

Density is in the eye of the beholder: visual versus semi-automated assessment of breast density on standard mammograms

M. B. I. Lobbes · J. P. M. Cleutjens · V. Lima Passos ·
C. Frotscher · M. J. Lahaye · K. B. M. I. Keymeulen ·
R. G. Beets-Tan · J. Wildberger · C. Boetes

Received: 2 August 2011 / Accepted: 24 October 2011 / Published online: 20 November 2011
© European Society of Radiology 2011

Abstract

Objectives Visual inspection is generally used to assess breast density. Our study aim was to compare visual assessment of breast density of experienced and inexperienced readers with semi-automated analysis of breast density.

Methods Breast density was assessed by an experienced and an inexperienced reader in 200 mammograms and scored according to the quantitative BI-RADS classification. Breast density was also assessed by dedicated software using a semi-automated thresholding technique. Agreement between breast density classification of both readers as well as agreement between their assessment versus the semi-automated analysis as reference standard was expressed as the weighted kappa value.

Results Using the semi-automated analysis, agreement between breast density measurements of both breasts in

both projections was excellent ($ICC > 0.9$, $P < 0.0001$). Reproducibility of the semi-automated analysis was excellent ($ICC > 0.8$, $P < 0.0001$). The experienced reader correctly classified the BI-RADS breast density classification in 58.5% of the cases. Classification was overestimated in 35.5% of the cases and underestimated in 6.0% of the cases. Results of the inexperienced reader were less accurate. Agreement between the classification of both readers versus the semi-automated analysis was considered only moderate with weighted kappa values of 0.367 (experienced reader) and 0.232 (inexperienced reader). **Conclusion** Visual assessment of breast density on mammograms is inaccurate and observer-dependent.

Keywords Breast cancer · Breast neoplasms · Breast density · Mammography · Threshold method

M. B. I. Lobbes (✉) · C. Frotscher · M. J. Lahaye ·
R. G. Beets-Tan · J. Wildberger · C. Boetes
Department of Radiology, Maastricht University Medical Center,
P.O.Box 5800, 6202 AZ Maastricht, The Netherlands
e-mail: marc.lobbes@mumc.nl

J. P. M. Cleutjens
Department of Pathology, Maastricht University Medical Center,
Maastricht, The Netherlands

V. Lima Passos
Department of Methodology and Statistics, Maastricht University,
Maastricht, The Netherlands

K. B. M. I. Keymeulen
Department of Surgery, Maastricht University Medical Center,
Maastricht, The Netherlands

J. Wildberger · C. Boetes
GROW School for Oncology and Developmental Biology,
Maastricht, The Netherlands

Introduction

Breast density on a conventional mammogram is mainly composed of two components: fatty and fibroglandular tissue. Fat has a lower X-ray attenuation coefficient than fibroglandular tissue, which is composed of connective tissue and epithelial cells. Therefore, on a conventional mammogram, fat appears dark, whereas fibroglandular tissue appears white [1].

In 1976, Wolfe was one of the first to demonstrate an association between breast density and the risk of developing breast cancer [2]. This association was confirmed by McCormack et al. In a large meta-analysis including 42 studies, they showed that the risk of developing breast cancer is increased in dense breasts. The magnitude of this risk can be as high as 4.6-fold for the most dense breasts

compared with the least dense category [3]. Recently, Boyd et al. found similar results, showing that women with fibroglandular tissue in more than 75% of the mammogram had an increased risk of developing breast cancer when compared with women with less than 10% breast density in the mammogram (odds ratio 4.7) [4]. In addition to this association, several studies have shown that the sensitivity of the mammogram for detecting breast cancer is decreased in dense breasts [5, 6]. For example, Carney et al. demonstrated in a study of 463,372 screening mammograms that the sensitivity and specificity for fatty breast mammography were 88 and 97% respectively. However, these numbers decreased in extremely dense breasts (sensitivity 62%, specificity 90%) [5].

Several qualitative methods to assess breast density have been proposed in the past years, for example the Wolfe and Tabar classifications [2, 7]. In 2003, the Breast Imaging Reporting and Data System (BI-RADS) quantitative classification was introduced [8]. Of these different breast density classification systems, the BI-RADS classification is widely used in clinical radiology practice. BI-RADS guidelines recommend an assessment of breast composition to be included in the analysis of the mammogram [8]. According to the quantitative BI-RADS breast density classification, breast density on mammogram can be divided into four categories: (1) almost entirely fat ($\leq 24\%$), (2) scattered fibroglandular densities (25–49%), (3) heterogeneously dense (50–74%), and (4) extremely dense ($\geq 75\%$).

In the past, (semi)automated systems to assess breast density have been proposed [1]. Nonetheless, they have not yet attained a strong hold in screening or clinical settings, mainly because these systems are time and labour intensive and require dedicated software programs and operator training [9]. In screening and clinical settings, radiologists therefore generally assess breast density in a visual manner ('eyeballing') as part of the total evaluation of the mammogram. In this study, we aimed to evaluate the accuracy of visual assessment of breast density (on standard digital mammograms) for both experienced and inexperienced readers as compared to the semi-automated assessment of breast density using a dedicated software program.

