Radiology

Mammographic Breast Density: Comparison of Methods for Quantitative Evaluation[1]

Oliver W. E. Morrish, MSc
Lorraine Tucker, DCR
Richard Black, MSc
Paula Willsher, DCR
Stephen W. Duffy, MSc
Fiona J. Gilbert, MD

Purpose: To evaluate the results from two software tools for measurement of mammographic breast density and compare them with observer-based scores in a large cohort of women.

Materials and Methods: Following written informed consent, a data set of 36281 mammograms from 8867 women were collected from six United Kingdom centers in an ethically approved trial. Breast density was assessed by one of 26 readers on a visual analog scale and with two automated density tools. Mean differences were calculated as the mean of all the individual percentage differences between each measurement for each case (woman). Agreement in total breast volume, fibroglandular volume, and percentage density was assessed with the Bland-Altman method. Association with observer's scores was calculated by using the Pearson correlation coefficient (r).

Results: Correlation between the Quantra and Volpara outputs for total breast volume was $r = 0.97$ ($P < .001$), with a mean difference of 43.5 cm^3 for all cases representing 5.0% of the mean total breast volume. Correlation of the two measures was lower for fibroglandular volume ($r = 0.86$, $P < .001$). The mean difference was 30.3 cm^3 for all cases representing 21.2% of the mean fibroglandular tissue volume result. Quantra gave the larger value and the difference tended to increase with volume. For the two measures of percentage volume density, the mean difference was 1.61 percentage points ($r = 0.78$, $P < .001$). Comparison of observer's scores with the area-based density given by Quantra yielded a low correlation ($r = 0.55$, $P < .001$). Correlations of observer's scores with the volumetric density results gave r values of 0.60 ($P < .001$) and 0.63 ($P < .001$) for Quantra and Volpara, respectively.

Conclusion: Automated techniques for measuring breast density show good correlation, but these are poorly correlated with observer's scores. However automated techniques do give different results that should be considered when informing patient personalized imaging.

Clinical trial registration no. ISRCTN 73467396

[1] From the East Anglian Regional Radiation Protection Service (O.W.E.M.), Department of Medical Physics and Clinical Engineering (R.B.), and Cambridge Breast Unit (P.W.), Cambridge University Hospitals NHS Foundation Trust, Cambridge, England; Department of Radiology, University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge CB2 0QQ, England (L.T., F.J.G.); and Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, England (S.W.D.).
Received June 26, 2014; revision requested August 21; final revision received September 18; accepted October 2; final version accepted October 10. Supported by the National Institute for Health Research's Health Technology Assessment Programme. **Address correspondence to** F.J.G. (e-mail: *fjg28@cam.ac.uk*).

The estimation and measurement of breast density has been reported in many studies during the past 2 decades (1–9). This has been partly driven by the role of breast density as a strong, independent, and modifiable risk factor of breast cancer (3–5). Most of the work has used mammograms for the evaluation of density where the appearance of the breast reflects variations in the relative amounts of fat, connective tissue, and epithelial tissue and their different x-ray attenuation characteristics (10). Breast density is expressed as a percentage of the mammogram occupied by the fibroglandular and stromal tissue (9).

Assessment of breast density is traditionally performed by a reader on the basis of a visual assessment of the standard two-view mammogram. Consistency of this measure requires an experienced observer to be able to correctly assess the relative proportions of glandular and fatty tissue while accounting for variations in breast shape, size, fibroglandular pattern, and the various radiographic factors used. Density scores are then given either on a continuous scale as a percentage (11) or within discrete ranges, such as the four-point Breast Imaging Reporting and Data System, BI-RADS (12), scale or the Boyd five-point scale (5). Studies suggest that training and experience are essential in ensuring that the scores are accurate and reproducible (12,13).

The introduction of full-field digital mammography technologies has provided an opportunity to implement automated breast density measurement algorithms, which had been initially developed for digitized analog mammograms (11,14). These algorithms work by applying thresholds to the pixel values within the digital image to identify the area of the image that contains the breast and to then determine the proportion of that breast which contains fibroglandular tissue. For example, the pixel values with the highest signal (radiation dose detected by the pixel) can identify the areas of the image where no breast tissue has attenuated the primary x-ray beam. The areas of lowest signal, on the other hand, represent areas where the x-rays have passed through a section of tissue that is relatively most attenuating (15).

Later developments have led to software that estimates the volume of dense fibroglandular tissue rather than just the area represented on the mammogram. By using the image pixel data in combination with information about the acquisition parameters (eg, compression paddle height, x-ray tube potential, target, and filter), more recent algorithms are capable of providing estimates of the relevant tissue volumes by means of derivation of the tissue composition represented by each pixel (16–18).

The researchers in one small study (19) have compared a range of volumetric software tools and found that the density measurements were in substantial agreement with data from breast magnetic resonance (MR) imaging. The aim of this work was to evaluate the results from two software tools for the measurement of mammographic breast density and compare them with observer-based scores in a large cohort of women.

Advances in Knowledge

- When two observers score breast density for the same woman using a visual analog scale, in 70% (449 of 638) of cases (women), their scores agree to within 10% but in 13% (85 of 638) of cases, there is disagreement about whether breast density was greater or less than 50%.
- There is good agreement for overall breast volume when we compare two software tools, with Quantra giving a median value of 953.5 cm^3 (range, 73.0–4986.5 cm^3) and Volpara giving a median value of 921.4 cm^3 (range, 33.4–5009.3 cm^3), but there is less agreement for median fibroglandular tissue volume (Quantra, 93.0 cm^3 [range, 4.0–1024.0 cm^3]; and Volpara, 71.6 cm^3 [range, 6.8–628.5 cm^3]) and median percentage breast density (Quantra, 9.5% [range, 1.4%–56.2%]; and Volpara, 7.7% [range, 2.5%–54.2%]), particularly as density increases, with Volpara having lower values compared with Quantra.

Implications for Patient Care

- Automated breast density measurements are objective and are not subject to observer variability, but different systems do not produce identical results.
- Different software tools and observers' scores result in variation in the number of women classified in the greater than 50% density category.

