16. Ciatto S, Andreoli C, Di Maggio C. Proficiency test in clinical mammography. Results of a consecutive series of 130 volunteer Italian radiologists. [In Italian.] Radiol Med 1999;98:255–8.
17. Andersson I, Ikeda DM, Zackrisson S, Ruschin M, Svahn T, Timberg P, et al. 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 2008;18: 2817–25.
18. Rafferty EA, Smith AP, Niklason LT. Assessing radiologist performance in dense versus fatty breasts using combined full-field digital mammography and breast tomosynthesis compared to full-field digital mammography alone. Radiologic Society of North America 95th Scientific Assembly and Annual Meeting; Chicago, IL. 2009. Available online at http://rsna2009.rsna.org (accessed December 2011).
19. Good WF, Abrams GS, Catullo VJ, Chough DM, Ganott MA, Hakim CM, et al. Digital breast tomosynthesis: a pilot observer study. AJR Am J Roentgenol 2008;190:865–9.
20. Gur D, Abrams GS, Chough DM, Ganott MA, Hakim CM, Perrin RL, et al. Digital breast tomosynthesis: observer performance study. AJR Am J Roentgenol 2009;193:586–91.
21. Corsetti V, Houssami N, Ghirardi M, Ferrari A, Speziani M, Bellarosa S, et al. Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: Interval breast cancers at 1year follow-up. Eur J Cancer 2011;47:1021–6.