New elastographic classification of breast lesions during and after compression

Eduardo de Faria Castro Fleury, Jose Carlos Vendramini Fleury, Sebastiao Piato, Decio Roveda Jr.

Breast ultrasonography has evolved considerably over the last few decades. From the limited diagnostic method that allowed only the discrimination of solid and cystic lesions, it has become a sophisticated tool for evaluating and classifying masses and cysts (1, 2). Ultrasonographic criteria that help differentiate benign from malignant lesions include shape, margin, orientation, echo pattern, posterior acoustic features, lesion boundary, and vascularization, as proposed in the latest edition of Breast Imaging Reporting and Data System (BI-RADS™) (3).

In this context, new research is being conducted on methods that may improve the diagnostic specificity and accuracy of ultrasonography without compromising its sensitivity. In 1991, Ophir et al. (4) introduced a new technique called elastography, which allowed the evaluation of soft tissue by elastic deformation, through the application of compression to the area of interest, bearing in mind that the mechanical properties of the tissue are important indicators for the diagnosis of malignant lesions, given that most breast carcinomas are harder than the adjacent tissue (5).

Currently there are two main lines of research to determine the clinical applicability of ultrasound elastography. The first line of research is based on the assessment of the mass size before and after compressing the target areas. This assessment uses software with which soft lesions appear lighter and firm lesions appear darker, and malignant lesions tend to be more evident than benign lesions (6–8). The other line of research is based on the use of a software that applies a different color spectrum to tissues according to their hardness, ranging from red to soft tissues, green to intermediate tissues and dark blue to hard tissues (9–11).

There is no consensus as to which is the best technique or classification defining its clinical application. The main limitations are the interobserver variability described in previous studies and the lack of a standard for how compression should be performed during the study.

This study proposes an elastographic classification system using not only the image of the lesion before and during compression, but also after parenchymal decompression and its correlation with the respective histological result. The introduction of the time point (decompression) into the study shows that the patient’s breast tissue type has less influence on the results, allowing evaluation criteria to be adopted with scores ranging from 1 to 4, based on the color variation observed in the areas of interest.

Materials and methods

Patients

This is a prospective trial. Images from 207 patients, with a mean age of 44 years (range, 17–83), were evaluated. They presented with 247
lesions during the ultrasonographic study, and were referred for percutaneous breast biopsy during the period from May 1st to October 30th 2007. The mean diameter of the lesions was 1.5 cm (median of 1.3 cm, ranging from 0.5 cm to 3.2 cm). The first 10 patients were excluded from the trial and were regarded as having contributed to the learning curve for the method. The results of the percutaneous biopsy were benign for these patients. Nine additional patients with nine lesions were excluded due to the presence of non-mass lesions with asymmetry and architectural distortions on the ultrasound. All mass lesions according to the BI-RADS™ lexicon were included in the study. Two-hundred and eleven (92.5%) core biopsies and 17 (7.5%) preoperative localizations were carried out after performing and documenting the elastographic study; there was histological confirmation in all cases.

The trial was approved by the institutional review board. Informed consent was obtained from all the participating patients.

**Equipment**

Both the conventional and the elastographic studies were performed by two radiologists with 6 and 17 years of experience in breast imaging, respectively. Examinations were performed using a Sonix SP (Ultrasonix Medical Corporation, Vancouver, Canada) US system and a 5–14 MHz multifrequency linear probe. For the elastography study, a special software was used for the Ultrasonix system, version 3.0.2 (Beta 1) updated to the commercial version 2.6.

**Elastographic study**

The elastographic study was performed with the patient lying in the same position used for conventional US examination (B-mode) and with the transducer positioned perpendicular to the region of interest (ROI). The target lesion was repeatedly compressed before examination to ensure that there was no lateral shift. After the activation of elastography, continuous manual compression was applied to the target region, perpendicular to the pectoral muscle, until tissue resistance was detected. When resistance was felt, manual pressure was interrupted, allowing spontaneous decompression of the breast parenchyma. The study area comprised the region from the subcutaneous tissue to the pectoral muscle and also the margins of the mass up to 0.5 cm.

