

Cite this article as:

Tagliafico A, Bignotti B, Tagliafico G, Tosto S, Signori A, Calabrese M. Quantitative evaluation of background parenchymal enhancement (BPE) on breast MRI. A feasibility study with a semi-automatic and automatic software compared to observer-based scores. *Br J Radiol* 2015; **88**: 20150417.

## FULL PAPER

# Quantitative evaluation of background parenchymal enhancement (BPE) on breast MRI. A feasibility study with a semi-automatic and automatic software compared to observer-based scores

<sup>1</sup>ALBERTO TAGLIAFICO, MD, <sup>2</sup>BIANCA BIGNOTTI, MD, <sup>3</sup>GIULIO TAGLIAFICO, PhD, <sup>4</sup>SIMONA TOSTO, MD, <sup>2</sup>ALESSIO SIGNORI, PhD and <sup>4</sup>MASSIMO CALABRESE, MD

<sup>1</sup>Institute of Anatomy, Department of Experimental Medicine, University of Genoa, Genoa, Italy

<sup>2</sup>Department of Health Sciences (DISSAL), University of Genoa, Genoa, Italy

<sup>3</sup>Azienda Sanitaria Locale Genova, Corso Scassi, Genova, Italy

<sup>4</sup>Department of Diagnostic Senology, Ist Istituto Nazionale per la Ricerca sul Cancro, IRCCS Azienda Ospedaliera Universitaria San Martino, Genoa, Italy

Address correspondence to: Dr Alberto Tagliafico

E-mail: [alberto.tagliafico@unige.it](mailto:alberto.tagliafico@unige.it)

**Objective:** To evaluate quantitative measurements of background parenchymal enhancement (BPE) on breast MRI and compare them with observer-based scores.

**Methods:** BPE of 48 patients (mean age: 48 years; age range: 36–66 years) referred to 3.0-T breast MRI between 2012 and 2014 was evaluated independently and blindly to each other by two radiologists. BPE was estimated qualitatively with the standard Breast Imaging Reporting and Data System (BI-RADS) scale and quantitatively with a semi-automatic and an automatic software interface. To assess intrareader agreement, MRIs were re-read after a 4-month interval by the same two readers. The Pearson correlation coefficient ( $r$ ) and the Bland-Altman method were used to compare the methods used to estimate BPE.  $p$ -value  $<0.05$  was considered significant.

**Results:** The mean value of BPE with the semi-automatic software evaluated by each reader was 14% (range: 2–79%) for Reader 1 and 16% (range: 1–61%) for Reader 2 ( $p > 0.05$ ). Mean values of BPE percentages for the automatic software were  $17.5 \pm 13.1$  ( $p > 0.05$  vs semi-automatic).

The automatic software was unable to produce BPE values for 2 of 48 (4%) patients. With BI-RADS, interreader and intrareader values were  $\kappa = 0.70$  [95% confidence interval (CI) 0.49–0.91] and  $\kappa = 0.69$  (95% CI 0.46–0.93), respectively. With semi-automated software, interreader and intrareader values were  $\kappa = 0.81$  (95% CI 0.59–0.99) and  $\kappa = 0.85$  (95% CI 0.43–0.99), respectively. BI-RADS scores correlated with the automatic ( $r = 0.55$ ,  $p < 0.001$ ) and semi-automatic scores ( $r = 0.60$ ,  $p < 0.001$ ). Automatic scores correlated with the semi-automatic scores ( $r = 0.77$ ,  $p < 0.001$ ). The mean percentage difference between automatic and semi-automatic scores was 3.5% (95% CI 1.5–5.2).

**Conclusion:** BPE quantitative evaluation is feasible with both semi-automatic and automatic software and correlates with radiologists' estimation.

**Advances in knowledge:** Computerized BPE quantitative evaluation is feasible with both semi-automatic and automatic software. Computerized BPE quantitative scores correlate with radiologists' estimation.

## INTRODUCTION

Background parenchymal enhancement (BPE) is the term used to describe the enhancement of the normal breast tissue. BPE is related to the volume and intensity of enhancement after intravenous contrast agent administration of the normal fibroglandular tissue.<sup>1,2</sup> The normal BPE of breast tissue is categorized in the Breast Imaging Reporting and Data System (BI-RADS) system as minimal, mild, moderate or marked.<sup>1,2</sup> BPE is thought to represent blood flow in the breast tissue and may represent breast activity.<sup>1,3</sup>

As for mammographic breast density, the “extent” of BPE varies among different females, and it is linked to the risk of developing breast cancer.<sup>3–5</sup> It has been suggested that the risk of cancer increases steadily with increasing BPE.<sup>3–5</sup> Recent studies suggest that increased BPE may also affect both reading breast MRI and risk of cancer.<sup>4,5</sup> Recently, it was found that greater BPE was associated with a higher probability of developing breast cancer in females at high risk for cancer.<sup>5</sup> Positive correlations between follow-up recommendations and BPE and between biopsy rates and BPE

have been shown.<sup>6,7</sup> Also, the degree of BPE affected the detection and staging of breast cancer using MRI.<sup>8</sup> A recent retrospective study<sup>9</sup> stated that parenchymal enhancement in the contralateral breast of patients with invasive unilateral breast cancer is significantly associated with long-term outcome, particularly in patients with certain subtypes at immunohistochemistry. Normally, BPE can be classified visually on the basis of the percentage of glandular tissue demonstrating enhancement on fat-suppressed or subtracted breast MRI. According to literature and the need for a more personalized approach to medicine, a more strict quantitative assessment, even with a threshold for defining the level of signal increase for BPE, is needed.<sup>10,11</sup> Indeed, radiologists' agreement for BPE qualitative evaluation is fair, and it may require training. Radiologists' assessment of BPE requires better standardization and reproducibility.<sup>12</sup> Therefore, the purpose of our study was to evaluate quantitative measurements of BPE on breast MRI using a semi-automatic and an automatic software and compare them with standard BI-RADS scores.