Materials and Methods

Support for this study was given by two companies—Hologic (Bedford, Mass) and Mātakina Technology (Wellington, New Zealand). Both provided automated breast density software and technical advice. The authors had full access to all data in the study and the information submitted for publication.

Study Data

Standard two-dimensional digital mammograms (n = 36281) were obtained from the TOMosynthesis with digital

Published online before print
10.1148/radiol.14141508 **Content code:** BR

Radiology 2015; 275:356–365

Abbreviations:
CC = cranial caudal
DICOM = Digital Imaging and Communications in Medicine
MLO = mediolateral oblique

Author contributions:
Guarantor of integrity of entire study, F.J.G.; 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, O.W.E.M., F.J.G.; clinical studies, O.W.E.M., P.W., F.J.G.; statistical analysis, O.W.E.M.; and manuscript editing, O.W.E.M., L.T., P.W., S.W.D., F.J.G.

Clinical trial registration no. ISRCTN 73467396

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

MammographY, or TOMMY, trial in the United Kingdom National Health Service Breast Screening Program (20) (clinical trial registration ISRCTN 73467396) by using the full-field digital mammography system (Selenia Dimensions; Hologic) installed in six centers in the United Kingdom. All the systems underwent regular quality control testing to demonstrate ongoing compliance with expected system performance (21). The trial was approved by the Scotland A Research Ethics Committee, and participants gave written informed consent.

Study Cohort

The study cohort included women aged 47–73 years who were recalled to an assessment clinic following abnormal screening mammography results and women aged 40–49 years with a family history of breast cancer who were attending annual screening. The imaging protocol for the trial included acquisition of standard two-view mammograms (cranial caudal [CC] and mediolateral oblique [MLO]) of both breasts, and these images were used for the density assessment. A complete set of images for a total of 8867 individual women were available for analysis.

Reader Assessment

In the study, 26 readers were asked to assess breast density on a visual analog scale, giving a score ranging from 0% to 100%. For each subject, a single score assigned by a single reader was obtained at review of the mammograms, except for subjects with a family history of breast cancer, in whom two readers assigned scores. These readers, who were not authors of this article, consisted of 21 radiologists, two breast physicians, and three advanced practitioner radiographers, each with a minimum of 2 years of experience reading at least 5000 mammograms per year. The visual analog scale is a 10-cm line on a paper form on which readers indicate their density score by making an appropriate mark, with the left end of the line representing 0% and the right end representing 100%. Each center then digitized and processed the forms to extract density scores for each subject on the basis of the position of the reader's mark (11). In four centers, this resulted in a scale of scores given to the nearest whole percentage; however, one center rounded scores to the nearest 5%, and another center rounded scores to the nearest 10%. When the readers were scoring the images, they were advised not to alter the window levels of images.

Density Assessment with Software

Two software packages were used to assess the breast density on each mammogram (Quantra version 2.0, Hologic; and Volpara version 1.4.2, Mātakina Technology) (17,22). The output from each program consisted of values for the absolute volume of fibroglandular tissue and overall breast volume, as well as the volumetric breast density on a per-image basis. In addition, with Quantra, an area-based score was determined.

To obtain the overall score for each examination, the absolute values of total breast volume and fibroglandular tissue volume for each of the four views (left and right, CC and MLO) were examined. For cases (women) in which no cancer was assessed as being present, the largest total breast volume and the largest fibroglandular tissue volume for each breast (either from the CC or MLO view) were separately determined, and the average tissue volumes of the two breasts were calculated. The same logic was applied to the scores determined with the Quantra system for area breast density (which should nominally be comparable with the observer scores).

For cases in which cancer was confirmed with histopathologic examination, results were used from the contralateral breast. We believe this procedure to mirror the behavior of observers. Also, the rejection of data from the affected breast reflects methods used in cancer cohort studies in which mammographic density is assessed from the contralateral mammogram (5,9). If no contralateral data were available, results from the affected breast were used. Volumetric density was calculated from the ratio of the fibroglandular tissue volume to the overall breast volume.

Statistical Analysis

Agreement in the measurements of absolute tissue volumes and density measurements was assessed by using the Bland-Altman method. The mean difference between each measurement was derived by subtracting the relevant Volpara value from the Quantra value and then calculating the mean value of this difference for all cases. The mean percentage difference was derived by expressing the difference for each case as a percentage of the mean of the two measurements and then calculating the mean percentage for all cases. The relationship between the various measurements was assessed by calculating the Pearson correlation coefficient (r). Logarithmic transformation of the data was performed where examination of residual plots indicated that the data would behave more linearly and variance would be more stable when compared with the untransformed data. These analyses were further performed on the fibroglandular tissue volumes above and below 300 cm^3 (mean volume) to assess the software's performance at greater fibroglandular volumes.

For the subset of women with a family history of breast cancer, breast density was assessed by two observers at each center. In these cases, agreement between the observers' scores was analyzed by examination of the absolute difference between the scores.

Results

A complete set of standard mammographic images (right MLO, left MLO, right CC, left CC) from 8867 individual women were available for analysis. However, the software was unable to produce scores for every image analyzed. Reasons for this inability were not available for the Quantra algorithm; Volpara supplied error messages for each image without a score. Reasons for algorithm failures were varied and are summarized in Table 1. Most of the unexpected errors appeared to be related to the incorrect use of certain exposure settings on the acquisition workstation, leading to inconsistent values being used in the

DICOM image header. Also, as the data set was very large, consisting of 36281 images, we were unable to manually exclude cases with additional views (eg, magnification or mosaic views).

The logical combination of maximal tissue volumes from each view resulted in scores being used from the MLO view in most of the cases. From the Quantra results, where there were 16957 cases of paired CC and MLO scores, the MLO view gave the largest total breast volume in 13953 cases and the largest fibroglandular tissue volume in 12347 cases. From the Volpara results where there were 16991 cases of paired CC and MLO scores, the MLO view gave the largest total breast volume in 14469 cases and the largest fibroglandular tissue volume in 9200 cases. When we combined these volumes from both breasts for each woman, there were therefore cases where the CC view from one breast was averaged with the MLO view from the other. For the Quantra results, this averaging occurred in 1471 cases for total breast volume and in 2495 cases for fibroglandular tissue volume. For the Volpara results, this averaging occurred in 1337 cases for total breast volume and in 2544 cases for fibroglandular tissue volume.

