European Journal of Radiology 82 (2013) 1724–1730

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Variability of breast density assessment in short-term reimaging with digital mammography

Won Hwa Kim [a], Woo Kyung Moon [a,*], Sun Mi Kim [b], Ann Yi [c], Jung Min Chang [a], Hye Ryoung Koo [a], Su Hyun Lee [a], Nariya Cho [a]

[a] Department of Radiology, Seoul National University Hospital, Seoul, Republic of Korea
[b] Department of Radiology, Seoul National University Bundang Hospital, Gyeonggi-do, Republic of Korea
[c] Department of Radiology, Seoul Metropolitan Government Seoul National University, Boramae Medical Center, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 14 March 2013
Received in revised form 22 April 2013
Accepted 5 May 2013

Keywords:
Breast
Breast density
Mammography/methods
Mammography/utilization
Reproducibility of results

ABSTRACT

Objective: To evaluate the variability of breast density assessments in short-term reimaging with digital mammography.
Materials and methods: In 186 women, short term (mean interval, 27.6 days) serial digital mammograms including CC and MLO views were obtained without any treatment. Mammographic density assessments were performed by three blinded radiologists for Breast Imaging Report and Data System (BI-RADS, grades 1–4) and visual percentage density (PD) estimation, and by one radiologist for computer-aided PD estimation. The variability of assessments was analyzed according to the age, breast density, and mammography types by multivariate logistic regression.
Results: In BI-RADS assessments, 29% (161 of 558) of breast density categories were assessed differently after short-term reimaging and the mean absolute difference in PD for CC and MLO view was 7.6% and 8.1% for visual assessments, and 7.4% and 6.4% for computer-aided assessments, respectively. Among all computer-aided assessments, 29% (54 of 186) of CC view and 22% (41 of 186) of MLO view assessments had discrepancy over 10% in PD. Younger age (<50), greater breast density (grades 3 and 4), and different mammography types were significantly associated with the variability.
Conclusion: Considerable variability in breast density assessments occurred in short-term reimaging with digital mammography, particularly in women with younger age and greater breast density and when examined using different types of mammography.

1. Introduction

Breast density on mammography is considered as an independent risk factor for breast cancer as well as a marker of decreased sensitivity of mammographic diagnoses [1]. Several studies in which quantitative methods for assessing density were used have shown that women with dense breasts have a four- to five-fold increased risk of breast cancer compared to women with fatty breasts [2]. There is also evidence that hormonal therapies, including tamoxifen, can change both mammographic density and the risk of breast cancer [3–6]. In one study, women in the tamoxifen group who experienced a 10% or greater reduction in breast density had 63% reduction in breast cancer risk (odds ratio, 0.37; $P = 0.002$), whereas those who took tamoxifen but experienced less than a 10% reduction in breast density had no risk reduction (odds ratio, 1.13; $P = 0.60$) [3]. Age and baseline density on mammography was associated with breast density change after treatment. In breast cancer prevention trials, the change in breast density on mammography has been increasingly used to monitor the effect of intervention [3–7]. However, for mammographic density change to be used as a surrogate biomarker of the effect of intervention that increase or decrease breast cancer risk, it is necessary to evaluate the validity of the current density assessment methods in advance [8].

Breast Imaging Reporting and Data System (BI-RADS) density categories or visual and computer-aided estimation of percentage density (PD) are commonly used to assess mammographic density [9–12]. In both film screen and digital mammography, computer-aided assessment of PD with interactive thresholding software has been known to be the most consistent method with higher inter- and intraobserver reproducibility [11–14]. However, reimaging itself, can lead to changes in mammographic density assessment, due to different positioning, breast compression, exposure factors, and image processing conditions [15,16]. Previous study also found that PD was significantly lower for digital

* Corresponding author at: Department of Radiology, Seoul National University Hospital, 101 Daehangno, Jongno-gu, Seoul 110-744, Republic of Korea. Tel.: +82 2 2072 3928; fax: +82 2 743 6385.
E-mail address: moonwk@snu.ac.kr (W.K. Moon).

0720-048X/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ejrad.2013.05.004

mammograms compared to film screen mammograms (mean PD, 32.2% vs. 40.3%) [5,17]. Moving into the digital era, as institutions adopt various types of digital mammography systems and researchers try to conduct large-scale prevention trials in multi-institutional settings, mammograms obtained from various digital mammography systems are expected to be more involved in studies of breast density [7]. However, to our knowledge, no report on the variability of breast density assessments with various types of digital mammography has been published. Moreover, we are unaware of any investigation into the effects of short-term repeat image acquisition in the same breast on density assessment with mammography.

Thus, the purpose of our study was to evaluate the variability of breast density assessments in short-term reimaging with digital mammography.