The elastographic imaging technique used comprises three phases: Step 1: Tissue is imaged with and without light compression, and RF (radiofrequency) data lines are acquired in the ROI. Step 2: A displacement is estimated between every two lines of RF data. Step 3: A strain value for every point in the ROI is estimated based on RF data deformation. Each pixel of the elasticity image was assigned one of 256 specific colors, depending on the magnitude of strain. The scale ranged from red for components with greatest strain (i.e., softest components) to blue for those with no strain (i.e., firmest components). Green indicated average strain in the ROI. The images obtained through elastography were superimposed to the B-mode images.

The images acquired were assessed in real time through “cinememory”, which allows retrospective evaluation of the behavior of the mass during compression and after decompression. Overall, study duration did not exceed two minutes per lesion.

**Elastographic classification**

The elastographic classification used a four-point scale according to the color variation during compression and after decompression of the ROI (Fig. 1). A score of 1 was assigned to lesions presenting the same color spectrum as the peripheral breast tissue (Fig. 2). A score of 2 was assigned to lesions presenting color variation of less than 50% of the lesion area (between 10% and 50%), usually around the margins, varying between yellow and green on the color scale; Score 4, lesions with no significant color variation during compression and decompression.

**Figure 1.** Classification of breast masses using scores and the assessment of color variation during compression and after decompression of the breast tissue. Score 1, lesions that presented the same spectrum of colors as the adjacent breast tissue; Score 2, lesions which, after decompression, had soft tissue color variations covering over 50% of the lesion; Score 3, lesions which, after decompression, presented a color variation on less than 50% of the area (between 10% and 50%), usually around the margins, varying between yellow and green on the color scale; Score 4, lesions with no significant color variation during compression and decompression.
Figure 2. a–c. Example of score 1. Mass in B-mode (a), during compression (b) and after decompression (c). There is no definition of the tissue mass when the elastographic study is associated with B-mode, with histology showing fibrocystic change in the breast of a 26-year-old woman. Biopsy samples were obtained by core biopsy.

Figure 3. a–c. Example of score 2. Mass in B-mode (a), during compression (b) and after decompression (c). The tissue mass presents an intense posterior acoustic shadow, and is poorly characterized in B-mode (a) during compression; the mass is better defined (b) than after decompression (c) and there is a change in color covering more than 50% of its area, which suggests it is benign. The histological study has confirmed that the tissue mass is benign, compatible with a fibroadenoma in involution in a 46-year-old woman. Fragments were obtained by core biopsy.

Figure 4. a–c. Example of score 3. Mass in B-mode (a), during compression (b), and after decompression (c). The ovoid mass limited to B-mode (a). During compression the mass is distinguished from normal tissue, more solid than the tissue (b), and after decompression (c) there is a change in color covering less than 50% of the tissue mass, suggesting that it is benign. The histological study has confirmed that the mass is benign, compatible with fibroadenoma in a 33-year-old woman. Biopsy samples were obtained by core biopsy.
during compression and after decompression of the parenchyma, appearing blue on both images (Fig. 5) (12). The classification system used by the authors is similar to the one proposed by Scaperrotta et al. (11); however, this proposed system relies on the assessment of the images during decompression periods, where the influence of the manner in which the compression is applied in the region of interest is smaller, thus simplifying the study systematization.

**Pathological diagnosis**

All samples obtained were sent for histological study, and were analyzed by a specialized breast pathologist with 17 years of experience. The lesions were divided into two groups: Group 1, benign lesions; and Group 2, malignant lesions. Group 1 was divided into three subgroups: Group 1a with fibrocystic alterations, Group 1b with fibroadenomas, and Group 1c with low malignant potential, including papillomas, radial scars, myoepitheliomas and sclerosing lesions, according to the classification proposed by Ellis et al. (13, 14).

**Statistical analysis**

Sensitivity, specificity, positive predictive value, and negative predictive value by elastography were evaluated in comparison with the histological results of the samples. Scores 1, 2, and 3 were considered negative, and score 4 was considered positive. For comparison purposes, these values were calculated for conventional ultrasound (B-Mode), with the lesions classified as BI-RADSTM categories 1, 2, and 3 considered negative, and 4 and 5 positive. The biopsied lesions were from patients with indication for biopsy referred from other services, and were classified as follows: 118 (51.8%) BI-RADSTM category 3, 104 (45.6%) BI-RADSTM category 4, and 6 (2.6%) BI-RADSTM category 5.