## METHODS AND MATERIALS

Partial support for this study was given by one company—Hologic (Bedford, MA) and by the University of Genova. The company provided two dedicated 5-megapixels workstations for data analysis and image evaluation. The University of Genova provided financial and technical support for the purpose of the study. The authors had full access to all data in the study and the information submitted for publication.

### Study population

Consecutive patients referred to breast MRI between December 2012 and December 2014 following EUSOMA (European Society of Breast Cancer Specialists) guidelines and enrolled in prospective ethically approved trials were included in this study.<sup>13</sup> All patients provided written informed consent for review of their medical records and diagnostic images. All patients were examined in the same period of the menstrual cycle.

### Inclusion criteria were:

- Negative results after MRI because the presence of breast malignancy on the MR images could have artificially increased BPE assessments.
- Written informed consent for review of their medical records and diagnostic images.

### Exclusion criteria were:

- Breast implants or prostheses were excluded because the automatic software was not able to assess BPE in these cases.
- Failure to obtain written informed consent.

## MRI PROTOCOL

MRI examinations were performed on a General Electric Signa HDx 3.0-T scanner (General Electric Medical Systems, Milwaukee, WI) with a dedicated eight-channel bilateral breast coil. The patient was placed in a prone position, without any compression of the breasts. MRI scan protocol included the following standard sequences:  $T_1$  turbo spin echo,  $T_2$  turbo spin echo, pre- and post-contrast agent (gadobenate dimeglumine, MultiHance® 0.5 M; Bracco Imaging, Milan, Italy) VIBRANT (Volume Imaging for Breast Assessment) as previously performed on a 3.0-T system.<sup>14,15</sup> Contrast material was power injected ( $0.1 \text{ mmol kg}^{-1}$  of body weight at a rate of  $2 \text{ ml s}^{-1}$ ) and followed by a 20-ml saline flush. All breast MRI findings were reported according to the level of suspicion of malignancy by using the American College of Radiology BI-RADS lexicon.<sup>16</sup> If MRI found a patient with MRI findings classified as BI-RADS 3, 4 or 5, lesion biopsy and subsequent treatment in case of confirmed breast cancer were possible. The complete MRI acquisition parameters are reported in Table 1. After the MRI examination, the non-enhanced images were subtracted from the first-acquired contrast material-enhanced images on a pixel-by-pixel basis using the commercially available software installed on the MRI scanner.

Time–intensity curves were created after contrast media injection, and then, after confirmation that the enhancement was higher in the early post-contrast fat-suppressed  $T_1$  weighted phase, the subtracted images obtained from the pre-contrast and from the early post-contrast fat-suppressed  $T_1$  weighted phases were used to estimate BPE. All images of the early post-contrast sequences were reviewed by two radiologists with 3 and 8 years' experience in breast MRI to assess the BPE independently and blindly to each other. The two radiologists (BB and AT) evaluated the BPE using the standard BI-RADS qualitative scale and a semi-automated software.<sup>1</sup> Then, on the same images, BPE was evaluated using an automated software. The volume and intensity of enhancement were considered in the global assessment of BPE and categorized, on the basis of proposed BI-RADS criteria, as minimal, mild, moderate or marked.<sup>1,4</sup> If the patient had a history of mastectomy, the remaining breast was used for assessment. In females who had both breasts, the level of MRI BPE was recorded for each breast. All breast MR images were re-read after a 4-month interval by the same two readers blindly and independently to the data collected previously to assess intrareader agreement. Images were randomly ordered to the readers to reduce biases.

The mean values of BPE obtained by the qualitative assessment were used to set the automatic version of the software as

Table 1. Sequence parameter settings for axial acquisitions

Sequence	TR (ms)	TE (ms)	Flip angle (°)	Matrix (mm)	Slice thickness (mm)	Spacing (mm)	Approximate scan time (s)
$T_2$	5200	103	90	350 × 350	4.0	3.0	200
$T_1$	600	9	90	350 × 350	4.0	3.0	300
VIBRANT	6.2	3	10	350 × 350	1.2	1.2	90

TE, echo time; TR, repetition time; VIBRANT, volume imaging for breast assessment.

suggested in the literature for breast density evaluation.<sup>17</sup> To avoid the bias of evaluating BPE with the semi-automated software immediately after the qualitative BI-RADS evaluation, images were randomized to reading.