2. Materials and methods

2.1. Subjects and mammograms

This retrospective study was approved by the Institutional Review Board of our institution, and informed patient consent was waived. Between January 2011 and December 2011, 211 consecutive women who had their first mammograms performed in other facilities underwent second mammograms in our institution with an interval of less than 2 months. Patients were referred for their clinical symptom in their breasts ($n = 104$) or suspicious finding on their first mammography ($n = 107$). Between first and second mammography, percutaneous image-guided core-needle biopsy was performed in all the 211 women and 101 women were diagnosed with breast cancer. In these 211 women, second mammograms were performed at the clinician's request to evaluate preoperative tumor extent ($n = 101$) or follow-up after benign biopsy ($n = 110$). For all patients, contralateral, unaffected breasts only were included to exclude breasts with post-biopsy changes. Patients who had undergone hormonal replacement therapy ($n = 19$) were excluded. In addition, six patients with abnormal clinical or mammographic findings on both breasts or had history of biopsy or surgery on both breasts were excluded. Finally, 186 patients with at least one unaffected breast were included in this study. The mean age of women was 48.0 years (standard deviation, 9.3 years; range, 26–78 years) and 120 (64.5%) of the women were premenopausal and 66 (35.5%) were postmenopausal. The mean interval time between first and second mammograms was 27.2 days (range, 5–15 days [$n = 47$], 16–30 days [$n = 69$], 31–45 days [$n = 44$], and 46–60 days [$n = 26$]).

All mammograms were taken from facilities which fulfill the standard of Mammography Quality Standards Requirements in Korea. First and second mammograms comprised two standard views (craniocaudal [CC] and mediolateral oblique [MLO]) for each breast. First mammography was performed in eight outside facilities using three types of digital mammography systems of computed radiography (CR) ($n = 115$), direct radiography with direct conversion type (direct DR) ($n = 60$), and direct radiography with indirect conversion type (indirect DR) ($n = 11$). The CR images were obtained with four CR systems manufactured by Fuji (Tokyo, Japan) ($n = 37$), Konica Minolta (Tokyo, Japan) ($n = 27$), Kodak (Rochester, NY) ($n = 26$), and Agfa (Agfa Gevaert N.V., Mortsel, Belgium) ($n = 25$). The direct DR images were obtained with Lorad Selenia (Hologic Inc., Bedford, MA) system. The indirect DR images were obtained with two DR systems; Senographe DS ($n = 7$) or Senographe 2000D ($n = 4$) (GE Medical Systems, Milwaukee, WI).

Second mammography was performed in our institution using direct DR ($n = 117$) and indirect DR ($n = 69$) systems. The direct DR system was Lorad Selenia (Hologic Inc.) and indirect DR system was Senographe DS 2000D (GE Medical Systems). Premium View (PV, GE Medical Systems) was applied as a post-processing algorithm to all images for indirect DR images in our institution and outside facility.

For breast density assessments, mammograms of the unaffected breasts (139 left breasts and 47 right breasts) were arranged in random order. To evaluate the intraobserver variability, 40 mammograms (20 pairs) from approximately 10% of patients ($n = 20$) were included twice in the data set randomly.

2.2. BI-RADS, visual, and computer-aided assessments of mammographic density

Three radiologists independently assigned a set of mammograms including CC view and MLO view from each patient into one of the four BI-RADS density categories (1 = almost entirely fatty, 2 = scattered fibroglandular densities, 3 = heterogeneously dense, 4 = extremely dense) [18]. Next, another three radiologists visually estimated percentage density (PD) on each mammogram using a set of reference mammograms by 10% in density modified from six categories [2,19]. Radiologists were blinded to the time sequence (first or second mammogram), technical information including type of mammography system, and demographic findings of subjects. The radiologists were specialized in breast imaging (range, 6–25 years; mean, 11 years) and have interpreted at least 2000 mammograms per year. Both BI-RADS assessment and visual estimation of PD were performed using a 5-megapixel (2560 × 2048 pixels) liquid crystal display system (ME511L, Totoku Electric) with picture archiving and communication system (PACS). For each mammogram, the radiologists were allowed to adjust the window and level settings on the computer screen.

For computer-aided assessments, a radiologist who specialized in breast imaging and had 2 years of experience with the quantitative density assessment measured each mammogram using a commercial software program (Cumulus®, Toronto, ON). Using this program, the total breast and glandular area were defined using a thresholding technique [20]. PD was calculated as the glandular area divided by the total breast area.

Since the positioning difference could be the major factor to effect the discrepant results in density assessments in serial mammograms [15], two radiologists who were specialized in breast imaging (6 and 25 years) retrospectively determined if CC and MLO serial mammograms had identical or dissimilar positioning. Mammograms with dissimilar positioning were defined as mammograms with one or more discrepancy in the followings details: inferior extent of the pectoral muscle relative to the posterior nipple line, presence of fibroglandular tissue at the posterior edge of the film, profile of the nipple, and opening of inframammary fold. A total of 155 pairs of mammograms were assessed as having identical positioning, and 31 pairs were assessed as having the dissimilar positioning.

2.3. Statistical analysis

Variability in three breast density assessment methods was appraised from the discrepancies between first and second mammograms. To exclude the effect of mammographic positioning on density assessment, we performed analyses for mammograms with only identical positioning and the results were provided separately.

For the variability of BI-RADS assessments, the frequencies of discrepant BI-RADS assessments categories of the breast between first and second mammograms among all paired assessment were quantified. For the variability of visual and computer-aided PD assessments, mean absolute and relative differences were computed for CC and MLO views. The absolute difference was the absolute value of relative difference between serial mammograms.