Evaluation of Virtual Touch Tissue Imaging Quantification (VTIQ), a new Shear Wave Velocity Imaging Method, for Breast Lesion Assessment by Ultrasound

Authors
Michael Golatta, Mirjam Schweitzer-Martin, Aba Harcos, Anne Stieber, Christina Gomez, Iris Maeder-Joepgen, Sarah Schott, Hans Junkermann, Geraldine Rauch, Florian Schütz, Christof Sohn, Joerg Heil

Affiliation:
University Breast Unit
Vossstrasse 9
69115 Heidelberg

Corresponding author:
Dr. Michael Golatta
University Breast Unit
Vossstrasse 9
69115 Heidelberg
Tel +49-6221-56-37637
Email Michael.Golatta@med.uni-heidelberg.de
Abstract

**Purpose:** To prospectively evaluate Virtual Touch Tissue Imaging Quantification (VTIQ) as a new elastography method concerning its intra- and inter-examiner reliability, its ability to differentiate benign from malignant lesions in comparison to and in combination with B-mode BIRADS® assessment.

**Materials and Methods:** B-mode ultrasound and VTIQ were performed by an experienced and an inexperienced examiner in 103 women with 104 breast lesions before biopsy. The study was approved by an independent ethics committee. All participants gave informed consent. Intra- and inter-examiner reliability of VTIQ was assessed. BIRADS® assessment was applied prior to five VTIQ measurements. The area under the receiver operating curve (AUC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (PNV) of BIRADS®, VTIQ and combined data were compared.

**Results:** Fifty-four of 104 lesions were malignant (51.9%). Intra-examiner reliability was consistent and inter-examiner agreement showed a strong positive correlation assessed with orthogonal regression (intercept=-0.91, slope=1.07, r=0.93). The mean VTIQ values in malignant lesions were significantly higher than in benign (7.73 m/s ± 1.02 versus 4.46 m/s ± 1.87; p<0.0001). The optimal cut-off for clinical decision making was 5.18 m/s, yielding a sensitivity of 98%, a specificity of 68%, a PPV of 77% and a NPV of 97%. The combination of BIRADS® (AUC=0.96) and VTIQ (AUC=0.94) led to improved test validity (AUC=0.98, p<0.0001).

**Conclusion:** VTIQ is a highly reliable method. There is a significant difference regarding the mean maximum velocity of benign and malignant lesions. Adding VTIQ to BIRADS® assessment improves the specificity.
Keywords

Breast ultrasound, shear wave elastography, cut-off value, breast lesions
Introduction

Breast ultrasound (US) is used to differentiate between benign and malignant lesions. The B-mode technique is commonly applied to evaluate morphologic features as for example shape and margin of a lesion (1-3). Unfortunately B-mode US suffers from low specificity (4). A newer development is elastography, which is a method for measuring the stiffness of tissue (5, 6). Within the last few years various studies have evaluated different methods of elastography showing that elastography has the same diagnostic performance as B-mode US and in combination with B-mode US it can help improve the differentiation of benign and malignant lesions (7-9). Since malignant lesions often alter tissue stiffness the basic hypothesis of all elastography techniques is that malignant tissue is stiffer than benign tissue and the measured stiffness can help differentiate between benign and malignant (10-12).

First studies on freehand elastography showed several disadvantages; most meaningful is the inter-examiner variability concerning the acquisition and interpretation and the qualitative strain information (13, 14). In spite of these limitations of the method the diagnostic accuracy of elastography was at least as good as that of standard B-mode BIRADS® assessment (15-17). In order to improve the diagnostic quality in terms of sensitivity, specificity and predictive value examiner independency and reproducibility have to be addressed. One possibility to improve the diagnostic accuracy of elastography would be a highly reproducible quantitative method which could be compared objectively. The Virtual Touch Tissue Imaging Quantification (VTIQ) method has the potential to technically overcome the limitations of former approaches because it is meant to be examiner independent, reproducible and results in an absolute measurement of tissue stiffness in the region of interest.