Before biopsies were performed, these lesions were reclassified in our service as follows: three (1.3%) BI-RADSTM category 1, 20 (8.8%) BI-RADSTM category 2, 138 (60.5%) BI-RADSTM category 3, 57 (25%) BI-RADSTM category 4, and 10 (4.4%) BI-RADSTM category 5. The lesions reclassified as BI-RADSTM category 1 were interpreted in our service as areas of fibrocystic changes interspersed in the heterogeneous breast tissue. However, biopsies were performed to confirm these findings.

Fisher’s exact test was used to test the association between the elastogram and the histological result, with

---

**Table 1.** Histological results. Distribution of the histological results according to group division

<table>
<thead>
<tr>
<th>Results</th>
<th>Group</th>
<th>Lesions</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fibrocystic changes</td>
<td>1a</td>
<td>65</td>
<td>28.5</td>
</tr>
<tr>
<td>Fibroadenoma</td>
<td>1b</td>
<td>112</td>
<td>49.1</td>
</tr>
<tr>
<td>Low malignant potential</td>
<td>1c</td>
<td>21</td>
<td>9.2</td>
</tr>
<tr>
<td>Malignancy</td>
<td>2</td>
<td>30</td>
<td>13.2</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>228</td>
<td>100.00</td>
</tr>
</tbody>
</table>

**Table 2.** Mean, median, and standard deviation of the elastography scores according to histological groups

<table>
<thead>
<tr>
<th>Histology</th>
<th>Group</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fibrocystic changes</td>
<td>1a</td>
<td>2.1</td>
<td>2.0</td>
<td>0.61</td>
</tr>
<tr>
<td>Fibroadenoma</td>
<td>1b</td>
<td>2.5</td>
<td>2.0</td>
<td>0.57</td>
</tr>
<tr>
<td>Low malignant potential</td>
<td>1c</td>
<td>2.8</td>
<td>3.0</td>
<td>0.70</td>
</tr>
<tr>
<td>Malignant</td>
<td>2</td>
<td>3.9</td>
<td>4.0</td>
<td>0.35</td>
</tr>
</tbody>
</table>
significance at \( P < 0.05 \), using the commercial software SPSS 12.0 (SPSS Inc., Chicago, USA).

The accuracy of the method was also determined using the parametric estimate of the area under the receiver operating characteristic (ROC) curve, using the commercial Stata 8.0 software (StataCorp, College Park, Texas, USA).

In order to assess the agreement between observers, a kappa test according to the criteria described by Landis and Koch was used (15). All significance probabilities (\( P \) values) shown were two-sided. Values <0.05 were considered statistically significant. The software SAS 9.1.3 (Statistical Analysis System, Cary, North Carolina, USA) was used for the calculations. For this purpose, all lesions were assessed again and reclassified by each of the observers using "cinememory".

### Results

**Pathological diagnosis**

Of the 228 lesions evaluated, 65 (28.5%) were included in Group 1a; 112 (49.1%) in Group 1b; 21 (9.2%) in Group 1c; and 30 (13.2%) in Group 2 (Table 1). Of the 30 malignant results (Group 2), 19 (63.4%) were invasive ductal carcinomas; nine (30.0%) were invasive lobular carcinomas; one (3.3%) was a papillary carcinoma; and one (3.3%) was a carcinoid tumor.

Of the 21 lesions classified in Group 1c, 13 (61.9%) were obtained during surgical excisional biopsy. Of the remaining eight lesions, five (23.8%) underwent excisional biopsy after the diagnosis from the percutaneous biopsy, and were considered benign. Imaging follow-up for one year was performed for the other three lesions (14.3%). Malignancy was not observed in any of these cases.

**Elastography scores**

The mean and median scores for the histological classification of the lesions were 2.1 and 2.0, respectively, for Group 1a; 2.5 and 2.0 for Group 1b; 2.8 and 3.0 for Group 1c; and 3.9 and 4.0 for Group 2 (Table 2).

Table 3 shows the frequency of the histological groups, according to the electrographic scores (Table 3).

The four (1.7%) false-negative results obtained by the elastogram were classified as score 3, with the following findings: two lobular carcinomas with diameters of 0.9 cm and 0.8 cm, respectively, one papillary carcinoma of 2.3 cm, and one carcinoid tumor of 1.2 cm.

