Original Investigation

Automated Volumetric Mammographic Breast Density Measurements May Underestimate Percent Breast Density for High-density Breasts

Kareem Rahbar, MD[1], Albert Gubern-Merida, PhD, James T. Patrie, MS, Jennifer A. Harvey, MD

Rationale and Objectives: The purpose of this study was to evaluate discrepancy in breast composition measurements obtained from mammograms using two commercially available software methods for systematic trends in overestimation or underestimation compared to magnetic resonance-derived measurements.

Materials and Methods: An institutional review board-approved, Health Insurance Portability and Accountability Act-compliant retrospective study was performed to calculate percent breast density (PBD) by quantifying fibroglandular volume and total breast volume derived from magnetic resonance imaging (MRI) segmentation and mammograms using two commercially available software programs (Volpara and Quantra). Consecutive screening MRI exams from a 6-month period with negative or benign findings were used. The most recent mammogram within 9 months was used to derive mean density values from "for processing" images at the per breast level. Bland-Altman statistical analyses were performed to determine the mean discrepancy and the limits of agreement.

Results: A total of 110 women with 220 breasts met the study criteria. Overall, PBD was not different between MRI (mean 10%, range 1%–41%) and Volpara (mean 10%, range 3%–29%); a small but significant difference was present in the discrepancy between MRI and Quantra (4.0%, 95% CI: 2.9 to 5.0, $P < 0.001$). Discrepancy was highest at higher breast densities, with Volpara slightly underestimating and Quantra slightly overestimating PBD compared to MRI. The mean discrepancy for both Volpara and Quantra for total breast volume was not significantly different from MRI ($p = 0.89$, 0.35, respectively). Volpara tended to underestimate, whereas Quantra tended to overestimate fibroglandular volume, with the highest discrepancy at higher breast volumes.

Conclusions: Both Volpara and Quantra tend to underestimate PBD, which is most pronounced at higher densities. PBD can be accurately measured using automated volumetric software programs, but values should not be used interchangeably between vendors.

Key Words: Breast density; mammography; breast MRI.

INTRODUCTION

Breast density decreases the sensitivity of mammography (1,2) and is a moderate independent risk factor for breast cancer (3–6). In current practice, evaluation of breast density from mammograms using Breast Imaging Reporting and Data System (BI-RADS) density categories (7,8) is somewhat subjective, with only a moderate inter-reader agreement (9–11). In the current edition of BI-RADS (8), more subjectivity is encouraged regarding upgrading of mammograms with focal areas of density to the heterogeneously dense category, resulting in lower inter-reader agreement (12). As public awareness and research continues on breast density, an accurate automated assessment of percent breast density (PBD) from mammograms is needed.

Mammographic breast density can be quantified using area or volumetric methods. In area-based methods, pixels of the mammogram are segmented into fat or breast tissues in a binary fashion (13). The area-based methods have consistently demonstrated a moderate statistically significant association with breast cancer risk with the women in the highest quartile of the population being about four times more likely to be diagnosed with breast cancer than women in the lowest-density quartile (5). A disadvantage of area-based methods is the lack of accounting for pixel depth, or the whiteness of the pixel.

Acad Radiol 2017; ■:■■–■■

From the Roper Radiologists, P.A., Charleston, South Carolina (K.R.); Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, The Netherlands (A.G.-M.); Department of Biostatistics, University of Virginia Health System, Charlottesville, Virginia (J.T.P.); Department of Radiology and Medical Imaging, University of Virginia Health System, Charlottesville, VA 22908 (J.A.H.). Received May 5, 2017; revised June 20, 2017; accepted June 20, 2017. This study was funded by the Congressionally Directed Medical Research Programs (CDMRP, grant no. BC100474). [1]Current affiliation- work performed while resident at UVa. **Address correspondence to:** J.A.H. e-mail: jharvey@virginia.edu

http://dx.doi.org/10.1016/j.acra.2017.06.002

Volumetric breast density software programs quantify mammographic breast density by evaluating the whiteness (pixel depth) of the mammogram, creating a quantitative density map to estimate the percent volume of breast tissue. These values will inherently be smaller than area methods because a pixel that would be valued as binary positive for fibroglandular tissue may range in value from 1% to 100% to account for the whiteness of the pixel.