The aim of this study is to prospectively evaluate this new method of Virtual Touch Tissue Imaging Quantification (VTIQ) concerning its intra- and inter-examiner reliability. In addition its ability of differentiating benign from malignant lesions on the basis of the lesion’s stiffness
will be evaluated and compared to conventional B-mode BIRADS® assessment and the combination of both methods.
Material and Methods

Virtual Touch Tissue Imaging Quantification (VTIQ)

Virtual Touch Tissue Imaging Quantification (SIEMENS Medical Solutions, Mountain View, CA, USA) is a new elastography technique. In the beginning elastography relied on manual compression and decompression applied by the examiner (18). To improve examiner independency and reproducibility the tissue compression is automated in VTIQ. The probe generates a longitudinal push pulse which causes minimal localized displacement which is tracked by a detection pulse (9, 19, 20). Therefore - compared to other available elastography techniques - measuring the shear wave propagation induced by the push pulse is meant to be the most standardized and examiner independent technique (14). As in Acoustic Radiation Force Impulse (ARFI) Imaging the 9 MHz probe (9L4 Siemens) with the technical possibility of VTIQ generates a low frequent longitudinal push pulse. The push pulse induces shear waves which travel perpendicular to the ultrasound beam (21). Rather than detecting the axial displacement (= ARFI) VTIQ measures the speed of the perpendicular shear waves by a detection pulse. Because the speed of the shear waves propagating through the tissue is proportional to the stiffness of the tissue, a colored map in the region of interest (ROI) gives information on the tissue stiffness in the region of interest (14, 21). The shear wave velocity can be quantitatively measured in meters per second (m/s) within the ROI, up to 8.40 m/s (19). Manual precompression of the tissue changes the elasticity and makes the tissue stiffer (22); to obtain the most reproducible and optimal images it is therefore essential to minimize the precompression. Because shear waves cannot propagate in vicious fluid no signal can be measured e.g. in cysts (23, 24).

Patients and study design
The patients involved in the study were referred to a specialized diagnostic breast clinic to clarify clinical symptoms or already assumed imaging abnormalities.

Before scanning with VTIQ the women underwent standard clinical routine imaging. This work-up consisted of a clinical examination, ultrasound and mammography (for women aged 40 or older). BIRADS® assessment was completed after the routine work-up and before VTIQ was applied. After being examined women of the age of 18 or older with focal breast lesions assessed BIRADS® 3 to 5 visible in standard ultrasound were invited to participate in the study. From May to August 2012 a total of 125 patients were screened for the study, twelve patients refused to participate, 113 patients with 114 lesions gave informed consent and were finally examined using VTIQ. The vote of an independent ethics committee has been received.

100 lesions categorized BIRADS® 3, 4a, 4b, 4c or 5 underwent ultrasound guided biopsy (BIP HistoCore® 14 G). In these cases histology was used as gold standard. Four cysts categorized BIRADS® 3 were aspirated. In these cases the clinical diagnosis was used as gold standard. Ten BIRADS® 3 lesions did not undergo biopsy and were therefore excluded from the study.

All patients were scanned with the ultrasound system ACUSON S3000 US unit equipped with the VTIQ software (SIEMENS Medical Solutions, Mountain View, CA, USA) both by a consultant specialized in breast imaging and an inexperienced examiner. The examiners had been trained by a SIEMENS sonographer. The first 25 patients scanned with the VTIQ were not included in the study to avoid influencing the study results by inexperience.

For VTIQ the consultant (= examiner 1) obtained the first set of five VTIQ images. Step by step the consultant selected the lesion with the 9 MHz probe at its largest dimension in B-mode, adjusted the size of the VTIQ measuring box to include the lesion and surrounding tissue. Next, the maximum velocity of the lesion was measured by selecting the ROI. The
second set of five VTIQ images was then obtained by the inexperienced examiner (= examiner 2) in the same way to test the reliability and the examiner independency of VTIQ.