#### Background parenchymal enhancement analysis with semi-automated software

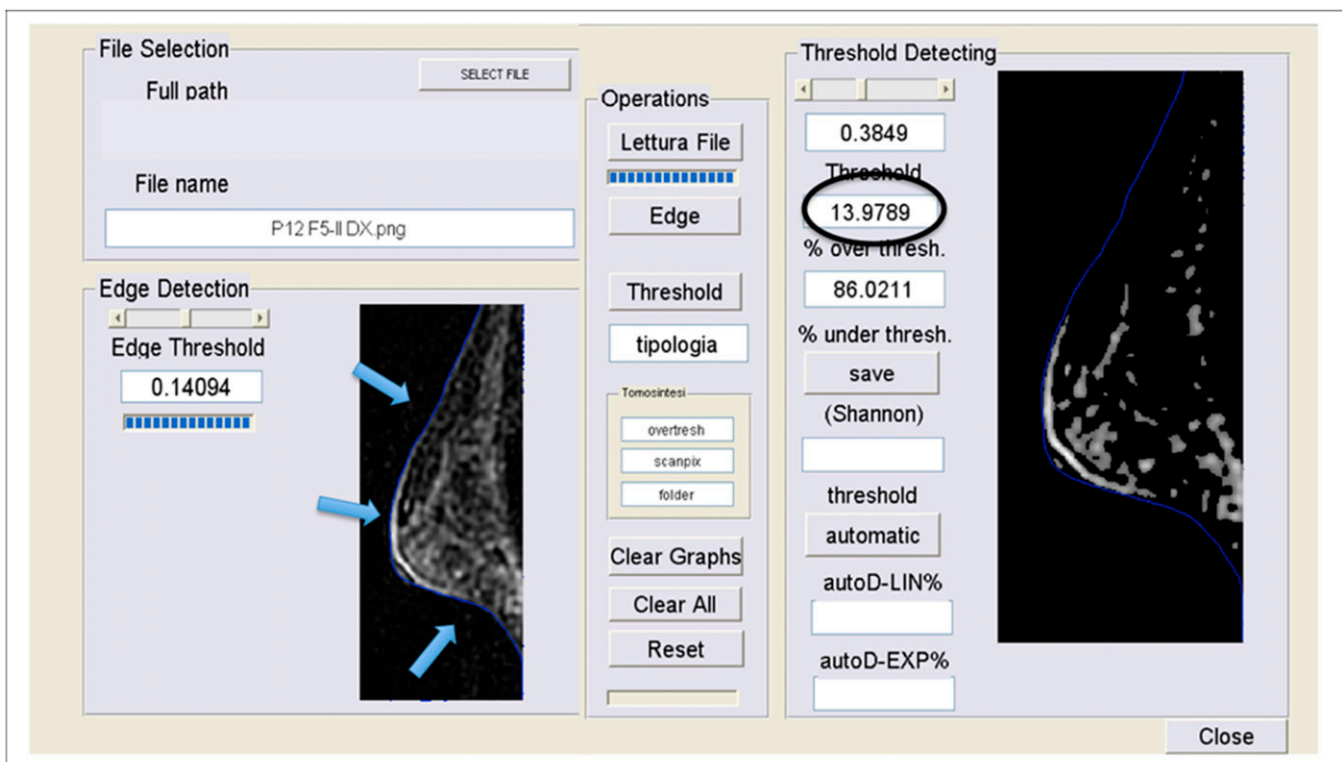
BPE was assessed using a semi-automated software (MedDensity<sup>®</sup>, developed by Giulio Tagliafico), which is a home-grown software previously validated for mammography, tomosynthesis and MRI.<sup>14,17–20</sup> For the purpose of the present study, the software interface was slightly adapted to assess BPE on MRI. The semi-automated version of the software tool (Figure 1) permits to adjust the image by manual threshold adjusting after breast border identifications. The software recognizes the breast boundaries on MRI images automatically by using an edge-tracking algorithm. All slices of both breasts of the early post-contrast sequences were evaluated. After the manual creation of a threshold value to separate enhancing from non-enhancing breast parenchyma, voxels were separated in two classes: voxel values above the threshold correspond to enhancing parenchyma and the voxel values below the threshold correspond to non-enhancing parenchyma. BPE was then calculated dividing the number of voxels in the enhancing parenchyma by the total number of voxels in the entire breast. A minimum training with 100 images was performed and considered sufficient.

#### Background parenchymal enhancement analysis with automated software

To assess BPE with a fully automated software, we defined indices that can be related to the reference of BPE evaluation performed by the radiologists. To set the BPE threshold of the automated software, we used the same cases included in the study using both breasts. The software works on a voxel-by-voxel analysis. The fully automated voxel-by-voxel analysis is objective and fully reproducible, but no correction from the user is possible after starting the calculation. The software is MRI manufacture independent. The region of interest of the breast is divided from the background using the edge-detecting algorithm to exclude the noise. The BPE is finally given using an inversion algorithm based on the correlation between the fully automated results and the reference obtained by the radiologists' BPE assessment. This software uses an algorithm based on the maximum entropy method, which has been demonstrated to be reliable for breast density percentage assessment on different imaging modalities.<sup>14,17–20</sup> This method is based on the maximum entropy thresholding developed by Shannon,<sup>14,17</sup> in which the threshold value corresponding to the maximum entropy value is chosen as an index of the image density<sup>14,17</sup> as measured by separating enhancing and non-enhancing voxels.

The time to perform both automatic and semi-automatic evaluation was calculated using a commercially available stopwatch.

Figure 1. Interface of the semi-automatic software. In this figure, the computer interface is demonstrated. Edges (arrows) of the breast are first identified, then the radiologist adjusts the threshold (in the semi-automated method) and finally the percentage of background parenchymal enhancement (BPE) is read (black circle). In the fully automated method, there is no need to adjust the threshold or to identify the breast edges. In this example, the BPE percentage is 13.9%.



## Statistical analysis

Statistical analysis was performed with commercially available software: IBM SPSS® Statistics v. 19 (IBM Corp., New York, NY; formerly SPSS Inc., Chicago, IL) and MedCalc (MedCalc 9 Software bvba, Ostend, Belgium). Errors of the automatic software were recorded and described. Differences between Reader 1 and Reader 2 scores obtained with the semi-automatic software were tested with Mann–Whitney *U* test, assuming a *p*-value <0.05 as significant. Images were read on per-patient basis: the multiple images for the same individual were analysed with software. Then, a mean value of the BPE for each patient was calculated by the single values of each MRI image. This procedure was performed to avoid the bias of underestimating variance. Finally, each patient had a single value of BPE per breast.

Intrareader and Interreader agreement was defined on the basis of Fleiss classification as follows: <0.40, poor; 0.40–0.59, moderate; 0.60–0.75, good; >0.75, excellent;<sup>21–23</sup> kappa ( $\kappa$ ) statistics were used. An agreement of 0.60 was considered the minimum threshold to perform the subsequent evaluations, and this value was selected in agreement with literature.<sup>4,12</sup> For each observation, the mean of the values obtained by the two radiologists was used. For the purpose of the study, the adjunct of a third radiologist was foreseen only if agreement results of qualitative data obtained by the two radiologists were considered different from the previously published results.<sup>4,17–19</sup>

The relationship among radiologists' BI-RADS evaluation, semi-automatic scores and automatic scores was assessed by calculating the Spearman's  $\rho$  rank correlation coefficient (*r*) For the purpose of the study, to correlate qualitative BI-RADS classes with quantitative BPE evaluation, we considered the proposed BI-RADS criteria (minimal, mild, moderate or marked) as following: minimal, <25% of BPE; mild, 25% ≤ BPE < 50%; moderate, 50% < BPE ≤ 75%; or marked, >75%.

