Research Article

Cancer Prevention Research

Quantitative Assessment of Breast Density: Transmission Ultrasound is Comparable to Mammography with Tomosynthesis

Rajni Natesan[1,2], James Wiskin[2], Sanghyeb Lee[2], and Bilal H. Malik[2]

Abstract

Elevated breast density is among the strongest independent predictors of breast cancer. Breast density scores are critical inputs in models used to calculate a patient's lifetime risk of developing breast cancer. Today, the only FDA-cleared technology for assessing breast density uses mammography. An alternative modality for breast density quantification is 3D transmission ultrasound (TU). In this retrospective study, we compared automated breast density calculations derived from TU using quantitative breast density (QBD) and mammography with tomosynthesis using VolparaDensity 3.1 for 225 breasts. Pearson correlation coefficients (r) and intraclass correlation coefficients were compared. Subset analyses of extremely dense breasts, premenopausal, and postmenopausal breasts were also performed. Comparative analysis between radiologist-derived density assessment and objective automated scores was performed. Calculations from TU and mammography with tomosynthesis for breast density, total breast volume (TBV), and fibroglandular volume (FGV) were strongly correlated ($r = 0.91$, 0.92, and 0.67, respectively). We observed moderate absolute agreement for FGV and breast density, and strong absolute agreement for TBV. A subset of 56 extremely dense breasts showed similar trends, however with lower breast density agreement in the subset than in the full study. No significant difference existed in density correlation between premenopausal and postmenopausal breasts across modalities. QBD calculations from TU were strongly correlated with breast density scores from VolparaDensity. TU systematically measured higher FGV and breast density compared with mammography, and the difference increased with breast density.

Impact: TU of the breast can accurately quantify breast density comparable with mammography with tomosynthesis.

Introduction

Breast density is one of the strongest predictors of breast cancer risk. The risk is 4 to 5 times greater in women with dense breast tissue in 75% or more of their breasts than those with little or no dense breast tissue (1–5). Accurate assessment of breast density plays a critical role in clinical decisions. For example, a woman with elevated breast density may have a higher lifetime risk score for breast cancer based on risk models such as Tyrer-Cuzick and may be recommended for supplemental screening. Recognizing the importance of breast density, newly enacted federal legislature will mandate mammography reports to include breast-density information to patients and physicians once the FDA establishes the new breast density reporting language (6). However, mammographic breast density assessment is typically based on subjective visual two-dimensional (2D) measurements that result in substantial inter- and intra-observer variability (7). These inconsistencies have led to the development of fully automated technologies to objectively measure volumetric breast density, which are arguably more reliable and accurate (8–10). For example, the FDA-cleared VolparaDensity Software (Volpara Health Technologies) is now available in the clinic for automatically measuring volumetric breast density from 2D and three-dimensional (3D) mammography. Validation studies have shown that VolparaDensity provides volumetric breast density assessment in full-field digital mammograms that can enhance risk assessment (8, 11, 12). Unfortunately, mammography uses low-dose ionizing radiation, which presents cumulative cancer risk over time; these risks are however minimal in light of the benefit of cancer detection in patients who undergo screening mammography (13, 14).

Another FDA-cleared breast imaging modality is transmission ultrasound (TU), which is radiation-free and generates 3D speed-of-sound maps using both reflection

[1]Department of Diagnostic Radiology, MD Anderson Cancer Center, Houston, Texas. [2]Department of Research and Development, QT Ultrasound LLC, Novato, California.

Corresponding Author: Rajni Natesan, UT-MD Anderson Cancer Center, Houston, TX 77030. Phone: 630-886-4441; E-mail: rnatesan@mdanderson.org

Cancer Prev Res 2019;12:871–6

doi: 10.1158/1940-6207.CAPR-19-0268

and transmission techniques (15, 16). Recent developments in TU instrumentation and algorithms have enabled marked improvements in clinical utility, with high-resolution speed-of-sound maps that can quantitatively identify tissue types and measure volumetric breast density using a continuous scale (15–18). This study evaluates the potential for TU to provide accurate breast density assessment as compared with VolparaDensity scores based on mammography with tomosynthesis.

Materials and Methods

In this retrospective study, we directly compared automated volumetric breast density calculations derived from TU with those derived from digital mammography with tomosynthesis to assess the accuracy of TU volumetric breast density measurements. Digital mammography with tomosynthesis was chosen as the reference because it represents the current standard of care. Clearance for the study was obtained from the Western Institutional Review Board.