The summary of results for the study cohort is shown in Table 2. An overall density score was assigned by observers in 8391 cases. Quantra scores were available for 8512 women, and 8532 women had scores calculated from Volpara software. Observer scores were therefore unavailable for 5.4% (476 of 8391) of the cohort, and Quantra and Volpara scores were not generated for 4.0% (355 of 8867) and 3.8% (335 of 8867) of the cohort, respectively.

Figure 1 illustrates the comparison of total breast volume, as measured with the Quantra and Volpara software, throughout the study population. The Pearson coefficient for correlation between Quantra breast volume and Volpara breast volume was $r = 0.97$ ($P < .001$). The mean difference between the values calculated for each case was 43.5 cm^3, and the mean percentage difference between the values was 5.0%, suggesting good agreement between the two systems.

Table 1

Summary of Error Messages from Volpara Given as a Reason for Not Obtaining a Density Score in the Analysis of 34755 Images

Error*	No. of Images	Comment
Breast implant present: yes	32	On review, not all of these images had implants present; however, all were stated as such in the DICOM image header, suggesting occasional errors in the use of the 'implant present' mode at the acquisition workstation
Invalid quantitative values—not a standard mammogram?	9	Images that were in this category were ones with implants present but not indicated in the DICOM header and spot compression images; a mammogram where dense tissue was found largely just behind the nipple also appeared in this category
Magnification factor too large (> 1.1)	770	These were all magnification images that were not removed from the image data set
Filter material invalid/unknown	5	These were specimen images identified as such at acquisition; this leads to there being no filter specified in the DICOM header
Possible mosaic	292	Many of these were mosaic view where there was only partial coverage of the breast; some of the images in this category were spot compression views
Possible mag view	3	All of these images turned out to be specimens where the specimen view was not selected at acquisition
No background—not a standard mammogram?	379	On review, these images were collimated so that the breast edge was missing
No energy—all fat found	1	This was an implant image not indicated at acquisition
Recorded breast thickness invalid	3	Two of these were specimen images where a breast thickness of 0 mm is used in the DICOM header; the other image appeared to be a magnification view where the x-ray system has not recognized it as such (ie, broad focus and grid are indicated with a thickness of 317 mm)

Note.—DICOM = Digital Imaging and Communications in Medicine.

* The error messages are as transcribed from the software.

Table 2

Summary of Results for the Total Cohort of Cases

Measurement	Observers	Quantra	Volpara
No. of cases*	8391	8512	8532
Breast volume (cm^3)	. . .	953.5 (73.0–4986.5)	921.4 (33.4–5009.3)
Fibroglandular tissue volume (cm^3)	. . .	93.0 (4.0–1024.0)	71.6 (6.8–628.5)
Area breast density (%)	37.0 (0, 100)	14.8 (0–76.5)	. . .
Volumetric breast density (%)	. . .	9.5 (1.4–56.2)	7.7 (2.5–54.2)

Note.—Unless otherwise indicated, numbers are median values, and numbers in parentheses are the minimum and maximum values, respectively, in the ranges.

* Cases = women.

Figures 2 and 3 show the distribution and comparison of the fibroglandular tissue volumes within the breast reported by the two algorithms. The Pearson coefficient for correlation between the natural logarithm

Figure 1

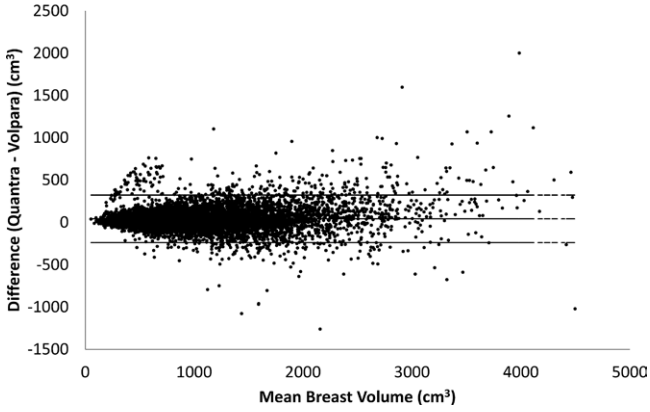

Figure 1: Bland-Altman plot of total breast volume reported by Volpara subtracted from that reported by Quantra compared with the mean of the two results. Middle dashed line = mean difference and top and bottom dashed lines = 95% limits of agreement (± 2 standard deviations).

Figure 2

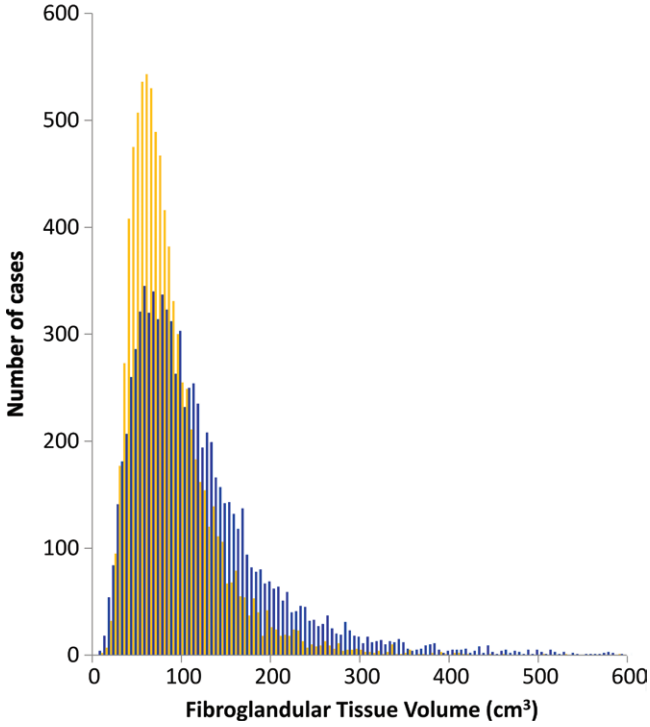

Figure 2: Histogram shows distribution of fibroglandular tissue volume across the study population, as measured by Quantra (blue) and Volpara (yellow). There were 30 cases in the Quantra data set and one case in the Volpara data set that had volumes greater than 600 cm^3, and these data are not shown on the histogram.