Materials and methods

Study inclusion

Consecutive digital mammograms of 200 women (mean age 51.6 years, range 23.9–91.2 years) were included. For this study, ethics approval and informed consent for the use of the (coded) images was waived according to Dutch law. All images were acquired on a dedicated mammography system (Senographe Essential, GE Healthcare, Waukesha,

WI, USA). The standard craniocaudal (CC) and mediolateral oblique (MLO) imaging projections were used. Both breasts had to be imaged for comparison purposes in the final analysis. Therefore, patients with (unilateral) mastectomy were excluded. Additionally excluded were mammograms of breasts that had undergone surgery of any kind, for instance lumpectomy because of breast cancer or benign breast diseases or breast reduction.

In the Netherlands, research covered by the Medical Research Involving Human Subjects Act must be submitted to an accredited medical ethics committee for approval. However, the Act does not cover retrospective research using (coded) data from patient's medical record or patient images. Therefore, our medical ethics committee concluded that the research proposal of the current study does not, under Dutch law, require medical ethics approval because there is no extra burden placed on research subjects for this study (decision number: METC 11-4-049).

Breast density was visually assessed on our dedicated mammographic workstations by an experienced mammoradiologist (C.B., 18 years experience) and by an (inexperienced) senior resident in radiology (C.F., 2 years of experience). Breast densities had to be scored in one of the four quantitative BI-RADS categories. Both readers did not receive any specific training in assessing breast density and were blinded to each other's results, as well as to the results of the semi-automated analysis. In addition, both readers had never interacted with the software programme before.

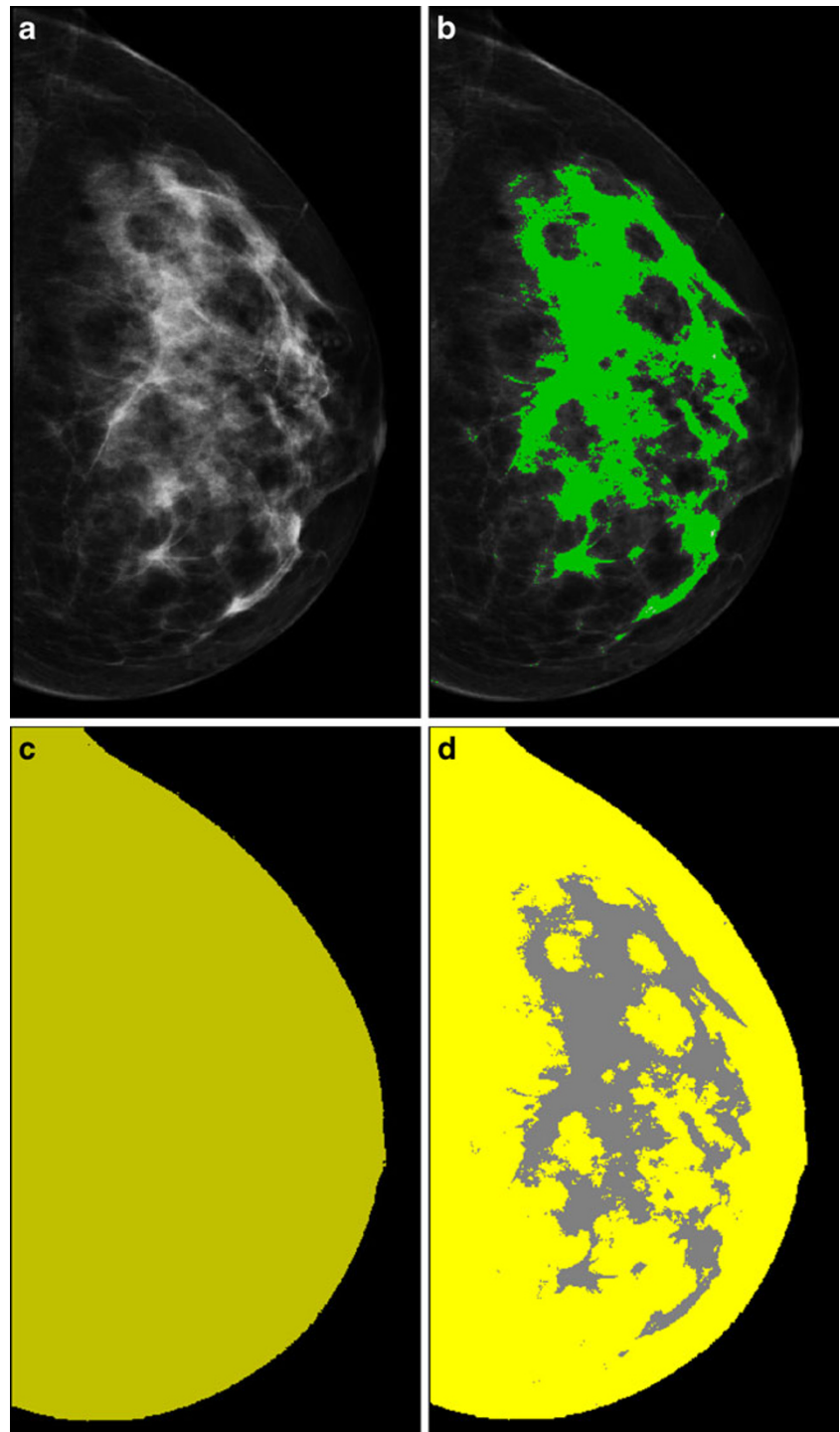
Breast density was analysed by using a so-called operator-dependent thresholding technique (Leica Qwin version 3, Leica Microsystems, Cambridge, UK). This approach is an extensively evaluated method for assessing breast density on standard mammograms [10–12]. More importantly, the association between increased risk for breast cancer in dense breasts has been demonstrated using the thresholding approach [4].