Agreement in the measurements of BPE scores with the semi-automated and with the automated approach was assessed using the Bland–Altman method. The mean difference between each measurement was derived by subtracting the semi-automatic value from the automatic value and then calculating the mean value of this difference for all cases. The mean percentage difference was derived by expressing the difference for each case as a percentage of the mean of the two measurements and then calculating the mean percentage for all cases.

## RESULTS

A complete set of MRI images from 48 females (mean age: 48 years; age range: 36–66 years) with no breast implants and negative results after MRI were available for analysis. Breast MRI

was performed in the first 15 days of the menstrual cycle in 28 premenopausal females, as suggested by current guidelines. 20 females were post-menopausal. 5 patients had a history of mastectomy so that the remaining breast was used for assessment. 6 patients were excluded before reaching the number of 48 patients.

The number of breast MRI examinations assessed as BPE minimal, mild, moderate or marked was:

–Reader 1: minimal, *n* = 38; mild, *n* = 7; moderate, *n* = 2; or marked, *n* = 1.

–Reader 2: minimal, *n* = 33; mild, *n* = 10; moderate, *n* = 4; or marked, *n* = 1.

Using the semi-automated software, the mean value of BPE was 14% (range: 2–79%) for Reader 1 and 16% (range: 1–61%) for Reader 2 (*p*-value .45). Mean value data regarding BPE quantitative analysis using the semi-automatic for both Reader 1 and Reader 2 and the automated software are reported in Table 2.

However, the automatic software was unable to produce BPE values for every image analysed. The automatic version of the software was not able to reliably identify the edges of 2 of 48 (4%) patients. In one patient, the error was located posteriorly, near the thoracic wall, and in another patient, the error was located anteriorly, near the nipple (Figure 2 and Figure 3, respectively). In the first case (error near the thoracic wall), the values of BPE were underestimated by the automatic software owing to the inclusion of background tissue below the threshold for BPE. These data were corrected using the semi-automatic version of the software. In the second case (error near the nipple), the values of BPE estimated by the automatic version of the software were similar to the values estimated by the semi-automatic software owing to the exclusion of normal breast parenchyma which left unchanged the ratio between enhancing and non-enhancing parenchyma. The semi-automatic version of the software, working under the supervision and control of the radiologists, was able to evaluate all images.

## Time

The mean time required for BPE evaluation was 10 ± 2 min for the automatic software and 14 ± 3 min for the semi-automatic software. The time reported is inclusive of all procedures needed to obtain the final data of BPE.

Agreement using the BI-RADS was: interreader,  $\kappa$  = 0.70 [95% confidence interval (CI) 0.49–0.91], and intrareader,  $\kappa$  = 0.69 (95% CI 0.46–0.93).

Agreement using the semi-automated software was: interreader,  $\kappa$  = 0.81 (95% CI 0.59–0.99), and intrareader,  $\kappa$  = 0.85 (95% CI 0.43–0.99).

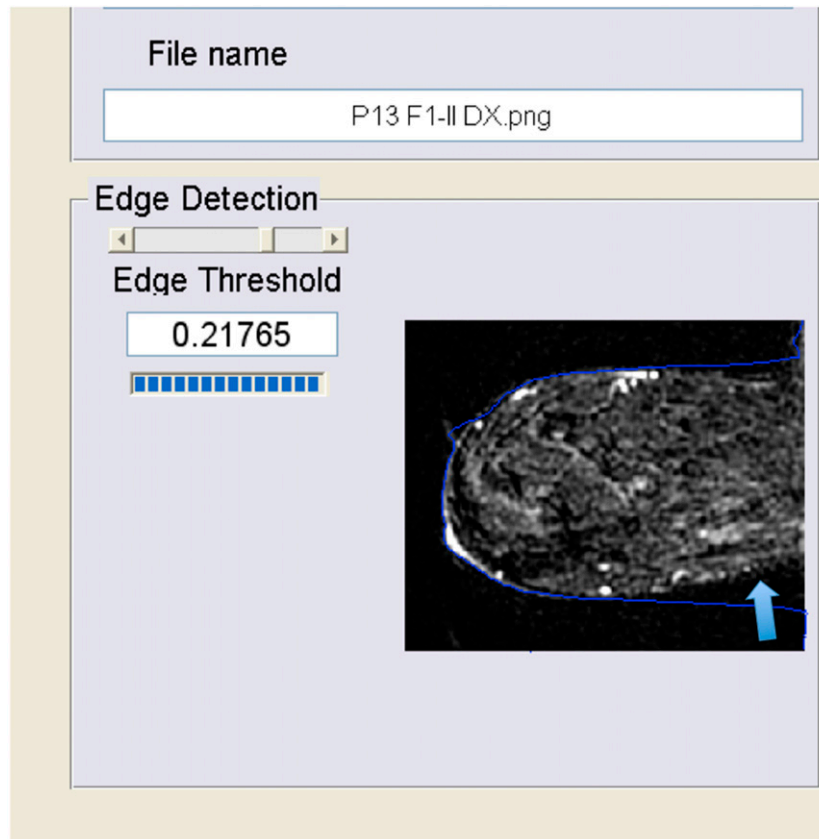
Table 2. Data regarding background parenchymal enhancement (BPE) quantitative analysis using the semi-automatic and the automatic software

Subjects ( <i>n</i> = 48)	Semi-automatic BPE (%)	Automatic BPE (%) <sup>a</sup>	Automatic vs semi-automatic
Images, <i>n</i> = 11,520	15 ± 11.9	17.5 ± 13.1	<i>p</i> > 0.05

Data are expressed in mean percentages ± standard deviation and represent the mean percentage values of BPE of the two readers.

<sup>a</sup>The automatic software registered errors for 2 patients (480 images).

Figure 2. Error in the edge detection near the pectoral region. In this case (error near the thoracic wall, arrow), the values of background parenchymal enhancement were underestimated by the automatic software owing to the inclusion of more background tissues. These data were corrected using the semi-automatic version of the software manually adjusting the edges.



