This copy is for personal use only. To order printed copies, contact *reprints@rsna.org*

Beyond BI-RADS Density: A Call for Quantification in the Breast Imaging Clinic[1]

Emily F. Conant, MD
Brian L. Sprague, PhD
Despina Kontos, PhD

The clinical assessment of breast density is becoming increasingly important in determining personalized screening regimens, because increased breast density limits the sensitivity and specificity of mammography and is also an independent risk factor for the development of breast cancer (1–4). Breast density notification legislation, currently passed in more than half of the states in the United States, mandates that women be notified of their breast density and be informed that breast density may limit the mammographic detection of breast cancers. Although the required wording of density notification varies according to state, most statements encourage shared decision making between women and their health care providers regarding the need for possible supplemental screening and breast cancer risk assessment (5,6). National legislation has also been under consideration that would require all mammography facilities to report breast density to physicians and patients and to require the Department of Health and Human Services to expand and intensify research on the cost-effectiveness of supplemental screening and best practices for imaging women with dense breasts (7).

An important concern surrounding breast density notification legislation and its potential effect on screening regimens is the subjective nature of clinical breast density ratings (8). Currently, breast density is typically determined by the interpreting radiologist who visually assesses the amount of radio-opaque or "dense" tissue from the "for presentation" (ie, vendor-processed) clinical mammogram. Beginning with the first version of the American College of Radiology Breast Imaging and Reporting Data System (BI-RADS) guidelines, summary descriptors of the breast composition or density were recommended for inclusion in the mammographic report (9). In the fourth edition of BI-RADS (10), quartile-based categories for breast density were defined on the basis of the perceived percentage of mammographically dense breast tissue in the entire area of the breast: 1, almost entirely fat (< 25% glandular tissue); 2, scattered fibroglandular densities (25%–50% glandular tissue); 3, heterogeneously dense tissue (51%–75% glandular tissue); and 4, extremely dense tissue (> 75% glandular tissue). For the purpose of most breast density state notification laws and according to convention in most research studies, the latter two density categories, heterogeneously dense and extremely dense tissue (breast with > 50% glandular density), are generally combined to represent dense breasts and the first two categories (breasts with less ≤ 50% glandular tissue) are considered to represent "nondense" breasts.

A new fifth edition of BI-RADS (11), published in 2013, recommends that breast imagers assign breast composition descriptors that better convey whether there are dense areas of tissue that could mask or obscure a cancer (12). The new definitions for the density categories exclude the numeric quartiles of percentages of dense area used in the fourth BI-RADS edition, potentially adding additional variability in breast density assessment. With these new definitions, breast density categories may be assigned on the basis of the densest region of breast tissue seen on the mammogram rather than the overall amount of dense tissue distributed throughout the entire breast. The new categories are as follows: a, almost entirely fatty; b, scattered areas of fibroglandular density; c, heterogeneously dense, which may obscure detection of small masses; and d, extremely dense, which lowers the sensitivity of mammography. It is also recommended that descriptive information be included in the mammographic report regarding the distribution of the

[1] From the Department of Radiology, Hospital of the University of Pennsylvania, Perelman School of Medicine, 3400 Spruce St, Philadelphia PA 10104 (E.F.C., D.K.); and Departments of Surgery and Radiology, University of Vermont Cancer Center, Burlington, Vt (B.L.S.). Received March 22, 2017; revision requested April 18; revision received June 3; accepted June 26; final version accepted June 27. **Address correspondence to** E.F.C. (e-mail: Emily.conant@uphs.upenn.edu).

E.F.C. and D.K. supported by the National Institutes of Health (U54CA163313). B.L.S. supported by the National Institutes of Health (U54CA163303r).

© RSNA, 2018

dense tissue such as "the dense tissue is located anteriorly" (11). This change in BI-RADS density definitions emphasizes the potential for the masking of cancer by locally dense areas of tissue, regardless of the size of the dense area. However, the exclusion of an estimate of the overall dense tissue area de-emphasizes the value of mammographic dense area in estimating a woman's risk of developing breast cancer (13). With the new fifth edition classification system, a woman who previously was categorized as having "less dense" breasts ("scattered" or "fatty") on the basis of the fourth edition percentage of dense area definitions who has only a focally dense area of glandular tissue could be classified as having c, or heterogeneously dense breasts, which, in most states, would trigger mandatory density notification and potentially prompt a discussion of supplemental screening.