In short, this technique translates the brightness of each pixel in a digitised mammographic image into a grey-level value. An operator selects a pixel of fibroglandular tissue. The grey-level value of this pixel is obtained, and appropriate threshold grey-level values are selected by the same operator to create an overlay representing the fibroglandular tissue within the breast. Next, the total breast area is automatically identified, since regions outside of the breast tissue area are completely black, i.e. they have a grey-level pixel value of 0. In the mediolateral oblique projections of the breast, the pectoral muscle is excluded in an additional, operator-dependent segmentation. The number of pixels is calculated for both the colour overlay of the fibroglandular tissue and the total area of the breast. Breast density is presented as the area percentage of the fibroglandular tissue within the breast and is calculated by dividing the number of pixels in the colour overlay by the

number of pixels in the total breast area and multiplying by 100 (Fig. 1). The thresholding approach in general was described more in detail by Yaffe [1]. The semi-automated analysis was performed by a experienced operator and radiologist (M.B.I.L., 5 years of experience using the thresholding approach and this software program), who was blinded to the results of both mammography readers.

Intra-observer reproducibility was assessed by re-analysing a subgroup of randomly selected mammograms ($n=50$) with an appropriate interval between these analyses of >6 months. For this purpose, patients received a different coding. To determine inter-observer reproducibility, the same sample of mammograms was evaluated by an inexperienced operator, M.J.L. (no experience in using the thresholding technique

Fig. 1 Semi-automated detection of mammographic breast density. **a** Standard craniocaudal mammogram of the left breast. **b** Detection of fibroglandular tissue. **c** Detection of total breast tissue. **d** Graphic overlay of (b) and (c). In this particular example, the mammographic breast density was 22% (BI-RADS density category 1)



and the software programme, and receiving only 1 h of operator training by M.B.I.L.). The latter was blinded to the results of both mammography readers and the semi-automated analysis performed.

Statistical analysis

Statistical analyses were performed by using SPSS 17.0 (SPSS, Chicago, IL, USA) and the SAS statistical package (SAS Institute, Cary, NC, USA). The correlation between breast densities of both breasts (as determined by the semi-automated analysis) and of both CC and MLO projections was assessed by calculating the intra-class correlation coefficient (ICC). For the final analysis, the CC projection was used, since it avoids the need for additional segmentation to exclude the pectoral muscle in the MLO projection, which might create bias within the measurements. Mean values of the breast density in the CC projections were assessed and translated into one of the four corresponding quantitative BI-RADS categories. The inter- and intra-observer reproducibility of the semi-automated analyses was also calculated by using the ICC. Bland-Altman plots were used to show the agreement between the various semi-automated analyses.

Furthermore, the agreement of the quantitative BI-RADS categorisation between the experienced and inexperienced reader was compared and expressed as the weighted kappa value. Finally, the accuracy of this quantitative BI-RADS categorisation between the readers was evaluated (with the semi-automated analysis as reference standard) and again expressed as weighted kappa value. *P*-values ≤ 0.05 were considered statistically significant.

Results

Using the semi-automated analysis, the ICC of the left-sided breast density in both CC and MLO projection was excellent: 0.91 [95% confidence interval (CI) 0.88–0.93]. Similar excellent results were acquired for the right breast in both projections: ICC 0.91 (95% CI 0.88–0.93). The breast densities of the left and right breast were comparable (ICC 0.92, 95% CI 0.89–0.94) in the CC projection. For the MLO projection, the ICC for the breast densities of the left and right breast was excellent (0.91, 95% CI 0.89–0.93). All these results were highly significant (all $P < 0.0001$) and comparable with results previously published [3]. Bland-Altman plots of all these analyses showed very good agreement (Fig. 2). Although there is a slightly larger disagreement for more dense breasts, the observed differences are still acceptable.

Reproducibility of the semi-automated analysis in a subset of mammograms ($n=50$) proved to be good as well.

The ICC of both CC and MLO projections of the left and right breast were all highly significant ($P < 0.0001$) with values ranging from 0.82 to 0.88. Even more interesting is the inter-observer comparison of the semi-automated analysis. In these analyses, the ICC of both CC and MLO projections of the left and right breast were also highly significant ($P < 0.0001$) with values ranging from 0.80 to 0.88. Although the observed differences in breast densities were slightly higher than in the initial semi-automated analyses, Bland-Altman plots still showed good agreement of all these analyses with acceptable differences in breast density measurements (Figs. 3 and 4).