**Statistical analysis**

This study is of explorative character. All statistical analyses are of descriptive nature. Statistical tests and resulting p-values are not adjusted for multiplicity and can therefore only be interpreted descriptively. To begin with, the study cohort was described by the measures of the empirical distribution. Depending on the scale level of the variable, mean and standard deviation, median and quartiles, minimum and maximum and absolute and relative frequencies are calculated. In a second step, the intra- and interexaminer reliability was assessed. To calculate the intraexaminer reliability, the VTIQ values of the five measurements of each examiner were compared by calculating the range of these values (minimal value - maximal value). A range of 0 thus corresponds to perfect agreement whereas a range larger than 0 reveals that the five measurements deviate. The mean range over all patients and corresponding standard deviation was calculated for both examiners. In order to assess interexaminer reliability, the means of the five measurements were compared between examiner 1 and 2. Pearson correlation coefficient was calculated and an orthogonal regression line was fitted. A slope close to 1 and a constant close to 0 would thus correspond to good agreement between raters.

In order to assess diagnostic accuracy for VTIQ and BIRADS®, a ROC analysis was performed for each of the diagnostic instruments. The optimal cut-off was determined by considering sensitivity and specificity, but focusing on sensitivity due to its higher clinical relevance. Moreover positive and negative predictive values were calculated. Thereby, we used the fact that our sample is representative for the study population and that for this reason prevalences can be estimated from the sample. To assess the improvement of diagnostic
accuracy by combining VTIQ and BIRADS® a logistic regression was performed. Using the logistic regression model, positive predictive values can be calculated for each combination of VTIQ and BIRADS® categories. The positive predictive values can thus be interpreted as a function of VTIQ and BIRADS® yielding a transformation of the two variables into a one dimensional space. An extended ROC analysis based on the positive predictive function can thus be performed in order to directly compare the improvement of the combination of VTIQ and BIRADS® to BIRADS® alone.

All analyses were done using software SAS JMP version 6.0.
Results

Description of the study cohort

The final analysis was based on 104 lesions in 103 patients (mean age 51 years ± 18.56, range 20-89 years). The mean size of the palpable lesions was 21 mm ± 10.06 (range 7-53 mm) and of the non-palpable 13 mm ± 5.98 (range 6-32 mm). Histological examination showed 50 benign (48.1%) and 54 malignant lesions (51.9%). Table 1 lists the absolute and relative frequencies of each BIRADS® category, the mean and standard deviation for the VTIQ values for both examiners and the histologic diagnoses of the examined benign and malignant breast lesions.
Table 1: Absolute and relative frequencies of each BIRADS® category and corresponding mean VTIQ values in m/s.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total n =104</th>
<th>Benign n=50</th>
<th>Malignant n=54</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIRADS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>15 (14.4 %)</td>
<td>15 (30 %)</td>
<td>0 (0 %)</td>
</tr>
<tr>
<td>4a</td>
<td>32 (30.8 %)</td>
<td>29 (58 %)</td>
<td>3 (5.6 %)</td>
</tr>
<tr>
<td>4b</td>
<td>13 (12.5 %)</td>
<td>4 (8 %)</td>
<td>9 (16.7 %)</td>
</tr>
<tr>
<td>4c</td>
<td>18 (17.3 %)</td>
<td>2 (4 %)</td>
<td>16 (29.6 %)</td>
</tr>
<tr>
<td>5</td>
<td>26 (25 %)</td>
<td>0 (0 %)</td>
<td>26 (48.1 %)</td>
</tr>
<tr>
<td>VTIQ Mean Examiner 1</td>
<td>6.16 ± 2.21</td>
<td>4.46 ± 1.87</td>
<td>7.73 ± 1.02</td>
</tr>
<tr>
<td>VTIQ Mean Examiner 2</td>
<td>5.71 ± 2.37</td>
<td>3.82 ± 1.79</td>
<td>7.46 ± 1.24</td>
</tr>
</tbody>
</table>
Intra- and interexaminer reliability

For each lesion, the range (maximum minus minimum) of the five measurements was calculated for the experienced examiner (= examiner 1) and for the less experienced examiner (= examiner 2). For examiner 1, the mean range was given by 1.05 ± 0.85, indicating that the intra-examiner measurement agreement deviates on average about one score unit while examiner 2 had a mean range of 0.92 ± 0.74 which corresponds with a slightly better agreement.