Unfortunately, there is a well-established and large degree of inter- and intrareader variability in the assignment of breast density, particularly among less-experienced readers (14–16). In a recent publication (17), we evaluated density assessment in clinical practice with 83 breast imaging radiologists in National Cancer Institute's Population-Based Research Optimizing Screening through Personalized Regimens, or PROSPR, network, which included data from more than 200000 screening mammograms among more than 145000 women. The rate of assignment to a "dense" category (category 3 or 4) by each individual radiologist ranged from 6.3% of screening examinations up to 84.5%. Statistical adjustment for patient variables such as age, body mass index, race, and ethnicity in each radiologist's patient mix had very little effect on this striking variability. In addition, among women with consecutive mammograms interpreted by different radiologists throughout an average span of 1.2 years, there was 17.2% discordance in dense versus nondense category assignments, with women shifting categories in both directions.

Most of the literature published to date linking breast density to mammographic screening performance and breast cancer risk has been based on clinical breast density assignments determined by using the fourth edition of BI-RADS or other subjective area-based assessments with the use of computer-assisted methods (eg, Cumulus; University of Toronto, Toronto, Canada) (1,4,18). With the introduction of the fifth edition of BI-RADS and the accompanying new density definitions emphasizing the masking effect of areas of increased breast density rather than the overall amount of dense breast tissue, there is concern that there will be even more variability in density assignments. Currently, there are few data on whether the assignment of a qualitative BI-RADS breast density category based on fifth edition definitions will be more or less variable than the estimate of percentage of the area of breast density used for assignment of fourth edition BI-RADS breast density categories, although one may argue that attempting to visually characterize local areas of dense tissue may be even more subjective than providing a global assessment of the overall amount of dense tissue. Indeed, a recent study (19) comparing repeat ratings with both the fourth and fifth edition BI-RADS density categories in the same patient set supported this assumption. The study results showed that interreader agreement, as assessed with κ statistics, was only 0.57 (95% confidence interval [CI]: 0.53, 0.61) with the fifth edition compared with 0.65 (95% CI: 0.61, 0.69) with the fourth edition ($P = .006$). In addition, when the fifth edition definitions were used, there was an approximately 10% increase in women deemed to have dense breasts compared with the use of the fourth edition definitions ($P < .0001$) (19).

The rapid implementation of digital breast tomosynthesis in breast clinics may also affect the clinical assessment of breast density, particularly because practices are increasingly replacing the conventional digital mammography component in favor of the dose-saving synthetic, two-dimensional images derived from the tomosynthesis acquisition. Synthetic two-dimensional images may have a very different appearance than conventional digital mammograms (20), and in two recent studies reporting outcomes of synthetic two-dimensional and digital breast tomosynthesis screening, there was a significant downgrade of BI-RADS density categories seen with synthetic two-dimensional tomosynthesis screening, with more women deemed as having nondense breasts compared with populations screened with either digital mammography alone or with combined digital mammography and digital breast tomosynthesis (21,22).

Other factors may also affect the subjective assessment of breast density. Breast density may appear quite variable with the different image formats, synthetic two-dimensional or standard-dose mammograms that are prepared for presentation, depending on which equipment and which reconstruction algorithms are used (23). Finally, results of studies have shown that the assignment of density categories by radiologists is also affected by the enactment of breast density notification legislation in their states. In one study (24), in the months immediately after the enactment of density legislation, radiologists classified fewer patients as having dense breasts.