Table 1 shows the inter-observer agreement of the quantitative BI-RADS classification of both mammographic readers. There was a disagreement between the quantitative BI-RADS categorisations of the experienced and inexperienced reader. In 83 of the 200 cases (42%), a different BI-RADS density category was assigned to the mammograms. The agreement of the experienced and inexperienced reader was therefore only moderate with a kappa value of 0.52 (Table 1).

When compared with the results of the semi-automated analyses, the experienced reader agreed with the quantitative BI-RADS category in 58.5% of the cases. The classification was overestimated in 35.5% of the cases and underestimated in 6.0% of the cases. In comparison, the inexperienced reader agreed less (42.0%) and generally overestimated the quantitative BI-RADS classification more than the experienced reader (56.0%). Agreement between the classification of both readers versus the semi-automated analysis was poor to moderate with weighted kappa values of 0.367 (experienced reader) and 0.232 (inexperienced reader, Table 2).

Discussion

In this study, reliability of visual assessment of breast density for both experienced and inexperienced readers was evaluated, as compared to the semi-automated assessment of breast density using a dedicated software program. Our results showed that there was disagreement between the quantitative BI-RADS categorisation of the experienced and inexperienced readers. When compared to the semi-automated analysis, the experienced reader agreed with the quantitative BI-RADS classification in 58.5% of the cases. The classification was overestimated in 35.5% of the cases, and underestimated in 6.0% of the cases. Results of the inexperienced reader were less accurate. Furthermore, the semi-automated assessment of breast density showed good intra- and interobserver reproducibility.

Breast density is an important risk factor for breast cancer development, independent of other breast cancer risk

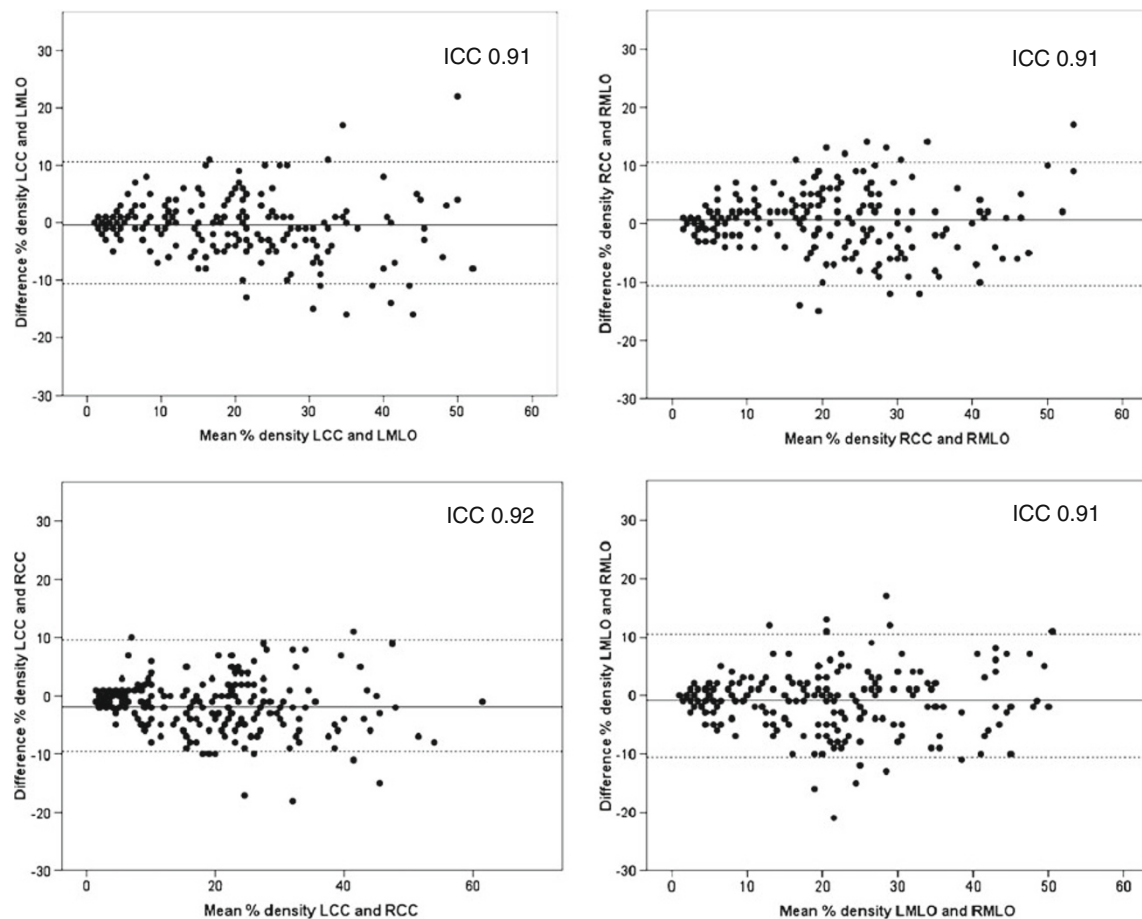


Fig. 2 Bland-Altman plots for the semi-automated analyses of breast densities of both breasts in two projections. The plots show very good agreement for all analyses performed. Intra-class correlation coefficients (ICC) of the measurements were highly significant