of Quantra fibroglandular volume and the natural logarithm of Volpara fibroglandular volume was $r = 0.86$ ($P < .001$). The mean difference between the values calculated for each case was 30.3 cm^3, and the mean percentage difference between the values was 21.2%, with Quantra giving the larger of the two values. The differences between the two measurements increase as the amount of fibroglandular tissue increases. This is particularly noticeable for data above 300 cm^3, where the mean difference between cases is 226.4 cm^3 and the correlation coefficient reduces to 0.15 ($P = .06$) if only those data are considered. The remaining data below 300 cm^3 have a mean difference between cases of 26.4 cm^3 and a correlation coefficient of 0.85 ($P < .001$). In addition, the range of results for fibroglandular volume is greater for the Quantra system, with most values lying between 0 and 400 cm^3, as opposed to results with Volpara, with values lying between 0 and 300 cm^3.

Figures 4 and 5 show the results for volumetric breast density. The Pearson coefficient for correlation between the natural logarithm of Quantra volumetric density and the natural logarithm of Volpara volumetric density was $r = 0.78$ ($P < .001$). The mean difference between the values calculated for each case was 1.61 percentage points, and the mean percentage difference between the values was 16.3%.

Figures 6 and 7 show the results for breast density estimated from the projected mammogram, the area-based breast density score assigned by the observers, and the Quantra software. The Volpara software did not give an area-based density result. The Pearson coefficient for correlation between the natural logarithm of the observer density scores and the natural logarithm of the Quantra area-based density was $r = 0.55$ ($P < .001$).

Figure 8 shows a comparison of the area-based breast density assigned a score by the observers and the volumetric density measurement from each program. The Pearson coefficient for correlation between the observer density scores and the natural logarithms of the volumetric densities was $r = 0.60$ ($P < .001$) for Quantra and $r = 0.63$ ($P < .001$) for Volpara, although there does appear to be a better correlation in breasts with lower density. The

Figure 3

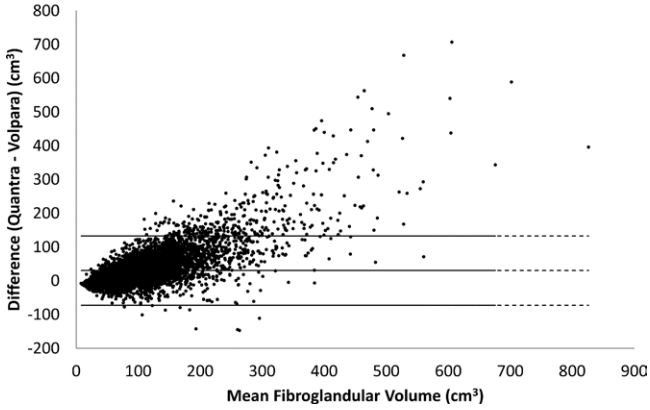

Figure 3: Bland-Altman plot of fibroglandular tissue volume reported by Volpara subtracted from that reported by Quantra compared with the mean of the two results. Middle dashed line = mean difference and top and bottom dashed lines = 95% limits of agreement (± 2 standard deviations).

Figure 4

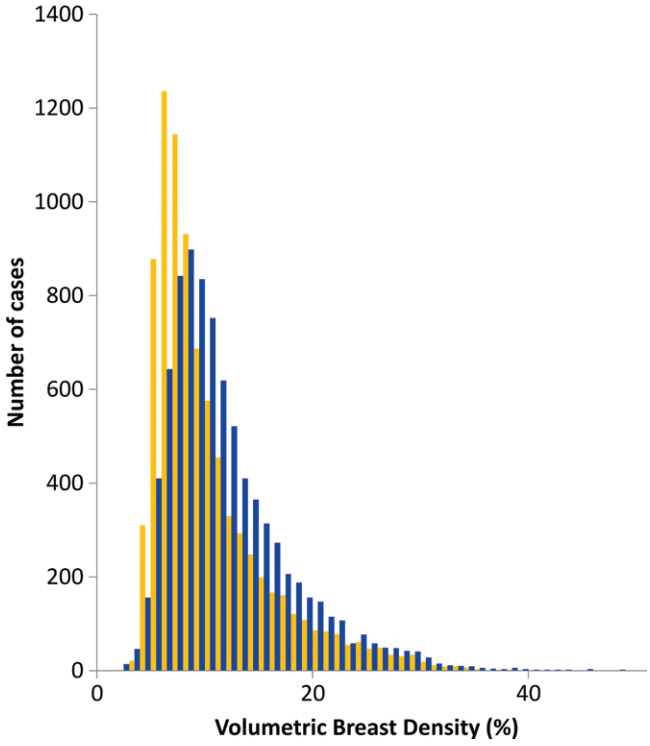

Figure 4: Histogram shows distribution of volumetric breast density across the study population, as measured by Quantra (blue) and Volpara (yellow). There was one case from each data set that had a volumetric density greater than 50%.

large numbers of values given at each 5% mark on the observer histogram in Figures 7 and 8 are a consequence of the way that the visual analog scale was processed in two centers, with one rounding results to within the nearest 5% and the other rounding results to the nearest 10%. Figure 9 shows how the Quantra outputs of area-based and volumetric density compare.

For the cases in women with a family history of breast cancer, density was assessed by two observers at each of the centers, giving us 638 cases with two scores for comparison. In 70% (449 of 638) of these cases, the score agreed to within 10%; however, in 8% of cases, the score disagreed by more than 20%. The 54 cases with more than 20% disagreement were found across the whole range of densities, with 33 of them in the density range of 50%–75%. In 13% (85 of 638) of cases, one reader assigned a score to the density above 50%, with the other reader assigning a score to the density below 50%.

Discussion

The two automated density assessment techniques have relatively good agreement in the evaluation of the overall breast volume; there was less agreement in the assessment of the fibroglandular volumes. We presume from the agreement in total volume that there are similarities in the way each algorithm identifies the breast against the background, while each applies their own corrections to account for compression paddle height and tilt and to estimate the volume at the edge of the breast where the paddle is not in contact with the breast (17,19,23).