The positive and negative predictive values, sensitivity, specificity, and diagnostic accuracy of the elastographic scores were 76.47%, 97.94%, 86.67%, 95.96%, and 94.74%, respectively (Table 4). When the scores were analyzed separately, we obtained an NPV of 100% for score 1, an NPV of 100% for score 2, an NPV of 94.52% for score 3, and a PPV of 76.47% for score 4. For the conventional study, we obtained 90% sensitivity, 79.80% specificity, and 81.14% diagnostic accuracy (Table 5).

For the elastogram, the ROC curves, revealed an area under the curve of 0.954, a confidence interval between 0.925 and 0.982, and error of 0.0146 (Fig. 6). For the conventional study, ROC curves revealed an area under the curve of 0.888, a confidence interval between 0.830 and 0.946, and error of 0.297 (Fig. 6).

The optimal cut-off point identified for the proposed classification was score 3, which is the cut-off point corresponding to the maximum value of the Youden index accounting for sensitivity, specificity, and diagnostic accuracy of 88%, 98.2%, and 96.5%, respectively, for observer 1, and 86.2%, 95.5%, and 94.2%, respectively, for observer 2. No interobserver statistically significant difference was observed at a significance level of 5%.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Frequency and percentage of histological groups according to elastography scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histology</td>
<td>Group</td>
</tr>
<tr>
<td>Fibrocystic changes</td>
<td>1a</td>
</tr>
<tr>
<td>Fibroadenomas</td>
<td>1b</td>
</tr>
<tr>
<td>Low malignant potential</td>
<td>1c</td>
</tr>
<tr>
<td>Malignant</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
<th>True-positive results (TP), true-negative results (TN), false-positive results (FP), and false-negative results (FN) for the elastography scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>Score 1</td>
</tr>
<tr>
<td>TP</td>
<td>0</td>
</tr>
<tr>
<td>TN</td>
<td>8</td>
</tr>
<tr>
<td>FP</td>
<td>0</td>
</tr>
<tr>
<td>FN</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5</th>
<th>True-positive results (TP), true-negative results (TN), false-positive results (FP), and false-negative results (FN) for the B-mode study</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI-RADS 1</td>
<td>BI-RADS 2</td>
</tr>
<tr>
<td>TP</td>
<td>0</td>
</tr>
<tr>
<td>TN</td>
<td>3</td>
</tr>
<tr>
<td>FP</td>
<td>0</td>
</tr>
<tr>
<td>FN</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
</tr>
</tbody>
</table>
Discussion

Over the last decade, elastography has become an important tool for the study of soft tissues, with the clinical perspective of detecting lesions and determining pathological tissue alterations, enabling the adequate treatment of lesions (17–20). The information acquired is similar to that obtained through manual palpation; however, the data from elastography studies is more sensitive and less subjective (21–25).

Studies conducted to evaluate lesions according to the size, using software with which soft lesions appear lighter, and hard lesions appear darker, and malignant lesions tend to be more evident than benign lesions, showed good diagnostic accuracy; however, the main limitation of the method was the interobserver variability. Moon et al. (8) carried out a study of 100 breast masses by continual elastography exam using a computer assisted diagnostic program through segmented images of the masses, evaluating the margins, changes in anteroposterior distance, and differences between the total area and hardness. These authors obtained a sensitivity of 85%, specificity of 88%, diagnostic accuracy of 87%, and negative predictive value of 90%. Based on a similar concept, but also considering the interobserver variability, Regner et al. (7) concluded that elastography has good potential for differentiating breast tissue lesions, but that one of its main limitations was low specificity for some readers (24%). Burnside et al. (6), presented similar results, and concluded that interobserver variability and image quality interfere with the observer performance.

In the present study, we did not consider tumor size measurements using elastography to be an adequate criterion for lesion analysis, mainly because during the process of the elastographic study, the definition of the margins of the lesion is poor, and this may influence the final result of the analysis. This may be one of the factors that contributed to the inter-observer variation reported in these series.