(all $P < 0.0001$). *Solid lines* represent the mean of the differences between the two analyses, *dotted lines* represent the boundaries of two times the standard deviation of the difference

factors [3]. Also, breast cancer is more difficult to detect in mammographically dense breasts [5, 6]. In our institution, mammograms are evaluated by radiologists and/or (supervised) residents, using the BI-RADS classification for breast density. However, our study results showed disagreement between radiologist (experienced reader) and resident (inexperienced reader) and that breast density is frequently overrated (even by a highly experienced reader). These findings are in line with previously published results [13]. In the majority of cases, the overestimation was only one BI-RADS category (data not shown). Although this may seem a negligible overestimation, a speculative (but nonetheless plausible) assumption is that overrating breast density might lead to more imaging [e.g. additional mammographic projections, ultrasound, or contrast-enhanced magnetic resonance (MR) mammography], more costs, and more patient anxiety.

Due to the improvements in MR mammography, it is worth considering which patients are at increased risk of developing breast cancer (i.e. patients with mammographically

dense breasts) and who might benefit from a shorter screening interval or additional MR mammography [14]. Despite the fact that this information is enclosed within the images, it is not used in current clinical settings or screening to identify high-risk patients, since the (visual) BI-RADS density classification is not suitable for the expression of breast cancer risk. Based on our current observations (which show a substantial disagreement between the visual and semi-automated assessment of breast density), we would prefer a (workstation) integrated (semi-)automated analysis of breast density to identify patients at high risk for developing breast cancer or in whom breast cancer is likely to be missed.

For this purpose, several (semi-)automated systems to assess breast density have been proposed in the past [1]. Of these, the so-called thresholding approach using (commercially available) software packages has been extensively studied in the past and is therefore frequently used in quantitative assessment of breast density [12]. Using this thresholding technique, Boyd et al. showed that women with mammographic breast density $\geq 75\%$ had an increased

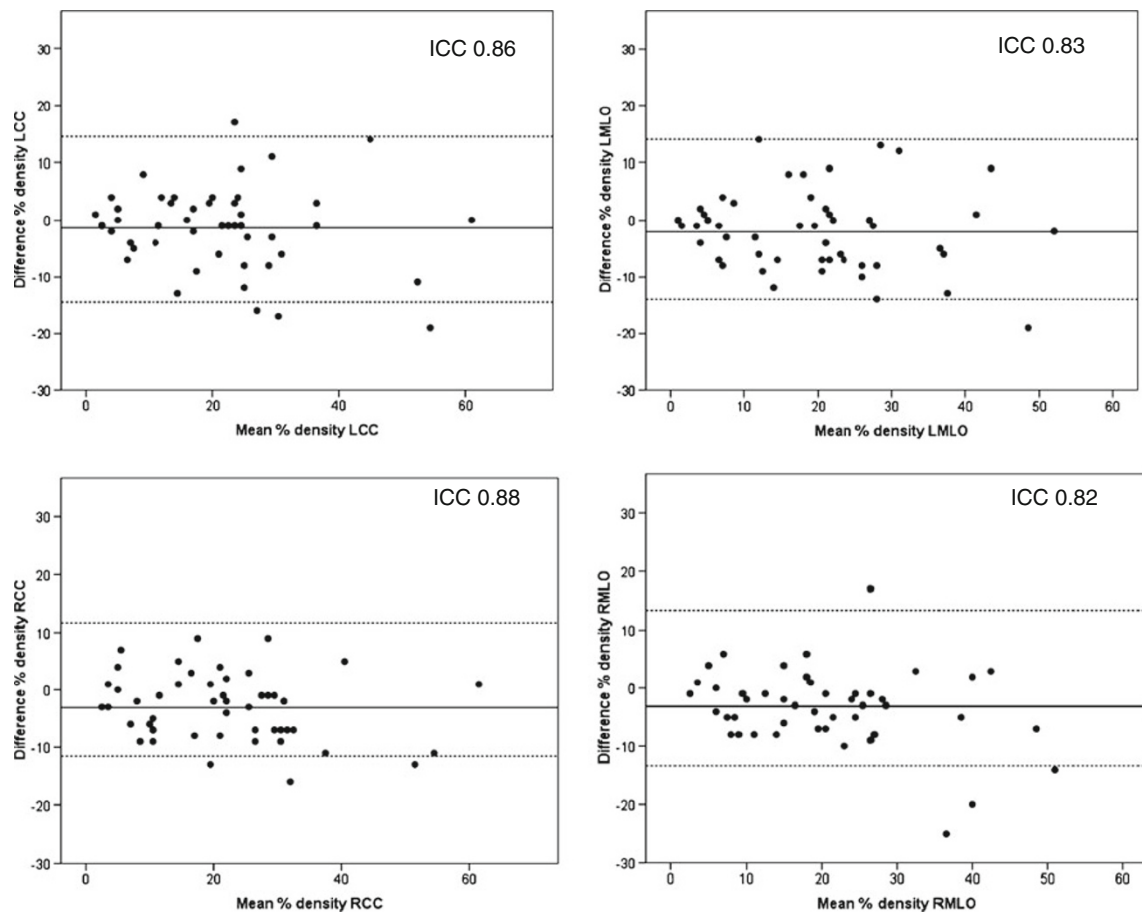


Fig. 3 Intra-observer agreement of the semi-automated analyses. Bland-Altman plots show good agreement for all analyses performed. Intra-class correlation coefficients (ICC) were highly significant