Although there was relatively good agreement at lower fibroglandular volumes, it became poorer as those volumes increased, particularly above 300 cm^3. The reasons for this discrepancy are most likely due to the differing ratios of fibroglandular, adipose, and skin tissues allocated to each pixel, on the basis of their relative x-ray attenuation with reference to pixels in the image defined as pure adipose or fatty tissue (19,22). As density increases, it becomes harder to identify these reference areas, and each manufacturer's solution to this problem is likely to result in differences in the final

volumetric density results (17). Further differences might be the way that skin and fibroglandular tissue are differentiated in the assessment of dense tissue (18,24,25).

While the specific corrections that each algorithm makes to account for paddle tilt are unknown because of their proprietary nature, Kallenberg et al (26) have shown that such corrections significantly influence fibroglandular tissue volumes and volumetric densities. In larger fatty breasts, the paddle tilt generates an inhomogeneity in the mammogram, which can cause dense tissue volumes to be overestimated, especially when intentionally flexible paddles are used (23). Conversely, in dense breasts, the increased height at the chest wall edge of the mammogram can lead to an underestimation of density (26).

There is low correlation between observer scores of breast density and automated analysis scores. Observer measurement of breast density has been shown in other studies to be affected by interobserver variability (12,25,27,28). When the histograms of the area-based measurements from both human observers and software analysis are examined, there is clearly a difference in the distribution of scores. This may in part be caused by the observers' application of a semivolumetric approach to the assessment rather than a purely area-based

Figure 5

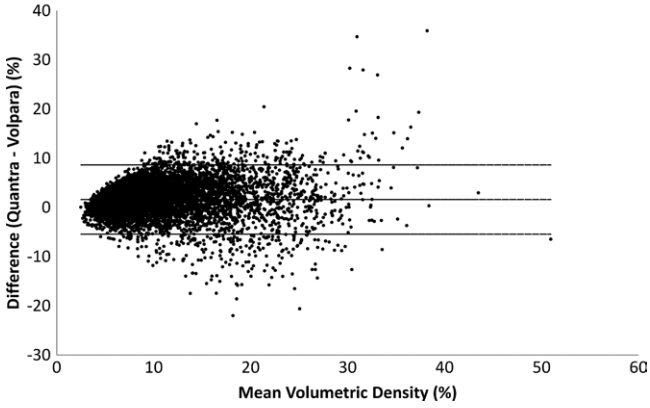

Figure 5: Bland-Altman plot of volumetric breast density reported by Volpara subtracted from that reported by Quantra compared with the mean of the two results. Middle dashed line = mean difference and top and bottom dashed lines = 95% limits of agreement (± 2 standard deviations).

Figure 6

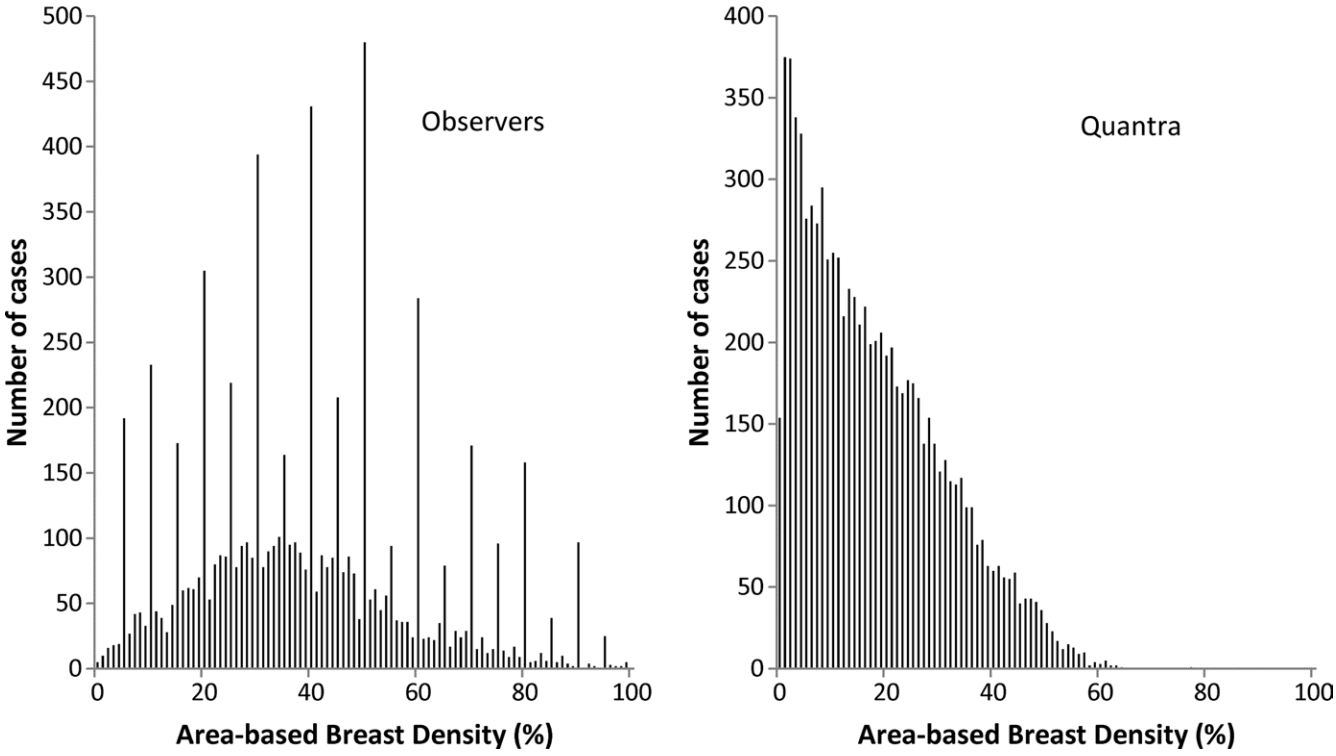

Figure 6: Histograms show distribution of area-based breast density across the study population from the observers (left) and the Quantra software (right).

one. However, when we compared the distribution with those found in the literature for visual and threshold methods (25,29), it is the software that produces the more comparable results.