The other research line based on the use of software that applies a different color spectrum to tissues according to their hardness showed good diagnostic accuracy, although there were no studies evaluating interobserver variability. Itoh et al. (9) proposed a classification for the lesions according to the color spectrum variation obtained in a dynamic form. A total of 111 lesions were assessed using a five-point classification system, where scores of 1, 2, and 3 were considered benign, and scores 4 and 5 were considered malignant. These authors obtained a sensitivity of 86.5%, specificity of 89.8% and diagnostic accuracy of 88.3%. Scaperrotta et al. (11), using a simplified three-point classification system based on the classification by Itoh et al. showed a sensitivity of 80% and a specificity of 80.9%, which was similar to findings of the ultrasound study. They concluded that elastography may be a useful aid to US for less experienced radiologists in the assessment of solid non-palpable breast lesions, especially BI-RADS® 3, for which specificity was higher (88.7%). The main limitation reported by the authors is that elastography is operator-dependent, and there may be an interobserver variability; however, all studies concluded that sonoelastography requires training and practice to learn the appropriate technique.

There are no reports on elastography studies using the images acquired after decompression, but only pre- and post-compression images in the region of interest. We believe that studies carried out during compression and after decompression, with compressions performed according to the proposed systematization, can not only standardize the study, but also can provide parameters for comparison between these two time points. We consider that the difficulty in quantifying the strength to be applied for parenchymal compression can be overcome by using gradual compression, until resistance is felt in the breast on which force was not applied, and taking into account the moment of spontaneous decompression.

According to Hooke’s law (26), when elastic deformity is created in a material (strain) by an external force (stress), the accumulated energy (elastic potential energy) allows the material to restore its original shape after decompression. Because it is cumulative, the elastic potential energy, which is the force used during the decompression period, does not depend on the manner (acceleration) with which compression is performed. This could minimize the main limitation of the method, which is the lack of a standard for the way in which the compression force should be applied.

The scoring system used in this study is similar to those used by Scaperrotta et al. (11), although in this study, images in the compression and decompression...
periods were compared. Scores 1 and 2 were considered benign, score 3 probably benign, and 4 suggestive of malignancy. We believe that the simpler the classification, the easier its applicability and reproducibility.

Analyzing the results obtained in this study, it is clear that there is strong statistical evidence of an association between the histological diagnosis and the scores proposed by the authors for elastography, whereby scores 1, 2, and 3 were considered negative for malignancy, and score 4 was considered suggestive of malignancy ($P < 0.001$). The ROC curve showed an area under the curve of 0.9539, demonstrating the excellent diagnostic accuracy of the method. Comparing this with the area obtained in other studies due to the exclusion of non-mass lesions with asymmetry and distortions. Such lesions are composed of healthy tissue interspersed within pathologic tissue, which may lead to false-negative results.

All of the false-negative results had been classified as score 3. Dimensions of the lesions that were misclassified varied in a great range (between 0.8 and 2.3 cm) which showed the histological type of the lesions influenced the results more than did their dimensions. In the previous reports the tumors initially classified as benign, particularly the carcinoid tumor and the papillary carcinoma, are lesions that are generally softer on manual palpation (27, 28). Our results differ from the ones reported in the literature, which showed that elastography has better accuracy for lesions smaller than 2.0 cm.

Our study demonstrated that fibroadenomas larger than 2.0 cm were classified with malignant scores, and conversely, that small malignant tumors tended to present benign scores. Figure 7 shows a ductal carcinoma of approximately 4.2 cm classified as score 4 by elastography, showing a better correlation with histological type than with size for the diagnosis of malignancy. Further studies are needed to confirm whether there is a higher correlation of elastographic scores with the histology of the lesion or with its size.

One of the limitations of our study was the small sample of malignant lesions compared to the benign lesions. However, this ratio is similar to the ratio observed in clinical practice. There was also a higher prevalence of lobular carcinomas than is described in the literature, and this may have influenced our results.

The classification by elastography proposed here, through the evaluation of tissue after compression and decompression of the breast parenchyma, can be an important tool, combined with ultrasonographic studies, for differentiating benign and malignant lesions among lesions classified as true masses according to the BI-RADS™ lexicon.

**Figure 7.** a–d. Example of breast carcinoma. B-mode study (a), histology (b), elastography during compression (c), and after decompression (d). Large irregular mass (4.2 cm) classified as BI-RADSTM 5. In the elastographic study, this lesion was classified as score 4 with no color variation observed during or after decompression. Histology of the biopsy specimen showed invasive ductal carcinoma.

### References