(all $P < 0.0001$). *Solid lines* represent the mean of the differences between the two analyses, *dotted lines* represent the boundaries of two times the standard deviation of the differences

risk of developing breast cancer when compared to women with breast density less than 10% (odds ratio 4.7, 95% confidence interval 3.0–7.4) [4]. We also used a thresholding approach using the Leica Qwin software package and showed similar results in breast densities of both breasts and in both projections [3]. In addition, our software demonstrated a good intra-observer agreement (Fig. 3). A major advantage of our semi-automated analysis is that even analysts with hardly any experience using our software can achieve good reproducibility (Fig. 4).

Although the thresholding approach is promising, it has several disadvantages. The assessment of the mammographic breast density using this technique is time and labour intensive. For example, our time for the assessment of breast density in a single patient study was estimated to be 5–8 min. Recently, Stone et al. demonstrated that assessing breast density of one breast in one projection is sufficient [15]. This suggestion is supported by the findings of our current study and enables the inclusion of women that have undergone unilateral breast surgery of any kind. Furthermore, the thresholding approach requires proper

operator-training to use the software, although our current results (based on our software program) suggest otherwise. In screening or clinical settings, this software needs to be integrated in the mammographic work stations. Because of these disadvantages, mammographic breast density is usually assessed visually, especially in screening/clinical settings and large studies in which a great number of mammograms need to be evaluated [16].

Previously, Martin et al. performed a similar study using a fully automated software package [17]. In this study mammograms of 65 women were analysed by seven radiologists. However, on close inspection, five of the radiologists had already interacted with the software programme, leaving two radiologists (albeit experienced in mammography reading) untrained for the breast density analysis. The breast density was overestimated by these two radiologists in 37% of the cases, as compared to 36% overestimation by the experienced radiologist in our study. So although it might be difficult to generalise our study results based on the results of two readers, they are in line with previously published results. In addition, there were

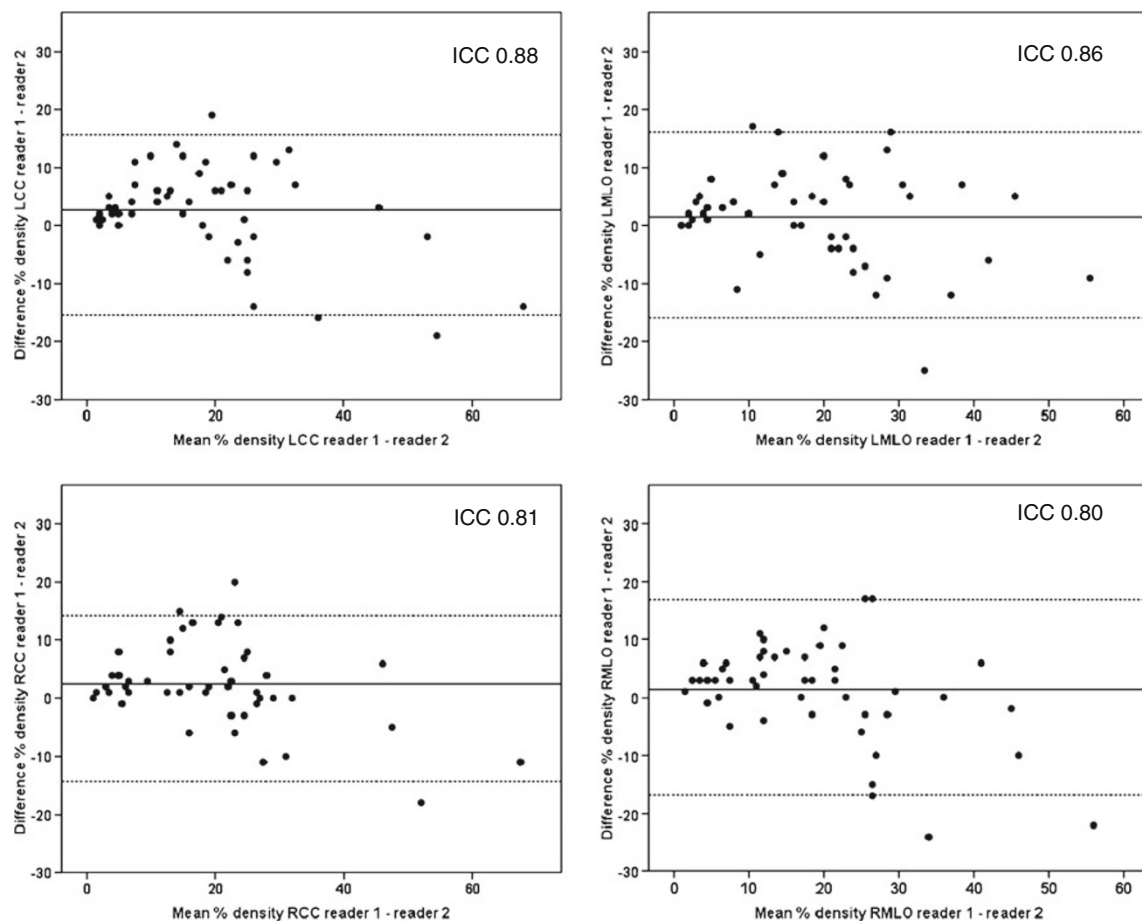


Fig. 4 Inter-observer agreement of the semi-automated analyses. Bland-Altman plots show good agreement for all analyses performed. Intra-class correlation coefficients (ICC) of the measurements were