Automated volumetric software programs can compute percent volume breast density using “for processing” information from digital mammograms (14,15). Volumetric-based methods are highly reliable (16) and should theoretically improve breast cancer risk prediction over area-based methods. Volumetric breast density is likewise associated with breast cancer risk, but the degree is variable (17,18).

Concern has been raised, however, about the ability of these widely available programs to accurately measure volumetric breast densities from two-dimensional (2D) mammographic images (19,20). Magnetic resonance imaging (MRI) produces signals related to the fat and water composition of the breast, with the water content highly correlated with the fibroglandular tissue volume (21). Several studies have shown that PBD derived from mammograms using automated volumetric software has high correlation with density derived from MRI (22–25). However, these studies have limitations in how correlation was made with the MRI (23), including only one volumetric software algorithm (22), and none evaluated for systematic variations for low- and high-density breasts.

The degree of discrepancy in breast composition measurements obtained from mammograms and MRI is useful to assess for systematic trends in the overestimation or underestimation of breast composition breast measurement methods and may explain in part the variability in the association between volumetric density and breast cancer risk. Adjustments of volumetric mammographic density software could then be made, which may result in improved breast cancer risk prediction.

The purpose of the present study was to evaluate discrepancy in breast composition measurements obtained from mammograms using two commercially available software methods for systematic trends in overestimation or underestimation compared to magnetic resonance (MR)-derived measurements.

MATERIALS AND METHODS

The present study was approved by our institutional review board and was compliant with the Health Insurance Portability and Accountability Act. A waiver of consent was granted. This retrospective analysis compared automated volumetric mammography-derived breast density measurements with density measurements derived from MRI examinations as the reference standard. Age and reported BI-RADS mammographic density (7) were obtained from the mammography reports to assess if the sample reflected a typical screening population.

Consecutive breast MRI examinations performed with the indication of high-risk screening between January 2012 and August 2012 were included if the woman was age 18 and older, asymptomatic, and had a final assessment BI-RADS category of 1 or 2 (negative or benign finding). Patients with a history of breast augmentation or mastectomy were excluded. Patients with a prior personal history of breast cancer treated with lumpectomy were included. Patients were also required to have a bilateral digital mammogram performed at our institution that included craniocaudal (CC) and mediolateral oblique (MLO) views within 9 months of the MRI study.

Mammographic breast density refers to the attenuation of the X-ray beam caused by the fibroglandular breast tissue. Although the volume of fibroglandular tissue is not identical to the breast density as perceived on the mammogram, in the present study, the term fibroglandular volume (FGV) (22,26) is used interchangeably to represent the volume of breast tissue as depicted on either MRI or mammography. Likewise, PBD refers to the percentage of the breast occupied by tissue that attenuates the X-ray beam, either by area (2D) or volume (three-dimensional) of the breast and typically refers to mammography. In the present study, the term PBD refers to the volume of fibroglandular tissue divided by the total breast volume (TBV), whether estimated from mammography or MRI.

MRI and Volumetric Processing

All breast MRI examinations were acquired using a 1.5- or 3.0-T MRI scanner (Avanto, Espree, Skyra; Siemens Medical Solutions, Malvern, PA) and either a 7- or 15-channel bilateral phase array breast coil (Invivo, Gainesville, FL; Siemens Medical Solutions, Malvern, PA). The examinations were performed with the patient lying in a prone position. Slice thickness for all subjects measured 1–2 mm in the axial plane. T1-weighted fat saturated noncontrast images were used to calculate the breast density measurements. The breast MRI examination also included a T2-weighted fat-saturated sequence and sequential T1-weighted fat-saturated sequences obtained after intravenous administration of gadolinium-based contrast materials that were not used in the present study.

PBD was calculated by segmenting the breast MR T1-weighted sequence into fibroglandular tissue and fatty breast tissue using a semiautomated method for each breast independently (27–29). All segmented images were then visually reviewed using ITK-snap (version 3.0.0, University of Pennsylvania, Philadelphia, PA) (30), to ensure that the segmented area included only the breast and that fat and fibroglandular tissue were appropriately segmented. Manual editing by one of the authors (K.R.), who was blinded to the results of the mammography-derived density measurements, was performed through ITK-snap to remove areas of chest wall and axilla that were outside of the breast. MR images of 31 women (28.2%) required manual removal of nonbreast areas. Sequences that were modified were then reprocessed for segmentation. For each breast, numerical values for the FGV