In order to assess the interexaminer reliability, an orthogonal regression between the mean measurements of the two examiners was performed. The resulting regression line has a slope of 1.07 and a constant of -0.91 indicating that the agreement was good but that the rating of examiner 2 was systematically lower than for examiner 1. Pearson's correlation coefficient was given by r=0.93.

Differentiation of benign and malignant lesions

Benign lesions showed a mean shear wave velocity of 4.46 m/s ± 1.87 and malignant lesions of 7.73 m/s ± 1.02 being significantly higher (p < 0.0001, figure 1). The mean velocity values for every BIRADS® category are shown in table 2. According to the ROC analysis the statistically recommended optimal cut-off, obtained by maximizing the sum of sensitivity and specificity, for VTIQ based on the mean of the five measurements of the first examiner was 7.13 m/s showing a sensitivity of 85% (46 of 54) and a specificity of 92% (46 of 50). In order to improve the sensitivity the optimal cut-off for clinical decision making concerning a low false negative rate was chosen to be 5.18 m/s. This cut-off yielded a sensitivity of 98% (53 of 54), a specificity of 68% (34 of 50), a positive predictive value (PPV) of 77% (53 of 69) and a negative predictive value (NPV) of 97% (34 of 35). Among the benign lesions 16 of 50 (32%)
showed a mean velocity higher than the cut-off of 5.18 m/s, ranging from 5.69 to 8.13 m/s. Within the malignant lesions one out of 55 (1.8%) showed a lower mean velocity of 3.06 m/s. Using standard B-mode BIRADS® assessment (with a cut-off level of BIRADS® 4a) sensitivity, by definition, would be 98% (in our study it was 100%) and specificity 30%, PPV 61% (54 of 89) and NPV 100% (15 of 15). The combination of B-mode BIRADS® and VTIQ showed a sensitivity of 91% and a specificity of 96%. Improving sensitivity to 98% for clinical decision making results in a specificity of 82%, a PPV of 86% (53 of 62) and a NPV of 98% (41 of 42). The area under the curve (AUC) for VTIQ alone was 0.94 and for BIRADS® 0.96 (figure 2). The combination of standard B-mode BIRADS® assessment and VTIQ resulted in the best discrimination between benign and malignant lesions (AUC=0.98; \( p < 0.0001 \), figure 3). The bivariate logistic regression analysis confirmed the finding, that BIRADS® and VTIQ together improve the prediction of the pathological result of the lesion (\( p \)-value of the model < 0.0001).
B-mode and VTIQ images of a 55-year-old woman show an invasive lobular carcinoma of 0.9 cm (G2). The maximum velocity of 8.40 m/s was measured in the center of the lesion.
Table 2: Mean velocity values for each BIRADS® category in m/s.

<table>
<thead>
<tr>
<th>BIRADS</th>
<th>Total n=104</th>
<th>Malignant n=54</th>
<th>Benign n=50</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.13 ± 1.47</td>
<td>-</td>
<td>3.13 ± 1.47</td>
</tr>
<tr>
<td>4a</td>
<td>4.89 ± 1.77</td>
<td>5.93 ± 2.58</td>
<td>4.78 ± 1.69</td>
</tr>
<tr>
<td>4b</td>
<td>7.33 ± 1.33</td>
<td>7.71 ± 0.93</td>
<td>6.47 ± 1.83</td>
</tr>
<tr>
<td>4c</td>
<td>7.61 ± 1.02</td>
<td>7.83 ± 0.79</td>
<td>5.81 ± 1.06</td>
</tr>
<tr>
<td>5</td>
<td>7.87 ± 0.77</td>
<td>7.87 ± 0.77</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 2: ROC-analysis of VTIQ alone.
Figure 3: ROC-analysis for the combination of VTIQ and BIRADS®.
Discussion

Up to now the differentiation of breast masses has been based on B-mode ultrasound which suffers from low specificity (4). As a possibility to overcome the low specificity and improve intra- and interexaminer reliability of ultrasound breast diagnostics we evaluated VTIQ as a new elastography method. We investigated the quantitative velocity values of different breast lesions with VTIQ and evaluated its diagnostic performance as a stand-alone method and in combination with standard B-mode BIRADS® assessment.