For the fully automated software, there was no need for intra-observer and interobserver agreement evaluation: the software gives the same result for the same image or image set.

Comparison of observer's scores with the BPE given by the automatic software yielded a positive correlation ( $r = 0.56$ ,  $p < 0.0001$ ). Correlations of observer's scores with the semi-automatic software results gave  $r$  values of 0.62 ( $p < 0.0001$ ). Correlations of the automatic software's scores with the semi-automatic software's results gave  $r$  values of 0.75 ( $p < 0.0001$ ).

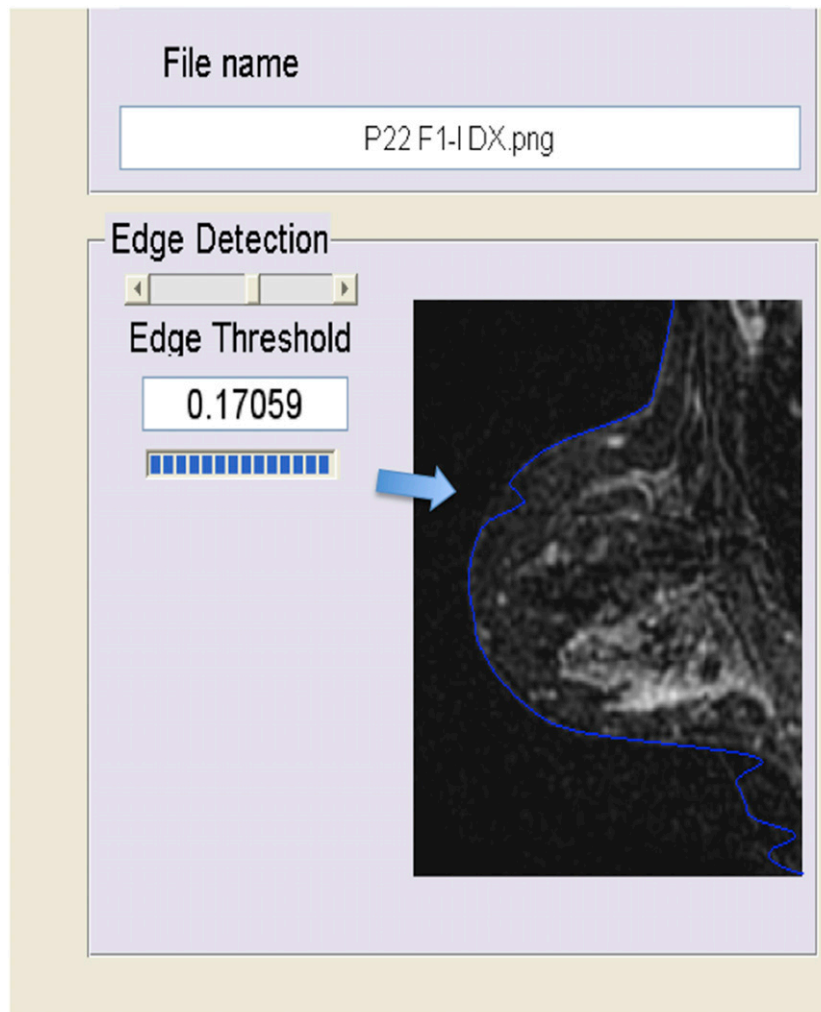
Figure 4 illustrates the comparison of BPE, as measured with the semi-automatic and automatic software, throughout the study population. The mean percentage difference between the values calculated for each case was 3.5% (95% CI 1.53–5.22), which in addition with a Pearson coefficient for correlation of 0.77 suggests a good overlap between the two methods. Figure 5 shows an example of BPE assessment with BI-RADS, semi-automatic and automatic method.

## DISCUSSION

The assessment of BPE is gaining importance owing to the recently described links between BPE and breast cancer risk. In a cohort of high-risk females with no history of breast cancer, BPE has been associated with an increased risk of developing breast cancer nine times higher in females with mild, moderate

or marked BPE than in females with minimal BPE.<sup>5</sup> A link between higher BPE values and breast cancer diagnoses was already noted in the past,<sup>4</sup> but only recently has BPE been associated with long-term outcome, particularly in patients with oestrogen receptor-positive, human epidermal growth factor receptor 2-negative breast cancer.<sup>9</sup> These results suggested that BPE could be an important biomarker of a female's individual breast cancer risk.<sup>5</sup> However, the medical literature lacks simple and reliable method to assess BPE quantitatively. A recent study on 50 BRCA1/2 carriers retrospectively analysed reported to have performed a fully automated quantitative measurement of BPE from breast dynamic contrast-enhanced MRI.<sup>24</sup> Therefore, this study describes a quantitative BPE percentage evaluation made with a semi-automatic and with automatic software and compares computerized score of BPE with radiologists' ones. The need to evaluate BPE with software derived by the moderate to low agreement described for BPE qualitative evaluation.<sup>4,12</sup> Indeed, even after training, the agreement for BPE assessment using a qualitative assessment increased from fair ( $\kappa = 0.36$ ) to moderate ( $\kappa = 0.48$ ).<sup>12</sup> In this study, using the semi-automatic software,  $\kappa$  values were higher than previously reported for both intraobserver and interobserver agreement ( $\kappa > 0.8$ ).<sup>12</sup> Also, when the semi-automatic software was compared with the automatic software, the differences between BPE scores obtained with the two methods could be considered similar and acceptable for both clinical and research purposes.<sup>25</sup> The drawback of