What does this variability in density assignments mean for an individual woman? Because approximately 80% of American women fall somewhere in the midrange of breast density (ie, in BI-RADS categories 2 to 3 or b or c) (25), the chance of a woman with intermediate density being upgraded or downgraded across the threshold to be considered as having dense breasts is highly dependent on multiple factors, such as which radiologist interpreted her mammographic study; whether BI-RADS fourth or fifth edition density definitions were used; and potentially, how the woman is imaged (digital mammography, digital mammography and digital breast tomosyntheses,

https://doi.org/10.1148/radiol.2017170644

Content code: BR

Radiology 2018; 286:401–404

Abbreviations:

BI-RADS = Breast Imaging and Reporting Data System
CI = confidence interval

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

or synthetic two-dimensional and digital breast tomosynthesis). This variability in density assignments may not only affect how often a women gets screened and with which modality or modalities but also whether or not she is considered to be at high risk, which could prompt additional targeted risk-reduction interventions. In addition, risk assessment models that incorporate breast density were developed primarily on the basis of the fourth edition BI-RADS density definitions (and very recently some with selected vendor-specific quantitative measures). It is unclear that fifth edition BI-RADS definitions will generate similar risk estimates.

To overcome the subjective nature of visual BI-RADS breast density assessment, fully automated software methods have been developed that generate robust and reproducible quantitative measures (26–28). U.S. Food and Drug Administration clearance has been obtained for some of these methods, and they are widely available with commercially implemented software that usually is available as an add-on to standard digital mammography systems. In addition, tools have been made available by the research community that can also be applied to the routinely stored processed (for presentation) mammograms, in addition to the raw (for processing) images typically required by the commercial vendor software. Such software packages can provide both absolute as well as percentage of area–based estimates of the mammographically dense tissue, including categorical BI-RADS–like assessments, and some algorithms can also provide volumetric breast density measures as a potentially more accurate assessment of the amount of fibroglandular (ie, dense) tissue in the breast. In addition to generating reproducible, quantitative metrics, these methods allow estimation of a continuous score, rather than a categorical assessment, which may allow for more granular density-based risk stratification for women. Results of recent retrospective studies (29,30) have demonstrated strong associations between volumetric measures of breast density with both cancer masking and cancer risk. For example, in a study sample of 1911 women with cancer and 4170 control subjects matched for age, race, examination date, and mammography machine, Brandt et al (31) showed that clinical BI-RADS density assessment and automated measures of volumetric breast density assessment have similar associations with breast cancer (odds ratios: BI-RADS, 2.3 [95% CI: 1.9, 2.8] vs volumetric breast density, 1.9 [95% CI: 1.5, 2.5]). In a retrospective study of screening-detected versus interval breast cancers, Destounis et al (32) showed that quantitative volumetric measures of breast density have a stronger association with the rate of development of interval cancers when compared with clinical BI-RADS assessment (odds ratios: BI-RADS, 3.54 [95% CI: 1.55, 8.10] vs volumetric breast density, 4.51 [95% CI: 1.92, 10.61]), suggesting potentially higher sensitivity in capturing the likelihood of a breast cancer being masked on a screening mammogram. Although results from large, prospective studies are lacking, fully-automated quantitative measures come with the benefit of high reproducibility. In a recent reader study including 1000 mammograms from 500 paired sequential screening examinations from the same patient, Holland et al (33) demonstrated high agreement of 90.4% in density assessments for the paired sets of mammograms by using automated quantitative measures compared with lower agreement (range, 86.2%–89.2% for the individual readers) for subjective assessments ($P < .05$) .

In addition, the broad availability of digital imaging has also availed the opportunity to further develop more refined quantitative measures of breast texture or complexity. Such measures go beyond merely calculating the area or volume of the dense tissue to also provide an assessment of the degree and pattern of breast density by using quantitative features of parenchymal texture and complexity beyond the dichotomous assessment of dense versus nondense tissue, including mapping of the dense parenchymal areas that could mask a cancer, which potentially is similar to the notion of the BI-RADS fifth edition definitions. These quantitative measures may ultimately augment the coarser assessment of conventional breast density ratings.