We examined retrospective data from women screened at the Marin Breast Health Trial Center between April 2018 and December 2018 for a total of 225 unilateral breast scans from 113 subjects. The distribution of age, race, ethnicity, and breast density are summarized in Table 1. Within a 90-day period, the participants received both a standard screening digital mammogram with tomosynthesis using a GE Senographe Essential/SenoClaire Scanner (GE Healthcare) and a full 3D inverse scattering ultrasound tomogram of the breast using a QT Scanner 2000 (QT Ultrasound). For both imaging modalities, the total breast volume (TBV), fibroglandular volume (FGV), and volumetric breast density were calculated for each breast.

Mammographic automated volumetric breast density calculations were provided by VolparaDensity 3.1, which uses a three-compartment model to segment the breast into skin, adipose tissue, and fibroglandular tissue. The VolparaDensity algorithm has been fully described in published literature (11, 19, 20). The estimated fibroglandular tissue volume is divided by the TBV to calculate the volumetric breast density, also known as the percent fibroglandular volume (%FGV). VolparaDensity measures volumetric breast densities ranging from approximately 0% to 35%, which are then binned into four Volpara Density grades VDG1, VDG2, VDG3, and VDG4 for breast density ranges of 0%–3.4%, 3.5%–7.4%, 7.5%–15.4 %, and 15.5 % and above, respectively (9, 12, 21). The Volpara Density grades can then be mapped to the four breast density categories of the American College of Radiology's Breast Imaging Reporting and Data System (BI-RADS), which are used for classification in clinical practice (22, 23). The BI-RADS ed.5 density category for all 225 cases was also determined by three board-certified and MQSA-certified breast radiologists and each case was assigned the median score across the three readers' assessments (22).

Table 1. Summary of age, race, ethnic, and breast density distributions

Age	
Median	54 years
Min	35 years
Max	82 years
Race distribution	
White	99
Asian	10
White/Pacific Islander	1
White/American Indian	1
Black	2
Ethnic distribution	
Not Hispanic/Latino	110
Hispanic Latino	3
Breast density distribution	
Category A	25
Category B	66
Category C	115
Category D	19

NOTE: The density distributions were as determined by radiologists using BI-RADS 5th edition atlas.

Ultrasound tomographic volumetric breast density calculations were performed using QBD (QT Ultrasound), a quantitative breast density three-compartment segmentation algorithm that isolates fibroglandular breast tissue volume from whole-breast volume using the speed of sound (18, 24). The TU scan is performed with the breast immersed in a water bath. Because skin and fibroglandular tissue both have a relatively high speed of sound compared with adipose tissue, skin must be segmented out and is differentiated from the fibroglandular tissue using its proximity to the water's border. The automated TU software then measures the volumetric breast density using a continuous scale, which has been shown to be more accurate than categorical density scales (25).

Next, we compared the volumetric breast measurements (TBV and FGV) and breast density calculations (%FGV) derived from both imaging modalities for all 225 breasts—using the Pearson correlation coefficient (r) to quantify the strength of association and the intraclass correlation coefficient (ICC) based on a two-way mixed-effects, absolute agreement, individual ratings model to quantify the level of agreement (26). All statistical analyses were performed using the R statistical package version 3.5.2 (R Core Team, 2018). Comparisons were visualized with scatter plots. Natural log transformations were performed on the data to normalize the distributions prior to analyses for correlation and agreement. These analyses were repeated for a subset of 56 extremely dense breasts, defined as having a VDG of 4 (BI-RADS density category D).

The QBD continuous measurements for this study were binned into four breast density ranges, corresponding to the four BI-RADS breast density categories. We used Deming linear regression to quantify the relationship between VolparaDensity and QBD scores. The resultant linear equation was used to convert the four VDG bins to respective

Table 2. Descriptive summaries of breast density, TBV, and FGV estimates by QBD and Volpara for the full set of breast cases

Measure	QBD	Volpara	Ratio
Breast density (%)			
Mean	22.7	11.2	2.3
SD	12.5	6.8	0.8
Median	18.9	10.5	2.0
IQR	15.3	10.6	0.9
Min	6.4	2.5	1.3
Max	66.2	30.0	5.1
TBV (cm^3)			
Mean	581.4	777.8	0.8
SD	345.9	467.5	0.3
Median	458.4	563.5	0.7
IQR	559.8	677.1	0.1
Min	92.0	103.0	0.4
Max	1,608.9	2,068.8	3.4
FGV (cm^3)			
Mean	105.8	66.9	1.8
SD	52.1	38.5	1.2
Median	96.6	57.8	1.5
IQR	52.8	39.0	0.7
Min	36.9	15.4	0.8
Max	383.0	288.1	11.6

NOTE: The ratio of QBD to Volpara is also summarized.