Figure 7

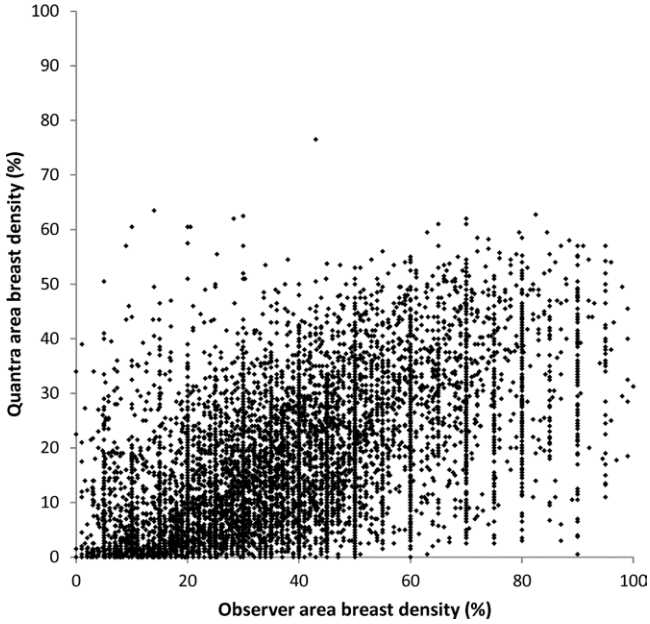

Figure 7: Scatter plot compares the observer and Quantra scores for the area-based breast density.

One possible technical reason for the difference between observer-based density scores and automated scores could be the processing of the displayed image. The software analyzes the raw digital data, whereas the observers make their estimation of density on processed images optimized for display at the workstation and have the ability to alter the window width and level of the gray scale applied to the image's pixels. This adjustment can substantially alter the image presentation. In this study, readers were advised not to alter the window levels.

In the current study, it is not possible to know what the ground truth is with regard to breast density measurement. Both of the automated density measurement tools used in this study have been shown to correlate with measurements made with MR imaging (17,19,24), with methods using a calibration phantom (21) and with visual assessment using Breast Imaging Reporting and Data System categories (17,30). Correlation, however, does not necessarily mean agreement between methods.

In recent work, Wang et al (19) retrospectively analyzed the breast density

Figure 8

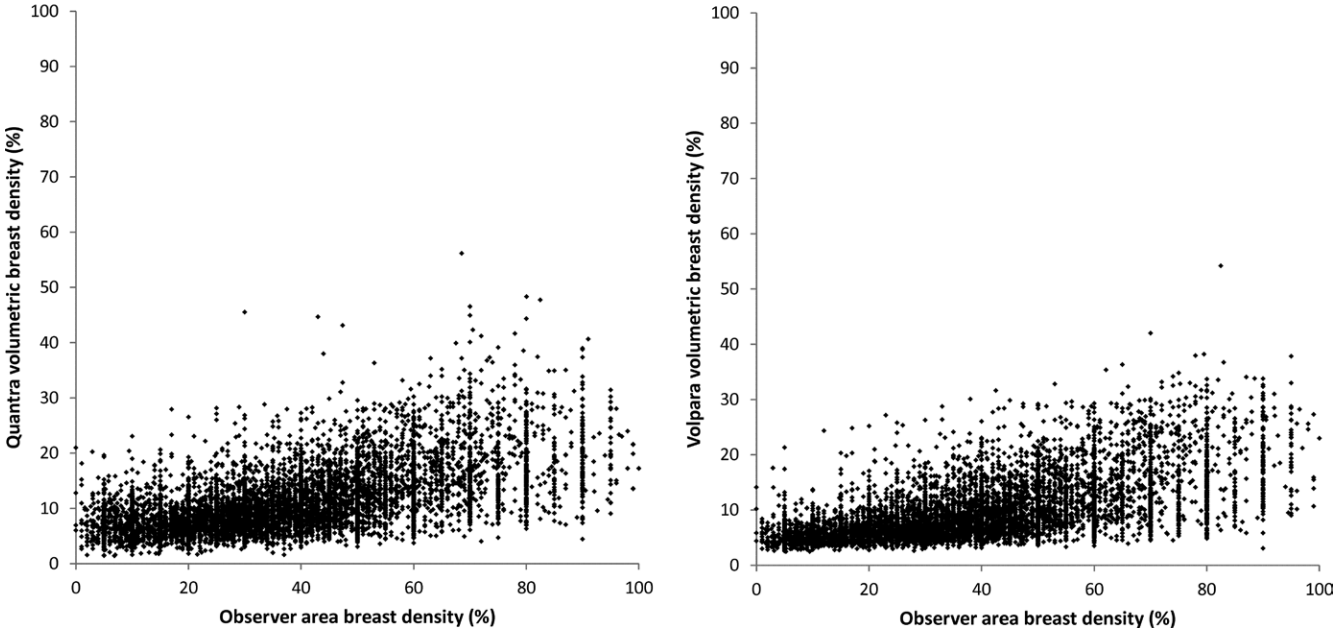

Figure 8: Scatter plots compare the observer area-based breast density scores with the volumetric measurements from Quantra (left) and Volpara (right).