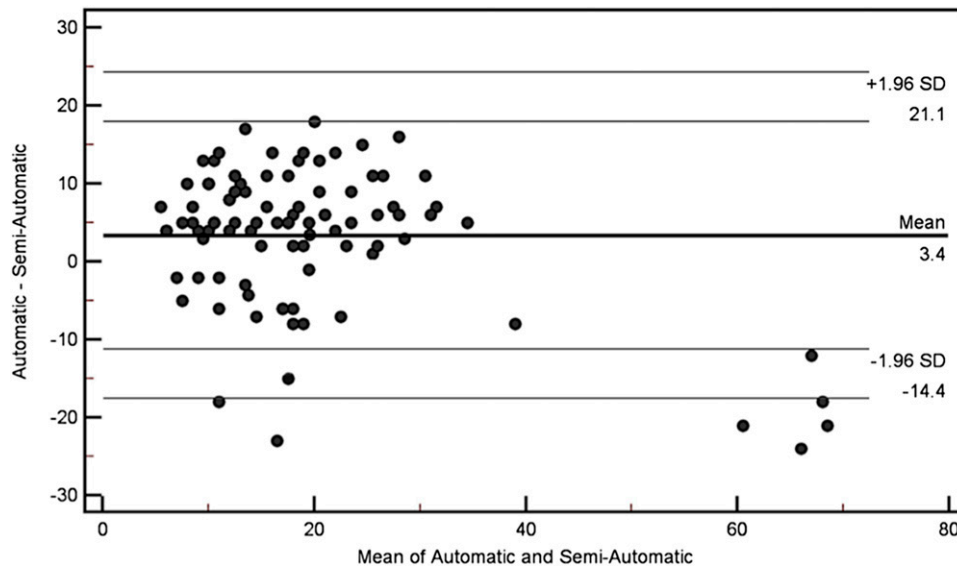
Figure 3. Error in the edge detection near the nipple. In this case (error near the nipple, arrow), the values of background parenchymal enhancement estimated by the automatic version of the software were similar to the values estimated by the semi-automatic software owing to the exclusion of normal breast parenchyma which left unchanged the ratio between enhancing and non-enhancing parenchyma.



the automatic software was the impossibility to evaluate patients with breast implants and errors in 2 of 48 patients near breast edges. These problems could theoretically be solved using the semi-automatic interface, in which the reader can adjust the edges of the tissue to the evaluated edges. Compared with literature, our results of interreader agreement for the qualitative evaluation ( $\kappa = 0.70$ ) were slightly better than those obtained after training<sup>12</sup> and from those reported in the study by King *et al*<sup>4</sup> who reported interreader agreement for BPE level between two readers of 0.47 for ordinal BPE and of 0.57 for collapsed BPE. However, in this study, the presence of breast malignancy on the MR images that could have artificially increased BPE assessments was excluded. The study by King *et al*<sup>4</sup> included patients with pathological MRI findings that could have reduced the agreement between readers. Compared with the study by Melsaether *et al*,<sup>12</sup> our qualitative results could be considered, as a whole, overlapping. In summary, the agreement between the two readers involved in this study confirms that the qualitative evaluation of BPE suffers from intrareader and interreader

variability. This result could be considered a step forward to radiologists' assessment of BPE towards a better standardization. In addition, comparing the semi-automatic software with the fully automated software, correlation values and Bland–Altman evaluations suggest that the latter method could be considered feasible and promising to assess BPE on MRI. On the basis of Bland–Altman plots, it could be hypothesized that, on average, the automatic method gave slightly higher measurements than the semi-automatic method. In addition, for females with particularly high BPE, there was a greater degree of disagreement between the methods with the automatic method giving lower measurements. However, these findings have to be demonstrated to be clinically relevant. If this trend is confirmed on larger series, it may have clinical implications. For example, if BPE is used in personalized medicine to estimate breast cancer risk, the automatic software may underestimate BPE in some patients. We believe that the adjunct of the semi-automatic interface is useful when the automatic version of the software reports errors. The correlation among observer scores and automated scores

Figure 4. Bland-Altman plot of background parenchymal enhancement reported by the semi-automated software subtracted from that reported by the automatic software compared with the mean of the two results. The line of the mean difference and top and bottom lines correspond to 95% limits of agreement [ $\pm 2$  standard deviations (SD)].

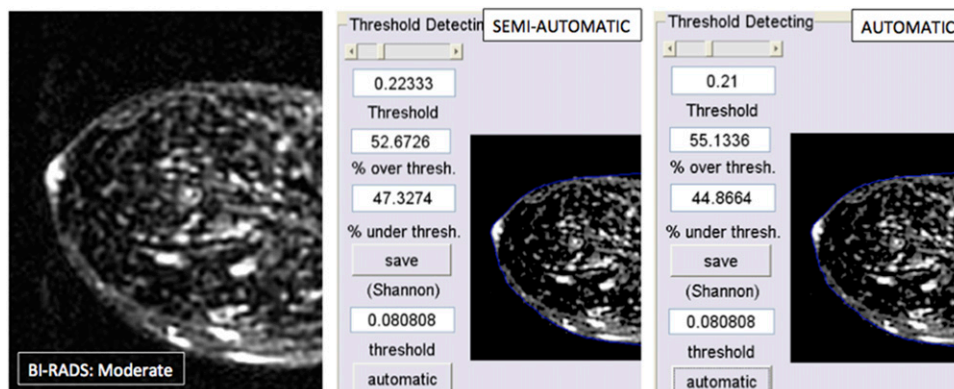


resulted relatively low, probably owing to interobserver variability. Interobserver variability has been claimed to reduce correlation also for breast density, which is a known risk factor for breast cancer.<sup>26</sup> However, we are not aware of a widely approved value of correlation coefficient on what is a sufficient correlation in terms of the clinical importance of agreement for a quantitative assessment of BPE on MRI owing to relatively new research topic we are dealing with. Further research is needed to address this issue. As for breast density, it is possible that the difference between observer-based BPE scores and software-based scores could be the appearance of the displayed image to radiologists who were allowed to adjust the window width and level of the MRI images. This adjustment can substantially alter the image presentation to the reader. Also, in the era of personalized medicine, the use of quantitative imaging biomarker is rapidly expanding. BPE calculated using BI-RADS classes is not a quantitative imaging biomarker, so it is important to assess BPE as a quantitative biomarker as suggested by the European Society

of Radiology and by Radiological Society of North America-Quantitative Imaging Biomarkers Alliance (RSNA-QIBA) Metrology Working Group.<sup>27</sup> For this reason, it is possible that, especially for the value of BPE near the threshold for the different BI-RADS classes (25%; 50%; 75%), some females are classified differently.