QBD bins. Both VDG and QBD density categories were compared with the respective BI-RADS density category for all the cases using Cohen kappa analysis which resulted in both an overall agreement score and kappa coefficient.

Results

For the full set of 225 breasts, QBD demonstrated a strong linear association with VolparaDensity for breast density with a Pearson correlation coefficient of 0.91. Tables 2 and 3 summarize our measurements.-VolparaDensity derived from digital mammography with tomosynthesis measured an average TBV of 777.8 ± 467.5 cm^3 (range: 103.0–2068.8 cm^3), average FGV of 66.9 ± 38.5 cm^3 (range: 15.4–288.1 cm^3), and average breast density of $11.2\% \pm 6.8\%$ (range: 2.5%–30.0%). The automated QBD calculations derived from TU measured the same 225 breasts to have an average TBV of 581.4 ± 345.9 cm^3 (range: 92.0–1608.9 cm^3), average FGV of 105.8 ± 52.1 cm^3 (range: 36.9–383.0 cm^3), and average breast density of $22.7 \pm 12.5\%$ (range: 6.4%–66.2%).

Table 3. Summary of association and agreement between QBD and Volpara for the full set, extremely dense breast subset, premenopausal subset, and postmenopausal subset of cases

Measure	Full set ($n = 225$)	Dense set ($n = 56$)	Premenopausal set ($n = 84$)	Postmenopausal set ($n = 136$)
Breast density (%)				
Pearson	0.91	0.69	0.92	0.91
ICC	0.49	0.12	0.48	0.48
TBV (cm^3)				
Pearson	0.92	0.80	0.97	0.89
ICC	0.84	0.74	0.87	0.82
FGV (cm^3)				
Pearson	0.67	0.65	0.74	0.75
ICC	0.44	0.51	0.46	0.40

The scatter plot in Fig. 1A demonstrates that QBD typically yielded (in 96% of cases) smaller TBV measurements than VolparaDensity. On average, the QBD TBV was 196.4 cm^3 less than VolparaDensity. Similarly, a Bland–Altman analysis indicates that the QBD TBV will be between 0.45 and 1.26 times the VolparaDensity TBV about 95% of the time, as shown in Fig. 1D. However, the TBVs derived from mammography with tomosynthesis and TU had a strong positive linear relationship with a Pearson correlation coefficient of 0.92 and a moderately high level of absolute agreement with an ICC of 0.84, as shown in Table 3.

The scatter plot in Fig. 1B shows that QBD routinely measured (in 96% of cases) a higher FGV than VolparaDensity with an average difference of 38.9 cm^3. Similarly, a Bland–Altman analysis indicates that the QBD FGV will be between 0.75 and 3.62 times the VolparaDensity FGV about 95% of the time, as shown in Fig. 1E. QBD demonstrated a moderately strong linear association with VolparaDensity for the FGVs with a Pearson correlation coefficient of 0.67 and a moderate absolute level of agreement with an ICC of 0.44. As a result, QBD consistently measured a higher volumetric breast density than VolparaDensity with an average %FGV difference of 11.5%, as shown in Fig. 1C. Similarly, a Bland–Altman analysis indicates that the QBD breast density will be between 1.21 and 3.93 times the VolparaDensity breast density about 95% of the time, as shown in Fig. 1F. Although QBD demonstrated a strong linear association with VolparaDensity for breast density, a moderate level of absolute agreement was observed with an ICC of 0.49.

The subset of 56 extremely dense breasts showed similar trends as the full set. However, the association and agreement between breast densities of the two modalities were notably less in the subset than in the full study with a Pearson correlation coefficient of 0.69 and an ICC of 0.12, as shown in Table 3.

A comparison of the QBD–Volpara correlation for the 84 premenopausal breasts ($r = 0.92$; $P < 0.0001$) and 136 postmenopausal (natural or surgical) breasts ($r = 0.91$; $P < 0.0001$) demonstrated no significant difference, indicating that correlation strength is likely not impacted by menopausal status in this study.