Figure 9

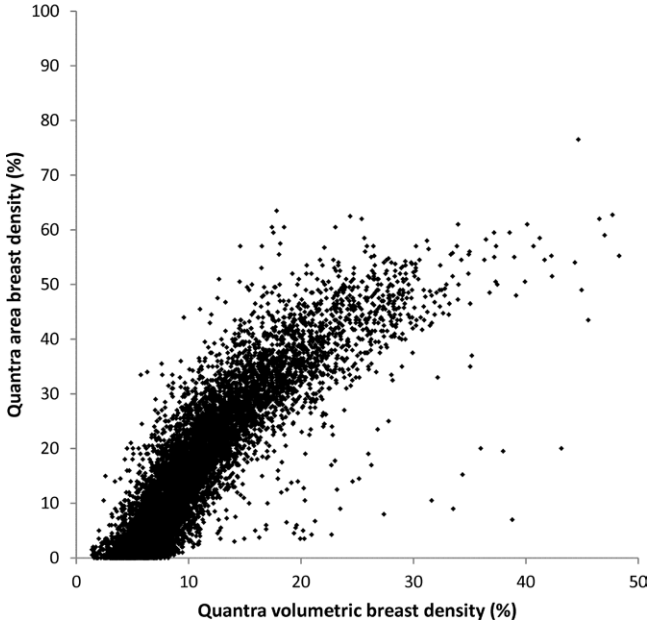

Figure 9: Scatter plot compares the area-based density score and the volumetric density given by Quantra.

of 99 women and compared the performance of Quantra and Volpara with that of a fuzzy-clustering segmentation method on MR images. Total breast volume with both mammographic measurement tools was highly correlated with MR imaging results ($r = 0.95$), although these researchers reported a greater difference between Quantra and Volpara than we have found. One possible explanation may be their use of the CC image and our use of the higher score from either the MLO or the CC view. This is also likely to have an effect on the volumetric density measurements that were reported to be a median of 22.0% and 13.3% for Quantra and Volpara, respectively, which are larger than those found in our work. For fibroglandular volume, the difference between the reported median values was 36.2 cm^3, with Quantra giving the larger score, which is similar to that found in our study.

One of the limitations of this study was that, despite the same cases being submitted to each of the three density measurement methods, the failure in some cases to provide a score means that the population used for comparing each of the three methods is different. Density is not routinely assessed in the National Health Service Breast Screening Programme, and so, it was sometimes overlooked during image reading. However, because each comparison of any two methods uses the results where the populations coincide, and are therefore identical, the effect of this difference is likely to be small. Also, the visual analog scale data were rounded to 5% in some cases and 10% in others, which presents another limitation of the study, although, again, the effect of this difference is unlikely to be large. Finally, to obtain percentage density scores and absolute volumes, we combined the density scores from each image given by the two software tools using the same logic for both. In their derivation of a category-based score for the subject, the software tools may combine the results using a different method than that used in the current study. The fact that the women selected for this study are from two different groups, those with abnormal mammographic results and those with a family history of breast cancer, is not thought to be a confounding factor.

Breast density has been correlated with risk of developing breast cancer (3–5,14,31). Because cancers originate in the fibroglandular tissue, it is proposed that measurements that describe its volume would be a better predictor of risk than those based on projected area (1,6,31). If volumetric density is to be used to estimate breast cancer risk, it is important that the measurements are reliable. Technical differences in the way in which each software package determines the fibroglandular tissue volume, and therefore the density, produce different values. This factor needs to be considered when density is used to inform patient personalized imaging.

Acknowledgments: The authors acknowledge all who contributed to this study: Aberdeen, Scotland: Karen A. Duncan, Herman Klaasen, Tanja Gagliardi, Jeanette Davidson, George Cameron, and Maureen G. C. Gillan. Cambridge, England: Sridevi Nagarajan and Emily Dixon. Glasgow, Scotland: Arachna Seth, Janet Litherland, Mabel Morrow, Linda McClure, and Ann Mumby. Guildford, England: Julie Cooke, Caroline Kissin, Caroline Taylor, Katherine Stoner, Phillipa Skippage, Victoria Cooke, and E. H. L. Mungutroy. King's College Hospital, London, England: Michael J. Michell, Juliet Morel, Jane Goligher, Rumana Rahim, Susan Dyson, and Asif Iqbal. Manchester, England: Yit Yoong Lim, Mary Wilson, Emma Hurley, Ursula Beetles, Sara Bundred, Susan M. Astley, Jin Zhou, Catriona Tate, and Elaine Harkness. St Bartholomew's Hospital, London, England: Hema Purushothaman, Tamara Suaris, and Francis McInally. National Coordinating Centre for Physics of Mammography: Kenneth C. Young and Celia Strudley.

Disclosures of Conflicts of Interest: O.W.E.M. Financial activities related to the present article: institution received a grant from National Institute for Health Research's Health Technology Assessment Programme; institution received payment for provision of Quantra software by Hologic and for provision of Volpara software by Mātakina Technology. Financial activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. **L.T.** disclosed no relevant relationships. **R.B.** Financial activities related to the present article: institution received a consulting fee or honorarium for data processing and computer services from University of Cambridge, Department of Radiology. Financial activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. **P.W.** disclosed no relevant relationships. **S.W.D.** disclosed no relevant relationships. **F.J.G.** Financial activities related to the present article: institution received a grant from National Institute for Health Research's Health Technology Assessment Programme; institution received payment for loaning of Quantra software to enable density measurements of the

mammograms by Hologic and for density measurements of mammograms using Volpara software by Mātakina Technology. Financial activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships.