This study has limitations. The first is that the evaluation was made only by two readers, but for the aim of the feasibility assessment, the involvement of two readers was discussed in the study protocol and deemed acceptable after comparing the results with the literature. Another limitation is that we used only one contrast media (gadobenate dimeglumine), which is a high relaxivity contrast material; therefore, it is not possible to estimate whether other contrast media with different relaxivity may give different results. It is known that gadobenate dimeglumine increases the enhancement of benign lesions.<sup>28</sup> For this reason, it is possible that BPE scores also could be overestimated by a computer-aided evaluation. Also, our method could

Figure 5. A comparison of Breast Imaging Reporting and Data System (BI-RADS) scale, semi-automatic and automatic evaluations. In this example, the background parenchymal enhancement (BPE) was scored moderate using BI-RADS, and BPE percentage was 52% and 55% using the semi-automatic and the automatic interfaces, respectively.



not perfectly match those already used in literature owing to slight technical differences and achievements.<sup>24,29,30</sup> Previous studies<sup>24,29,30</sup> tried to exclude the fat to assess BPE percentage, considering only the fibroglandular part of the breast, whereas our study did not. Our study considered the whole breast. It is not known if the different methods influence the possibility to use BPE as a biomarker. Also, a completely reliable comparison with the previous studies may not be appropriate, considering the differences in study population, readers and software interfaces.<sup>9,24</sup> Finally, we were not able to find a threshold of percentage BPE to differentiate qualitative BI-RADS categories, and we did not investigate a possible correlation with BPE and breast cancer risk. Also, all patients were examined in the same phase of the menstrual cycle, but even small irregularities in the menstrual cycle can alter BPE: this could be a confounding

factor in this study. Further research is needed to clarify these issues and to assess whether computerized BPE evaluation is influenced by different contrast materials. This research is needed especially if BPE percentages will be considered to assess breast cancer risk.

In conclusion, computer-assisted BPE evaluation was feasible with both semi-automatic and automatic software and correlated with radiologists' estimation. Semi-automatic evaluation permits adjustments when breast boundaries are hard to recognize on automatic software.

## FUNDING

PRA 2013 by University of Genova and AIRC Associazione Italiana Ricerca sul Cancro IG 15697: grants to Alberto Tagliafico.

## REFERENCES

- Morris EA. Diagnostic breast MR imaging: current status and future directions. *Radiol Clin North Am* 2007; **45**: 863–80. doi: [10.1016/j.rcl.2007.07.002](https://doi.org/10.1016/j.rcl.2007.07.002)
- Kuhl C. The current status of breast MR imaging. Part I. Choice of technique, image interpretation, diagnostic accuracy, and transfer to clinical practice. *Radiology* 2007; **244**: 356–78. doi: [10.1148/radiol.2442051620](https://doi.org/10.1148/radiol.2442051620)
- Pike MC, Pearce CL. Mammographic density, MRI background parenchymal enhancement and breast cancer risk. *Ann Oncol* 2013; **24**: viii37–viii41. doi: [10.1093/annonc/mdt310](https://doi.org/10.1093/annonc/mdt310)
- King V, Brooks JD, Bernstein JL, Reiner AS, Pike MC, Morris EA. Background parenchymal enhancement at breast MR imaging and breast cancer risk. *Radiology* 2011; **260**: 50–60. doi: [10.1148/radiol.11102156](https://doi.org/10.1148/radiol.11102156)
- Dontchos BN, Rahbar H, Partridge SC, Korde LA, Lam DL, Scheel JR, et al. Are qualitative assessments of background parenchymal enhancement, amount of fibroglandular tissue on MR images, and mammographic density associated with breast cancer risk? *Radiology* 2015; **276**: 371–80. doi: [10.1148/radiol.2015142304](https://doi.org/10.1148/radiol.2015142304)
- DeMartini WB, Liu F, Peacock S, Eby PR, Gutierrez RL, Lehman CD. Background parenchymal enhancement on breast MRI: impact on diagnostic performance. *AJR Am J Roentgenol* 2012; **198**: W373–80. doi: [10.2214/AJR.10.6272](https://doi.org/10.2214/AJR.10.6272)
- Hambly NM, Liberman L, Dershaw DD, Brennan S, Morris EA. Background parenchymal enhancement on baseline screening breast MRI: impact on biopsy rate and short-interval follow up. *AJR Am J Roentgenol* 2011; **196**: 218–24. doi: [10.2214/AJR.10.4550](https://doi.org/10.2214/AJR.10.4550)
- Uematsu T, Kasami M, Watanabe J. Does the degree of background enhancement in breast MRI affect the detection and staging of breast cancer? *Eur Radiol* 2011; **21**: 2261–7. doi: [10.1007/s00330-011-2175-6](https://doi.org/10.1007/s00330-011-2175-6)
- van der Velden BH, Dmitriev I, Loo CE, Pijnappel RM, Gilhuijs KG. Association between parenchymal enhancement of the contralateral breast in dynamic contrast-enhanced MR imaging and outcome of patients with unilateral invasive breast cancer. *Radiology* 2015; **276**: 675–85. doi: [10.1148/radiol.15142192](https://doi.org/10.1148/radiol.15142192)
- Uematsu T, Kasami M, Watanabe J. Should breast MRI be performed with adjustment for the phase in patients' menstrual cycle? Correlation between mammographic density, age, and background enhancement on breast MRI without adjusting for the phase in patients' menstrual cycle. *Eur J Radiol* 2012; **81**: 1539–42. doi: [10.1016/j.ejrad.2011.04.059](https://doi.org/10.1016/j.ejrad.2011.04.059)
- Morrish OW, Tucker L, Black R, Willsher P, Duffy SW, Gilbert FJ. Mammographic breast density: comparison of methods for quantitative evaluation. *Radiology* 2015; **275**: 356–65. doi: [10.1148/radiol.14141508](https://doi.org/10.1148/radiol.14141508)
- Melsaether A, McDermott M, Gupta D, Pysarenko K, Shaylor SD, Moy L. Inter- and intra-reader agreement for categorization of background parenchymal enhancement at baseline and after training. *AJR Am J Roentgenol* 2014; **203**: 209–15. doi: [10.2214/AJR.13.10952](https://doi.org/10.2214/AJR.13.10952)
- Sardanelli F, Boetes C, Borisch B, Decker T, Federico M, Gilbert FJ, et al. Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. *Eur J Cancer* 2010; **46**: 1296–316. doi: [10.1016/j.ejca.2010.02.015](https://doi.org/10.1016/j.ejca.2010.02.015)
- Tagliafico A, Bignotti B, Tagliafico G, Astengo D, Martino L, Airdi S, et al. Breast density assessment using a 3T MRI system: comparison among different sequences. *PLoS One* 2014; **9**: e99027. doi: [10.1371/journal.pone.0099027](https://doi.org/10.1371/journal.pone.0099027)
- Tagliafico A, Rescinito G, Monetti F, Villa A, Chiesa F, Fiscì E, et al. Diffusion tensor magnetic resonance imaging of the normal breast: reproducibility of DTI-derived fractional anisotropy and apparent diffusion coefficient at 3.0 T. *Radiol Med* 2012; **117**: 992–1003. doi: [10.1007/s11547-012-0831-9](https://doi.org/10.1007/s11547-012-0831-9)
- American College of Radiology BI-RADS Committee. *Breast imaging reporting and data system*. 4th edn. Reston, VA: American College of Radiology; 2003.
- Tagliafico A, Tagliafico G, Tosto S, Chiesa F, Martinoli C, Derchi LE et al. Mammographic density estimation: comparison among BI-RADS categories, a semi-automated software and a fully automated one. *Breast* 2009; **18**: 35–40. doi: [10.1016/j.breast.2008.09.005](https://doi.org/10.1016/j.breast.2008.09.005)
- Tagliafico A, Tagliafico G, Astengo D, Cavagnetto F, Rosasco R, Rescinito G, et al. Mammographic density estimation: one-to-one comparison of digital mammography and digital breast tomosynthesis using fully automated software. *Eur Radiol* 2012; **22**: 1265–70. doi: [10.1007/s00330-012-2380-y](https://doi.org/10.1007/s00330-012-2380-y)
- Tagliafico A, Tagliafico G, Astengo D, Airdi S, Calabrese M, Houssami N. Comparative estimation of percentage breast tissue density for digital mammography, digital breast tomosynthesis, and magnetic resonance imaging. *Breast Cancer Res Treat* 2013; **138**: 311–17. doi: [10.1007/s10549-013-2419-z](https://doi.org/10.1007/s10549-013-2419-z)
- Tagliafico AS, Tagliafico G, Cavagnetto F, Calabrese M, Houssami N. Estimation of percentage breast tissue density: comparison between digital mammography (2D full field digital mammography) and digital breast tomosynthesis according to different