The data points corresponding to correlation of VolparaDensity and QBD values were used in a Deming linear regression analysis to generate a linear fit. The resulting fit equation is:

$$\text{QBD} = (1.95 \times \text{VolparaDensity}) + 0.842$$

This fit equation was used to generate QBD range cutoffs as defined by VDG categories and were calculated as below 7.7%, 7.7%–15.4%, 15.5%–31.0%, and 31.0% and above, corresponding to VolparaDensity ranges of below 3.5%, 3.5%–7.4 %, 7.5%–15.4%, and 15.5% and above, respectively. Both VDG and QBD density categories were

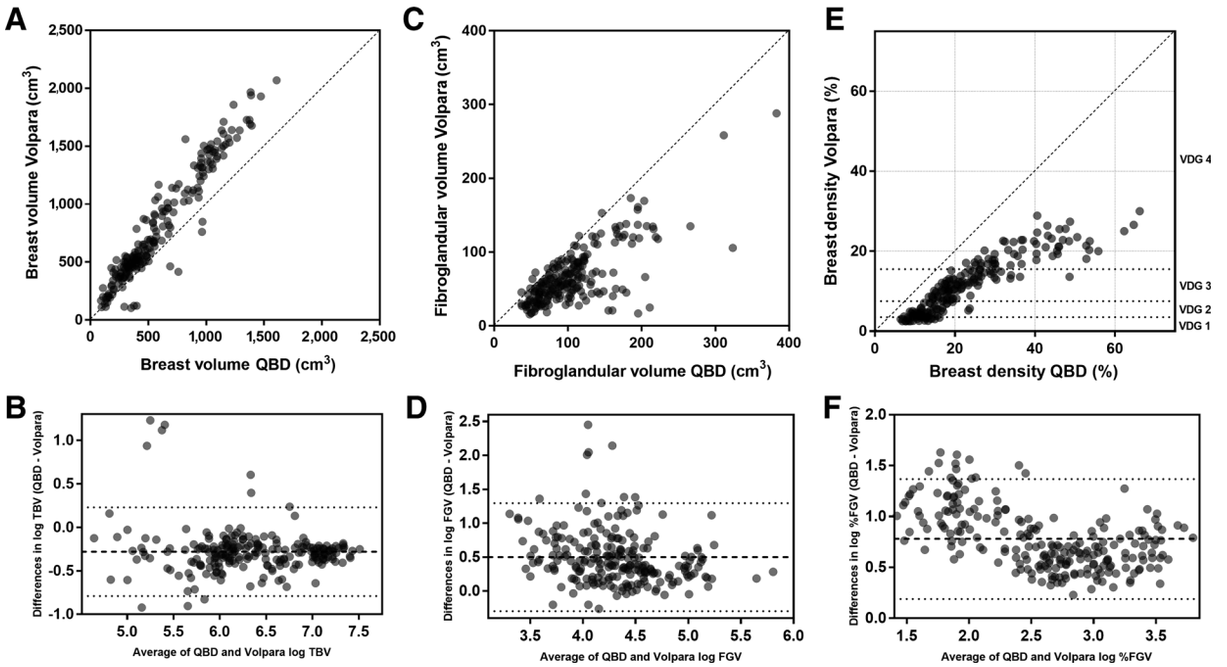

Figure 1.

Scatter plot of values calculated by VolparaDensity and QBD algorithms for TBV (**A**), FGV (**C**), and breast density (**E**). Bland–Altman plot of agreement between VolparaDensity and QBD after log transformation for TBV (**B**), FGV (**D**), and breast density (**F**).

compared with BI-RADS density categories. The comparison between VDG and BI-RADS resulted in an overall agreement of 58.7% and weighted kappa value of 0.56. A similar comparison between QBD density categories and BI-RADS resulted in an overall agreement of 62.7% and weighted kappa value of 0.53.

Pearson correlation coefficient and intraperson correlation between left and right breasts were also calculated. The QBD comparison between respective right and left breasts resulted in Pearson correlation coefficient of 0.97 [95% confidence interval (CI), 0.95–0.98] and ICC of 0.96 (95% CI, 0.94–0.98), indicative of strong relative and absolute agreements. In comparison, VolparaDensity showed similar results with Pearson correlation coefficient of 0.98 and (95% CI, 0.96–0.98) and ICC of 0.97 (95% CI, 0.96–0.98).