The large variability in the assignment of visually assessed breast density has wide-reaching implication for the individual woman making personal screening choices, health care providers who counsel women, and policymakers determining screening recommendations. In 2012, the American College of Radiology issued a statement warning about the potential unintended harms of mandatory breast density notification due in part to the lack of reproducibility of visual breast density assessment (34). While all women should be aware of their own breast density and cancer risk estimation, robust and reproducible metrics are needed to capture both the potential for areas of masking as well as the overall amount of glandular tissue and how it relates to risk estimation to guide decisions on supplemental screening for the individual woman. At the present, it is clear from the subjective nature of BI-RADS density assessments that the fate of a woman's referral for supplemental screening may be more dependent on who interprets her mammogram than on the actual amount and distribution of her breast glandular tissue.

Although research is yet to provide fully conclusive results, we believe that the adoption in breast imaging clinics of automated, U.S. Food and Drug Administration–cleared methods for quantification of mammographic density will greatly reduce the variability of breast density ratings. The implementation of such quantitative measures will, therefore, help not only in the development and standardization of density thresholds that prompt supplemental screening, but also in the refinement of breast cancer risk assessment and personalized screening algorithms. This effort will require continued collaborations among academics, radiology professional societies, and industry to develop quantitative software that is affordable, well validated, and broadly available. Ultimately, the incorporation of automated quantitative measures of breast density will lead to more effective clinical care and more robust outcomes research than the current,

subjective assignment of BI-RADS density categories by providing reproducible estimates of both the risk of masking a cancer as well as the risk of developing breast cancer—two important factors in determining personalized breast cancer screening algorithms (35).

Disclosures of Conflicts of Interest: E.F.C. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: grant from NCI, personal fees from Hologic. Other relationships: disclosed no relevant relationships.**B.L.S.** disclosed no relevant relationships. **D.K.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: U.S. patent 12/488871 issued. Other relationships: disclosed no relevant relationships.