References

1. Assi V, Warwick J, Cuzick J, Duffy SW. Clinical and epidemiological issues in mammographic density. Nat Rev Clin Oncol 2012; 9(1):33–40.
2. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med 2003; 138(3):168–175.
3. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006;15(6):1159–1169.
4. Boyd NF, Byng JW, Jong RA, et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. J Natl Cancer Inst 1995;87(9):670–675.
5. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007; 356(3):227–236.
6. Boyd N, Martin L, Gunasekara A, et al. Mammographic density and breast cancer risk: evaluation of a novel method of measuring breast tissue volumes. Cancer Epidemiol Biomarkers Prev 2009;18(6):1754–1762.
7. Cuzick J, Warwick J, Pinney E, et al. Tamoxifen-induced reduction in mammographic density and breast cancer risk reduction: a nested case-control study. J Natl Cancer Inst 2011; 103(9):744–752.
8. Boyd NF, Martin LJ, Yaffe MJ, Minkin S. Mammographic density and breast cancer risk: current understanding and future prospects. Breast Cancer Res 2011;13(6):article no. 223. http://breast-cancer-research.com/content/13/6/223. Accessed December 12, 2014.
9. Vachon CM, van Gils CH, Sellers TA, et al. Mammographic density, breast cancer risk and risk prediction. Breast Cancer Res 2007;9(6):article no. 217. http://breast-cancer-research.com/content/9/6/217. Accessed December 12, 2014.
10. Johns PC, Yaffe MJ. X-ray characterisation of normal and neoplastic breast tissues. Phys Med Biol 1987;32(6):675–695.
11. Sukha A, Berks M, Morris J, et al. Visual assessment of density in digital mammograms. In: Martí J, Oliver A, Freixenet J, Martí R, eds. Digital mammography: 10th International Workshop, IWDM 2010, Girona, Catalonia, Spain, June 16–18, 2010—proceedings. Heidelberg, Germany: Springer, 2010; 414–420.
12. Martin KE, Helvie MA, Zhou C, et al. Mammographic density measured with quantitative computer-aided method: comparison with radiologists' estimates and BI-RADS categories. Radiology 2006;240(3):656–665.
13. Conant EF, Li D, Gavenonis S, et al. A comparative study of the inter-reader variability of breast percent density estimation in digital mammography: potential effect of reader's training and clinical experience. In: Martí J, Oliver A, Freixenet J, Martí R, eds. Digital mammography: 10th International Workshop, IWDM 2010, Girona, Catalonia, Spain, June 16–18, 2010—proceedings. Heidelberg, Germany: Springer, 2010; 114–120.
14. Byng JW, Yaffe MJ, Jong RA, et al. Analysis of mammographic density and breast cancer risk from digitized mammograms. RadioGraphics 1998;18(6):1587–1598.
15. Yaffe MJ. Advanced applications of digital mammography. In: Pisani ED, Yaffe MJ, Kuzmiak CM, eds. Digital mammography. Philadelphia, Pa: Lippincott Williams & Wilkins, 2004; 67–76.
16. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. IEEE Trans Med Imaging 2006;25(3):273–282.
17. Hartman K, Highnam R, Warren R. Jackson V. Volumetric assessment of breast tissue composition from FFDM images. In: Krupinski EA, ed. Digital Mammography: 9th International Workshop, IWDM 2008 Tucson, AZ, USA, July 20–23, 2008—proceedings. Heidelberg, Germany: Springer, 2008; 33–39.
18. Yaffe MJ, Boone JM, Packard N, et al. The myth of the 50-50 breast. Med Phys 2009;36(12):5437–5443.
19. Wang J, Azziz A, Fan B, et al. Agreement of mammographic measures of volumetric breast density to MRI. PLoS ONE 2013;8(12):e81653. doi:10.1371/journal.pone.0081653. Published December 4, 2013. Accessed December 9, 2013.
20. Gilbert FJ, Tucker L, Gillan MCG, et al. TOMMY trial: a comparison of TOMosynthesis with digital MammographY in the UK NHS Breast Screening Programme. In: Health Technology Assessment Programme: NIHR HTA Programme 18 December 2012. Version 3.0, September 18, 2012. Funding Reference Number: 09/22/182. Southampton, England: National Institute for Health Research, 2012.
21. Strudley CJ, Young KC, Oduku JM, Looney P, Barnard A, Gilbert FJ. Development of a quality control protocol for digital breast tomosynthesis systems in the TOMMY trial. In: Maidment ADA, Bakic PR, Gavenonis S, eds. Breast imaging: 11th International Workshop IWDM 2012, Philadelphia, PA, USA, 2012—proceedings. Heidelberg, Germany: Springer, 2012; 330–337.
22. Highnam R, Brady M, Yaffe MJ, Karssemeijer N, Harvery J. Robust breast composition measurement–Volpara. In: Martí J, Oliver A, Freixenet J, Martí R, eds. Digital mammography: 10th International Workshop, IWDM 2010, Girona, Catalonia, Spain, June 16–18, 2010—proceedings. Heidelberg, Germany: Springer, 2010; 342–349.
23. Kallenberg MGJ, Karssemeijer N. Compression paddle tilt correction in full-field digital mammograms. Phys Med Biol 2012; 57(3):703–715.
24. Wang J, Aziz A, Newitt D, et al. Comparison of Hologic's Quantra volumetric assessment to MRI breast density. In: Maidment ADA, Bakic PR, Gavenonis S, eds. Breast imaging: 11th International Workshop IWDM 2012, Philadelphia, PA, USA, 2012—proceedings. Heidelberg, Germany: Springer, 2012; 619–626.
25. Sergeant JC, Warwick J, Evans DG, et al. Volumetric and area-based breast density measurement in the Predicting Risk of Cancer at Screening (PROCAS) study. In: Maidment ADA, Bakic PR, Gavenonis S, eds. Breast imaging: 11th International Workshop IWDM 2012, Philadelphia, PA, USA, 2012—proceedings. Heidelberg, Germany: Springer, 2012; 228–235.
26. Kallenberg MGJ, van Gils CH, Lokate M, den Heeten GJ, Karssemeijer N. Effect of compression paddle tilt correction on volumetric breast density estimation. Phys Med Biol 2012;57(16):5155–5168.
27. Wang XH, Good WF, Chapman BE, et al. Automated assessment of the composition of breast tissue revealed on tissue-thickness-corrected mammography. AJR Am J Roentgenol 2003;180(1):257–262.
28. Makaronidis J, Berks M, Sergeant J. Assessment of breast density: reader performance using synthetic mammographic images. In: Manning DJ, Abbey CK, eds. Proceedings of SPIE: medical imaging 2011—image perception, observer performance and technology assessment. Vol 7966. Bellingham, Wash: SPIE—The International Society for Optical Engineering, 2011;article no. 796603. http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=726666. Accessed December 12, 2014.
29. McCormack VA, Highnam R, Perry N, dos Santos Silva I. Comparison of a new and existing method of mammographic density measurement: intramethod reliability and associations with known risk factors. Cancer Epidemiol Biomarkers Prev 2007;16(6):1148–1154.
30. Seo JM, Ko ES, Han BK, Ko EY, Shin JH, Hahn SY. Automated volumetric breast density estimation: a comparison with visual assessment. Clin Radiol 2013;68(7):690–695.
31. Shepherd JA, Kerlikowske K, Ma L, et al. Volume of mammographic density and risk of breast cancer. Cancer Epidemiol Biomarkers Prev 2011;20(7):1473–1482.