Discussion

Our study observed strong positive correlations between breast density calculations derived from TU and digital mammography with tomosynthesis, indicating that the two imaging modalities provide comparable breast density assessments. QBD routinely measured smaller TBVs than VolparaDensity. This is likely due to differences in the field of view for the two imaging modalities. Standard mammographic mediolateral oblique views capture additional anatomy beyond the fibroglandular tissue, which may be included in the VolparaDensity calculations. Despite these differences, we observed a strong positive linear relationship and a moderately high level of absolute agreement between the two modalities for TBV.

QBD routinely measured larger FGVs than VolparaDensity. This discrepancy may be due to the challenges of estimating volumetric breast density using synthesized tomosynthesis because the 2D images do not fully capture the 3D heterogeneity of fibroglandular breast tissue, particularly for highly dense breasts. In intrinsic 3D imaging modalities like TU, image reconstruction allows improved differentiation between adipose and fibroglandular tissues. Despite this discrepancy, QBD still demonstrated a moderately strong correlation with VolparaDensity for FGV.

Given these trends in TBV and FGV, QBD consistently calculated larger volumetric breast densities than VolparaDensity and this trend was observed more strongly in the subset of extremely dense breasts. The range of QBD values is approximately twice as large as the range of VolparaDensity values. We also note that there is a similar spread of values when MRI is used to estimate breast density compared with digital mammography–based methods (21, 27). This indicates that QBD may be more sensitive than VolparaDensity for detecting variations in breast density within the dense breast cohort.

Our comparison of QBD and BI-RADS density grades resulted in an agreement similar to that between VDG and BI-RADS density grades. The agreement between VDG and BI-RADS density categories has been reported before and shows a wide span of results with kappa values ranging

from 0.45 to 0.8 (28–30). The comparison reported in this study resulted in a kappa value of 0.56 which is within the published range.

A limitation of our study design was the 90-day time window allowed between the mammogram with tomosynthesis and ultrasound tomography, as we did not account for any physiologic changes that may have occurred between the breast scans. A second limitation of this study is the absence of MRI-based density measurements, as previously published studies used MRI as a ground truth modality to estimate the accuracy of breast density for measurements using other modalities (12, 21).

In summary, we observed strong positive correlations between automated breast density calculations derived from TU and digital mammography with tomosynthesis. Our validation study demonstrates that TU can quantitatively measure volumetric breast density using speed of sound and provide comparable density assessments to mammography with tomosynthesis. Further research is needed to understand how QBD may contribute to the output of risk models such as the Tyrer-Cuzick and Breast Cancer Surveillance Consortium (BCSC) models.

Disclosure of Potential Conflicts of Interest

R. Natesan is co-president and chief medical officer at QT Ultrasound, reports receiving a commercial research grant from Carestream-RSNA, and has ownership interest (including patents) at QT Ultrasound. J. Wiskin is a principal scientist at QT Ultrasound, LLC. B.H. Malik is a principal scientist at QT Ultrasound LLC. No potential conflicts of interest were disclosed by the other author.

Authors' Contributions

Conception and design: R. Natesan, B.H. Malik
Development of methodology: R. Natesan, J. Wiskin
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): R. Natesan
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): R. Natesan, B.H. Malik
Writing, review, and/or revision of the manuscript: R. Natesan, J. Wiskin, B.H. Malik
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): R. Natesan
Study supervision: R. Natesan
Other (development of breast density estimation algorithm): J. Wiskin
Other (design and implementation of breast density calculation software): S. Lee

Acknowledgments

This work was supported in part by the NIH (R01 CA138536).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 24, 2019; revised September 19, 2019; accepted October 16, 2019; published first October 23, 2019.