highly significant (all $P < 0.0001$). *Solid lines* represent the mean of the differences between the two analyses, *dotted lines* represent the boundaries of two times the standard deviation of the differences

other interesting differences in the analyses performed by us and Martin et al. In the study of Martin et al., 6% of the images could not be analysed due to technical difficulties versus no technical difficulties in our study. Furthermore, the 95% limits of agreement between the trained radiologists and the automated analyses of the study of Martin et al. were rather large: -16% to $+27\%$. Visual inspection of

the Bland-Altman plots of our study show that our limits of agreement were much smaller (Fig. 2).

Our study included relatively small numbers of dense breasts, i.e. quantitative BI-RADS categories 3 and 4. The study of Martin et al. described similar results and included 15 additional women with dense breasts in their study after their initial inclusion of 50 women. Previously published

Table 1 Inter-reader agreement of breast density according to BI-RADS classifications. Results of the experienced reader are presented in the columns, results of the inexperienced reader are presented in the rows

		Experienced reader				
		BI-RADS 1	BI-RADS 2	BI-RADS 3	BI-RADS 4	Total
Inexperienced reader	BI-RADS 1	66 (33.0%)	2 (1.0%)	2 (1.0%)	0 (0.0%)	70 (35.0%)
	BI-RADS 2	19 (9.5%)	36 (18.0%)	1 (0.5%)	0 (0.0%)	56 (28.0%)
	BI-RADS 3	5 (2.5%)	38 (19.0%)	10 (5.0%)	0 (0.0%)	53 (26.5%)
	BI-RADS 4	1 (0.5%)	3 (1.5%)	14 (7.0%)	3 (1.5%)	21 (10.5%)
	Total	91 (45.5%)	79 (39.5%)	27 (13.5%)	3 (1.5%)	200 (100.0%)

The overall weighted kappa was 0.521, a moderate value (95% confidence interval 0.446–0.597)

Table 2 Classification of the results of the experienced and inexperienced readers and the software analysis. The agreement of the respective readers' results and the semi-automated software is presented as the weighted kappa value

	Experienced reader	Inexperienced reader	Software analysis
Total patients classified	200 (100.0%)	200 (100.0%)	200 (100.0%)
BI-RADS 1 classification (%)	91 (45.5%)	70 (35.0%)	127 (63.5%)
BI-RADS 2 classification (%)	79 (39.5%)	56 (28.0%)	68 (34.0%)
BI-RADS 3 classification (%)	27 (13.5%)	53 (26.5%)	5 (2.5%)
BI-RADS 4 classification (%)	3 (1.5%)	21 (10.5%)	0 (0.0%)
Correct classification (%)	117 (58.5%)	84 (42.0%)	
Classification overestimated (%)	71 (35.5%)	112 (56.0%)	
Classification underestimated (%)	12 (6.0%)	4 (2.0%)	
Weighted kappa	0.367	0.232	

larger studies show slightly higher numbers of dense breasts in their populations, presumably owing to the larger size of the populations used when compared to our study population [18]. For our study, we opted not to include additional women with dense breasts to prevent inclusion bias.

Our study has several limitations. First, a true golden standard for the assessment of breast density is lacking. There is no accurate way to determine breast density other than histopathologic analysis of mastectomy specimens. It is obvious that these specimens are not available for this study, and previously published studies assessing breast density with various computerised methods also lack this (true) golden standard. As we have shown in this study, our software programme acquired similar results to other, more validated programmes. This is why we have chosen to use our software programme as the reference standard (not golden standard) to which we compared the visual assessment of breast density.

A second limitation of our study is that the association between mammographically dense breasts and risk of developing breast cancer has not yet been demonstrated with our software programme. Due to the comparability of acquired results with more validated software programs, we expect that breast density (as assessed by our software program) is also associated with increased breast cancer risk. However, larger studies using our software programme are in progress to prove this hypothesis.

Finally, our analysis remains a semi-automated (and not fully automated) technique, requiring input from an operator and is therefore at risk of introducing observer-dependent bias. Recently, a commercially available software tool (Quantra, Hologic, Bedford, MA, USA) to automatically assess breast density was compared to the Cumulus software program. This study showed a strong density correlation between both breasts and for both methods, suggesting that fully automated assessment of breast density could also aid in breast cancer risk estimation [19]. This was confirmed in another study by Pinker et al., demonstrating that breast density, as assessed with this

automated analysis, was strongly associated with breast cancer risk in women younger than 50 years, but not older than 50 years [20]. In line with these developments, another fully automated software program was launched earlier this year: Volpara (Matakina International, Wellington, New Zealand). However, the major drawback of these recently available programs is their limited availability on only GE or Hologic digital mammography systems.