- BI-RADS categories. *Br J Radiol* 2013; **86**: 20130255. doi: [10.1259/bjr.20130255](https://doi.org/10.1259/bjr.20130255)
21. Fleiss, JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971; **76**: 378–82. doi: [10.1037/h0031619](https://doi.org/10.1037/h0031619)
  22. Gwet KL. *Handbook of inter-rater reliability*. 4th edn. Gaithersburg, MD: Advanced Analytics, LLC; 2014.
  23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–74.
  24. Wu S, Weinstein SP, DeLeo MJ 3rd, Conant EF, Chen J, Domchek SM, et al. Quantitative assessment of background parenchymal enhancement in breast MRI predicts response to risk-reducing salpingo-oophorectomy: preliminary evaluation in a cohort of BRCA1/2 mutation carriers. *Breast Cancer Res* 2015; **17**: 67. doi: [10.1186/s13058-015-0577-0](https://doi.org/10.1186/s13058-015-0577-0)
  25. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; **8**: 135–60. doi: [10.1191/096228099673819272](https://doi.org/10.1191/096228099673819272)
  26. Pettersson A, Graff RE, Ursin G, Santos Silva ID, McCormack V, Baglietto L, et al. Mammographic density phenotypes and risk of breast cancer: a meta-analysis. *J Natl Cancer Inst* 2014; **106**: diu078. doi: [10.1093/jnci/dju078](https://doi.org/10.1093/jnci/dju078)
  27. Sullivan DC, Obuchowski NA, Kessler LG, Raunig DL, Gatsonis C, Huang EP, et al; RSNA-QIBA Metrology Working Group. Metrology standards for quantitative imaging biomarkers. *Radiology* 2015; **12**: 142202. Epub ahead of print. doi: [10.1148/radiol.2015142202](https://doi.org/10.1148/radiol.2015142202)
  28. Carbonaro LA, Pediconi F, Verardi N, Trimboli RM, Calabrese M, Sardanelli F. Breast MRI using a high-relaxivity contrast agent: an overview. *AJR Am J Roentgenol* 2011; **196**: 942–55. doi: [10.2214/AJR.10.4974](https://doi.org/10.2214/AJR.10.4974)
  29. Wu S, Weinstein SP, Conant EF, Schnall MD, Kontos D. Automated chest wall line detection for whole-breast segmentation in sagittal breast MR images. *Med Phys* 2013; **40**: 042301. doi: [10.1118/1.4793255](https://doi.org/10.1118/1.4793255)
  30. Wu S, Weinstein SP, Conant EF, Kontos D. Automated fibroglandular tissue segmentation and volumetric density estimation in breast MRI using an atlas-aided fuzzy C-means method. *Med Phys* 2013; **40**: 122302. doi: [10.1118/1.4829496](https://doi.org/10.1118/1.4829496)