References

1. Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000;92:1081–7.
2. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer.N Engl J Med 2007;356:227–36.
3. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomark Prev 2006;15:1159–69.
4. Kerlikowske K, Scott CG, Mahmoudzadeh AP, Ma L, Winham S, Jensen MR, et al. Automated and clinical breast imaging reporting and data system density measures predict risk for screen-detected and interval cancers: a case-control study.Ann Intern Med 2018; 168:757–65.
5. Tice JA, Cummings SR, Ziv E, Kerlikowske K. Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population.Breast Cancer Res Treat 2005;94:115–22.
6. Cision.DenseBreast-Info.org hails Federal law that mandates that mammography reports include breast density information. Available from: https://www.prnewswire.com/news-releases/densebreast-infoorg-hails-federal-law-that-mandates-that-mammography-reports-include-breast-density-information-300797661.html.
7. Nicholson BT, LoRusso AP, Smolkin M, Bovbjerg VE, Petroni GR, Harvey JA. Accuracy of assigned BI-RADS breast density category definitions.Acad Radiol 2006;13:1143–9.
8. Alonzo-Proulx O, Mawdsley GE, Patrie JT, Yaffe MJ, Harvey JA. Reliability of automated breast density measurements.Radiology 2015;275:366–76.
9. Ng KH, Lau S. Vision 20/20: mammographic breast density and its clinical applications.Med Phys 2015;42:7059–77.
10. Holland K, van Gils CH, Mann RM, Karssemeijer N. Quantification of masking risk in screening mammography with volumetric breast density maps.Breast Cancer Res Treat 2017; 162:541–8.
11. Highnam R, Brady M Sir, Yaffe MJ, Karssemeijer N, Harvey J. Robust breast composition measurement - VolparaTM. In:Martí J, Oliver A, Freixenet J, Martí R, editors.Digital Mammography. IWDM 2010. Lecture Notes in Computer Science. Berlin,Heidelberg: Springer;2010. p. 342–9.
12. Wang J, Azziz A, Fan B, Malkov S, Klifa C, Newitt D, et al. Agreement of mammographic measures of volumetric breast density to MRI.PLoS One 2013;8:e81653.
13. Preston DL, Mattsson A, Holmberg E, Shore R, Hildreth NG Jr, Boice JD Jr. Radiation effects on breast cancer risk: a pooled analysis of eight cohorts.Radiat Res 2002;158: 220–35.
14. Council NR.Health risks from exposure to low levels of ionizing radiation: BEIR VII phase 2. Washington, DC:The National Academies Press;2006.
15. Malik B, Terry R, Wiskin J, Lenox M. Quantitative transmission ultrasound tomography: imaging and performance characteristics.Med Phys 2018;45:3063–75.

16. Lenox MW, Wiskin J, Lewis MA, Darrouzet S, Borup D, Hsieh S. Imaging performance of quantitative transmission ultrasound. Int J Biomed Imaging 2015;2015:454028.
17. Malik B, Klock J, Wiskin J, Lenox M. Objective breast tissue image classification using quantitative transmission ultrasound tomography.Sci Rep 2016;6:38857.
18. Wiskin JW, Borup DT, Iuanow E, Klock J, Lenox MW. 3-D nonlinear acoustic inverse scattering: algorithm and quantitative results.IEEE Trans Ultrason Ferroelectr Freq Control 2017;64: 1161–74.
19. 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; 25:273–82.
20. Jeffreys M, Harvey J, Highnam R. Comparing a new volumetric breast density method (Volpara™) to cumulus. In:Martí J, Oliver A, Freixenet J, Martí R, editors.Digital Mammography. IWDM 2010. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer;2010. p. 408–13.
21. Gubern-Merida A, Kallenberg M, Platel B, Mann RM, Marti R, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms: a validation study.PLoS One 2014;9: e85952.
22. D'Orsi CJ SE, Mendelson EB, Morris EA.ACR BI-RADS® atlas, breast imaging reporting and data system. Reston, VA:American College of Radiology;2013.
23. Winkler NS, Raza S, Mackesy M, Birdwell RL. Breast density: clinical implications and assessment methods.Radiographics 2015;35:316–24.
24. Wiskin J, Malik B, Natesan R, Lenox M. Quantitative assessment of breast density using transmission ultrasound tomography. Med Phys 2019;46:2610–20.
25. Abdolell M, Tsuruda KM, Brown P, Caines JS, Iles SE. Breast density scales: the metric matters.Br J Radiol 2017;90:20170307.
26. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial.Tutor Quant Methods Psychol 2012;8:23–34.
27. Moon WK, Chang J-F, Lo C-M, Chang JM, Lee SH, Shin SU, et al. Quantitative breast density analysis using tomosynthesis and comparison with MRI and digital mammography.Comput Methods Programs Biomed 2018;154:99–107.
28. Raza S, Mackesy MM, Winkler NS, Hurwitz S, Birdwell RL. Effect of training on qualitative mammographic density assessment. J Am Coll Rad 2016;13:310–5.
29. van der Waal D, den Heeten GJ, Pijnappel RM, Schuur KH, Timmers JMH, Verbeek ALM, et al. Comparing visually assessed BI-RADS breast density and automated volumetric breast density software: a cross-sectional study in a breast cancer screening setting.PLoS One 2015;10:e0136667.
30. McEntee MF, Damases CN.Mammographic density measurement: a comparison of automated volumetric density measurement to BIRADS.SPIE2014. DOI - 10.1117/12.2042966.