In summary, our results showed that there is a disagreement in quantitative BI-RADS breast density classification between experienced and inexperienced readers and the semi-automated software analysis. In order to accurately assess breast density in a reproducible and observer-independent manner, we would recommend the use of an integrated software tool, which can be applied in both screening and clinical settings. This semi-automated analysis of breast density might aid in identifying patients at high risk for developing breast cancer and/or patients who can benefit from additional MR mammography because of an unreliable mammogram. Currently, the thresholding approach to assess breast density on standard mammograms is preferable, for instance as assessed by the dedicated software program used in this study or any other commercially available software package. In order to rapidly assess breast density and to include patients who have undergone mastectomy, assessment of a single breast in one projection (preferably the CC projection) is sufficient [15]. This proposal is also supported by our current findings. In turn, future studies could investigate whether patients with mammographically dense breasts (i.e. those at high risk of developing breast cancer) can benefit from this accurate breast density assessment, for instance by shortening the screening interval or by performing additional (MR) imaging.

Conclusion

The results of our study showed that visual assessment of breast density on mammograms is inaccurate and observer-dependent. Semi-automated analysis of breast density

reduces inter-observer variation of breast density classification on mammograms.

References

1. Yaffe MJ (2008) Measurement of mammographic density. *Breast Canc Res* 10:209–218
2. Wolfe JN (1976) Breast patterns as an index of risk for developing breast cancer. *Am J Roentgenol* 126:1130–1139
3. McCormack VA, Dos Santos SI (2006) Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomark Prev* 15:1159–1169
4. Boyd NF, Guo H, Martin LJ, Sun L, Stone J et al (2007) Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 356:227–236
5. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R et al (2003) Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 138:168–175
6. Buist DS, Porter PL, Lehman C, Taplin SH, White E (2004) Factors contributing to mammography failure in women aged 40–49 years. *J Natl Cancer Inst* 96:1432–1440
7. Gram IT, Funkhouser E, Tabar L (1997) The Tabar classification of mammographic parenchymal patterns. *Eur J Radiol* 24:131–136
8. D'Orsi CJ, Mendelson EB, Ikeda DM et al (2003) Breast Imaging Reporting and Data System (BI-RADS) Atlas, 4th ed. American College of Radiology, Reston, VA
9. Heine JJ, Carston MJ, Scott CG, Brandt KR, Wu F et al (2008) An automated approach for estimation of breast density. *Cancer Epidemiol Biomark Prev* 17:3090–3097
10. Byng JW, Yaffe MJ, Jong RA, Shumak RS, Lockwood GA et al (1998) Analysis of mammographic breast density and breast cancer risk from digitized mammograms. *Radiographics* 18:1587–1598
11. Hovhannisyan G, Chow L, Schlosser A, Yaffe MJ, Boyd NF et al (2009) Differences in measured mammographic density in the menstrual cycle. *Cancer Epidemiol Biomark Prev* 18:1993–1999
12. Aitken Z, McCormack VA, Highnam RP, Martin L, Gunasekara A et al (2010) Screen-film mammographic density and breast cancer risk: a comparison of the volumetric standard mammogram form and the interactive threshold measurements methods. *Cancer Epidemiol Biomark Prev* 19:418–428
13. Nicholson BT, LoRusso AP, Smolkin M, Bovbjerg VE, Petroni GR et al (2006) Accuracy of assigned BI-RADS breast density category definitions. *Acad Radiol* 13:1143–1149
14. Boetes C (2010) Update on screening breast MRI in high-risk women. *Magn Reson Imaging Clin N Am* 18:241–247
15. Stone J, Ding J, Warren RML, Duffy SW (2010) Predicting breast cancer risk using mammographic density measurements from both mammogram sides and views. *Breast Cancer Res Treat* 24:551–554
16. Kerlikowske K, Cook AJ, Buist DSM, Cummings SR, Vachon C et al (2010) Breast cancer risk by breast density, menopause, and postmenopausal hormone therapy use. *J Clin Oncol* 28:3830–3837
17. Martin KE, Helvie MA, Zhou C, Roubidoux MA, Bailey JE et al (2006) Mammographic density measured with quantitative computer-aided method: comparison with radiologists' estimates and BI-RADS categories. *Radiology* 240:656–665
18. Pisano ED, Hendrick RE, Yaffe MJ, Baum JK, Acharyya S et al (2008) Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology* 246:376–383
19. Kontos D, Bakic PR, Acciavatti RJ, Conant EF, Maidment ADA (2010) A comparative study of volumetric and area-based breast density estimation in digital mammography: results from a screening population. *LNCS* 6136:378–385
20. Pinker K, Perry N, Milner S, Mokbel K, Duffy S (2010) Validation of a new automated volumetric breast density measurement system as a marker of breast density risk. *Breast Canc Res* 12(Suppl 3):O1