Two previous studies combining a different quantitative Shear Wave Elastography (SuperSonic Imagine, Aix-en-Provence, France) with B-mode ultrasound showed improved specificity without loss of sensitivity and an increasing AUC for the combination of these two methods (7, 8). For intrarater and interrater reliability prior studies have shown high correlation coefficients for Shear Wave Elastography (intrarater correlation coefficient 0.87, interrater correlation coefficient 0.87) (15, 16).

To use VTIQ on a regular basis in clinical routine it is important that it is highly reliable. The interexaminer measurement agreement for VTIQ had a high positive correlation even between experienced and inexperienced examiners (r=0.93). Also the intraexaminer measurement agreement was consistent.

To differentiate benign from malignant lesions a cut-off yielding a high sensitivity and specificity is needed. Different levels of cut-off values (mainly in kPa) have been published being specific to ultrasound / elastography systems which make it impossible to compare the cut-off values between different systems.

BIRADS® 3 is defined to be benign in more than 98%. To compare VTIQ with standard B-mode based BIRADS® assessment and the combination of these two methods we set the sensitivity of both methods at 98%. For VTIQ this resulted in a cut-off of 5.18 m/s yielding a
specificity of 68% compared to 30% for BIRADS®. The combination of VTIQ with BIRADS® yielded a specificity of 82%. Combining VTIQ as a diagnostic method with the B-mode based BIRADS® assessment clearly improved the specificity without loss of sensitivity.

Using another Shear Wave Elastography technique (SuperSonic Imagine, Aix-en-Provence, France) in combination with B-mode based BIRADS® resulted in an AUC between 0.962-0.982 being statistically significantly superior to either B-mode based BIRADS® or Shear Wave Elastography alone (7, 8, 16). Combining VTIQ with B-mode based BIRADS® in our study came to a comparable AUC of 0.98. The combination increased the specificity without loss of sensitivity.

It is difficult to objectively quantify the amount of pressure applied while obtaining the VTIQs. Our results show a systematic difference between the measured values of the two examiners indicated by the drift of -0.91 in the orthogonal regression analysis. This could be referred to the different amount of pressure applied by the two examiners. In spite of the difference the test performances - applied by the two different examiners - are the same (AUC (examiner 1) =0.94 versus AUC (examiner 2) =0.94).

For applying VTIQ in the clinical routine we suggest scanning and assessing breast lesions with the standard BIRADS® categories. Once BIRADS® assessment has been applied the velocity of the lesion can additionally be measured with VTIQ. If the measurement is above or below the range of ±1 m/s of the cut-off of 5.18 m/s one measurement is sufficient and might be used to up- or downgrade the BIRADS® category by one. If the measured value is within this range the measurement should be repeated to confirm the value. However this study was limited by the number of cases included. To evaluate the legitimization of up- or
downgrading BIRADS® categories by adding VTIQ a larger study with a higher number of cases for each BIRADS® category is needed.

**Conclusion**

VTIQ is a highly reliable method concerning intra- and interexaminer agreement. There is a significant difference with respect to the mean maximum velocity of benign and malignant lesions showing malignant lesions to be stiffer. Adding VTIQ to the BIRADS® assessment improves the specificity.

**Conflict of interest**

Financial activities related to the present article: institution received grant for research support and loan of prototype ultrasound machine (S3000) to do research from Siemens Healthcare AG.

Financial activities not related to the present article: M.G. received payment for lectures from Siemens Healthcare AG.

Otherwise there is no potential conflict of interest.
Literature