References

1. Kerlikowske K, Zhu W, Tosteson AN, et al. Identifying women with dense breasts at high risk for interval cancer: a cohort study. Ann Intern Med 2015;162(10):673–681.
2. Harvey JA, Yaffe MJ, D'Orsi C, Sickles EA. Density and breast cancer risk. Radiology 2013;267(2):657–658.
3. Ng KH, Lau S. Vision 20/20: Mammographic breast density and its clinical applications. Med Phys 2015;42(12):7059–7077.
4. Boyd NF, Martin LJ, Bronskill M, Yaffe MJ, Duric N, Minkin S. Breast tissue composition and susceptibility to breast cancer. J Natl Cancer Inst 2010;102(16):1224–1237.
5. Kressin NR, Gunn CM, Battaglia TA. Content, readability, and understandability of dense breast notifications by state. JAMA 2016;315(16):1786–1788.
6. Slanetz PJ, Freer PE, Birdwell RL. Breast-density legislation--practical considerations. N Engl J Med 2015;372(7):593–595.
7. Breast Density and Mammography Reporting Act of 2015, §370. 2015.
8. Haas JS, Kaplan CP. The divide between breast density notification laws and evidence-based guidelines for breast cancer screening: legislating practice. JAMA Intern Med 2015;175(9):1439–1440.
9. Burnside ES, Sickles EA, Bassett LW, et al. The ACR BI-RADS experience: learning from history. J Am Coll Radiol 2009;6(12):851–860.
10. American College of Radiology. ACR BI-RADS Atlas—Mammography. 4th ed. Reston, Va: American College of Radiology, 2003.
11. American College of Radiology. ACR BI-RADS Atlas—Mammography. 5th ed. Reston, Va: American College of Radiology, 2013.
12. Winkler NS, Raza S, Mackesy M, Birdwell RL. Breast density: clinical implications and assessment methods. RadioGraphics 2015; 35(2):316–324.
13. Ekpo EU, Ujong UP, Mello-Thoms C, McEntee MF. Assessment of interradiologist agreement regarding mammographic breast density classification using the fifth edition of the BI-RADS Atlas. AJR Am J Roentgenol 2016;206(5):1119–1123.
14. Ciatto S, Bernardi D, Calabrese M, et al. A first evaluation of breast radiological density assessment by QUANTRA software as compared to visual classification. Breast 2012; 21(4):503–506.
15. Ng KH, Yip CH, Taib NA. Standardisation of clinical breast-density measurement. Lancet Oncol 2012;13(4):334–336.
16. Ooms EA, Zonderland HM, Eijkemans MJ, et al. Mammography: interobserver variability in breast density assessment. Breast 2007; 16(6):568–576.
17. Sprague BL, Conant EF, Onega T, et al. Variation in mammographic breast density assessments among radiologists in clinical practice: a multicenter observational study. Ann Intern Med 2016;165(7):457–464.
18. Kerlikowske K, Zhu W, Hubbard RA, et al. Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. JAMA Intern Med 2013;173(9):807–816.
19. Irshad A, Leddy R, Ackerman S, et al. Effects of changes in BI-RADS density assessment guidelines (fourth versus fifth edition) on breast density assessment: intra- and inter-reader agreements and density distribution. AJR Am J Roentgenol 2016;207(6):1366–1371.
20. Zuckerman S, Maidment AD, Weinstein SP, McDonald E, Conant EF. Imaging with synthetic mammography: differences, advantages and pitfalls compared to digital mammography. Radiology 2018;28#(#):###–###.
21. Zuckerman SP, Conant EF, Keller BM, et al. Implementation of synthesized two-dimensional mammography in a population-based digital breast tomosynthesis screening program. Radiology 2016;281(3):730–736.
22. Aujero MP, Gavenonis SC, Benjamin R, Zhang Z, Holt JS. Clinical performance of synthesized two-dimensional mammography combined with tomosynthesis in a large screening population. Radiology 2017;283(1):70–76.
23. Keller BM, Nathan DL, Gavenonis SC, Chen J, Conant EF, Kontos D. Reader variability in breast density estimation from full-field digital mammograms: the effect of image postprocessing on relative and absolute measures. Acad Radiol 2013;20(5):560–568.
24. Bahl M, Baker JA, Bhargavan-Chatfield M, Brandt EK, Ghate SV. Impact of breast density notification legislation on radiologists' practices of reporting breast density: a multistate study. Radiology 2016;280(3):701–706.
25. Sprague BL, Gangnon RE, Burt V, et al. Prevalence of mammographically dense breasts in the United States. J Natl Cancer Inst 2014;106(10):dju255.
26. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. Automated analysis of mammographic densities. Phys Med Biol 1996;41(5):909–923.
27. Heine JJ, Carston MJ, Scott CG, et al. An automated approach for estimation of breast density. Cancer Epidemiol Biomarkers Prev 2008;17(11):3090–3097.
28. Keller BM, Nathan DL, Wang Y, et al. Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation. Med Phys 2012;39(8):4903–4917.
29. Heine JJ, Scott CG, Sellers TA, et al. A novel automated mammographic density measure and breast cancer risk. J Natl Cancer Inst 2012;104(13):1028–1037.
30. Keller BM, Chen J, Daye D, Conant EF, Kontos D. Preliminary evaluation of the publicly available Laboratory for Breast Radiodensity Assessment (LIBRA) software tool: comparison of fully automated area and volumetric density measures in a case-control study with digital mammography. Breast Cancer Res 2015;17(1):117.
31. Brandt KR, Scott CG, Ma L, et al. Comparison of clinical and automated breast density measurements: implications for risk prediction and supplemental screening. Radiology 2016;279(3):710–719.
32. Destounis S, Johnston L, Highnam R, Arieno A, Morgan R, Chan A. Using volumetric breast density to quantify the potential masking risk of mammographic density. AJR Am J Roentgenol 2017;208(1):222–227.
33. Holland K, van Zelst J, den Heeten GJ, et al. Consistency of breast density categories in serial screening mammograms: A comparison between automated and human assessment. Breast 2016;29:49–54.
34. American College of Radiology. ACR statement on reporting breast density in mammography reports and patient summaries. http://www.acr.org/About-Us/Media-Center/Position-Statements/Position-Statements-Folder/Statement-on-Reporting-Breast-Density-in-Mammography-Reports-and-Patient-Summaries. Published 2012. Accessed March 16, 2016.
35. Gastounioti A, Conant EF, Kontos D. Beyond breast density: a review on the advancing role of parenchymal texture analysis in breast cancer risk assessment. Breast Cancer Res 2016;18(1):91.