Prediction of Pathological Complete Response to Neoadjuvant Chemotherapy for Primary Breast Cancer Comparing Interim Ultrasound, Shear Wave Elastography and MRI

Prädiktion der pathologischen Komplettremission nach neoadjuvanter Chemotherapie bei primärem Mammakarzinom mittels Interim Ultraschall, Shearwave-Elastografie und MRT

ABSTRACT

Background Prediction of pathological complete response (pCR) of primary breast cancer to neoadjuvant chemotherapy (NACT) may influence planned surgical approaches in the breast and axilla. The aim of this project is to assess the value of interim shear wave elastography (SWE), ultrasound (US) and magnetic resonance imaging (MRI) after 3 cycles in predicting pCR.

Methods 64 patients receiving NACT had baseline and interim US, SWE and MRI examinations. The mean lesion stiffness at SWE, US and MRI diameter was measured at both time points. We compared four parameters with pCR status: a) Interim mean stiffness ≤ or > 50 kPa; b) Percentage stiffness reduction; c) Percentage US diameter reduction and d) Interim MRI response using RECIST criteria. The Chi square test was used to assess significance.

Results Interim stiffness of ≤ or > 50 kPa gave the best prediction of pCR with pCR seen in 10 of 14 (71 %) cancers with an interim stiffness of ≤50 kPa, compared to 7 of 50 (14 %) of cancers with an interim stiffness of >50 kPa, (p < 0.0001) (sensitivity 59 %, specificity 91 %, PPV 71 %, NPV 86 % and diagnostic accuracy 83 %). Percentage reduction in stiffness was the next best parameter (sensitivity 59 %, specificity 85 %, p < 0.0004) followed by reduction in MRI diameter of >30 % (sensitivity 50 % and specificity 79 %, p = 0.03) and % reduction in US diameter (sensitivity 47 %, specificity 81 %, p = 0.03). Similar results were obtained from ROC analysis.

Conclusion SWE stiffness of breast cancers after 3 cycles of NACT and changes in stiffness from baseline are strongly associated with pCR after 6 cycles.

ZUSAMMENFASSUNG

Hintergrund Die Prädiktion der pathologischen Komplettremission (pCR) bei primärem Mammakarzinom nach neoadjuvanter Chemotherapie (NACT) kann die geplante chirurgische Herangehensweise in Brust und Axilla beeinflussen. Ziel dieser
Arbeit ist die Beurteilung des Wertes von Interim Shearwave-Elastografie (SWE), Ultraschall (US) und Magnetresonanztomografie (MRT) nach 3 Zyklen für die Vorhersage einer pCR. **Methoden** Bei 64 Patienten mit NACT wurden Anfangs- und Interim-Untersuchungen mittels US, SWE und MRT durchgeführt. Die mittlere Steifigkeit der Läsion mittels SWE-, US- und MRT-Durchmesser wurde zu beiden Zeitpunkten bestimmt. Wir verglichen vier Parameter mit dem pCR-Status: a) Interim mittlere Steifigkeit ≤ oder > 50 kPa, b) prozentuale Reduktion der Steifigkeit, c) prozentuale Reduktion des US-Durchmessers und d) Interim MRT-Antwort mittels RECIST-Kriterien. Der Chi-Square-Test wurde zur Beurteilung der Signifikanz angewandt. **Ergebnisse** Die Interim-Steifigkeit von ≤ oder > 50 kPa führte zur besten Prädiktion einer pCR, mit einer pCR bei 10 von 14 (71 %) Karzinomen mit einer Interim-Steifigkeit von ≤ 50 kPa im Vergleich zu 7 von 50 (14 %) Karzinomen mit einer Interim-Steifigkeit von > 50 kPa. (p < 0,0001) (Sensitivität 59 %, Spezifität 91 %, positiver Vorhersegewert 71 %, negativer Vorhersagewert 86 % und diagnostische Treffsicherheit 83 %). Die prozentuale Reduktion der Steifigkeit war der zweitbeste Parameter (Sensitivität 59 %, Spezifität 85 %, p = 0,0004) gefolgt der Reduktion des MRT-Durchmessers um > 30 % (Sensitivität 50 % und Spezifität 79 %, p = 0,03) und einer prozentualen Reduktion des US-Durchmessers (Sensitivität 47 %, Spezifität 81 %, p = 0,03). Ähnliche Ergebnisse wurden mittels ROC-Analyse erzielt. **Schlussfolgerung** Die SWE-Steifigkeit von Mammakarzinomen nach 3 NACT-Zyklen und die Veränderungen der Steifigkeit im Bezug zum Anfangswert sind eng mit einer pCR nach 6 Zyklen assoziiert.

Advances in Knowledge
SWE stiffness and changes in stiffness after 3 cycles of NACT are strongly associated with pCR after 6 cycles. Percentage change in a combination of change in US diameter and reduction in stiffness is a useful parameter for assessing interim response to NACT.

Introduction
Pathological complete response (pCR) is increasingly common after neoadjuvant chemotherapy (NACT) for primary invasive breast cancer. Early prediction of pCR may influence planned surgical approaches in both the breast and axilla. Knowledge of response before the end of chemotherapy is helpful when planning complex autologous breast reconstructions especially if pre-reconstruction sentinel node biopsy is contemplated.

Interim assessment of response is usually carried out using dynamic contrast-enhanced magnetic resonance imaging (MRI) which has been shown to be superior to mammography and clinical examination in this clinical setting [1]. In many series MRI is also superior to ultrasound (US) [1, 2]. Standard MRI assessment uses RECIST criteria which are based on changes in unidimensional measurements. However, volumetric, functional and texture-based MRI assessments have been shown to be superior at interim prediction, particularly in luminal and hybrid subgroups [3 – 6]. MRI has the disadvantages of being expensive, time-consuming and there are availability issues at many centers.

Shear wave elastography (SWE) is an ultrasound imaging method which allows reproducible quantification of lesional and peri-lesional stiffness and it has been shown to aid benign/malignant differentiation of breast masses [7, 8]. It has also been shown to be an independent predictor of nodal metastasis and response to NACT at baseline [9, 10]. SWE is fast to perform and interpret.

One previous study has assessed the value of SWE in predicting pCR at the end of NACT [11]. The authors found that SWE performed similarly to MRI and better than US. Two previous studies looked at interim assessment with SWE. One study assessed interim prediction with SWE and strain elastography and response [12], while the other compared interim SWE alone to response [13]. Neither of these studies compared the SWE findings with those of MRI or grayscale US, the modalities routinely used in this clinical setting. We therefore hypothesized that SWE may be useful in the interim assessment of response in women undergoing NACT. The aim of this project was to assess the value of interim SWE compared to US and MRI after 3 cycles in predicting pCR after 6 cycles of NACT.

Materials and methods
64 patients with breast cancer receiving NACT were recruited into an ethically approved study which included baseline and interim US, SWE and MRI examinations. No patients were excluded after initial inclusion. Interim imaging examinations were performed after 3 cycles of chemotherapy, with patients proceeding to 6 cycles in total. All US scans were performed by one of five breast radiologists or one advanced radiography practitioner trained to perform and interpret breast ultrasound using an Aixplorer® ultrasound system (SuperSonic Imagine, Aix en Provence, France). These practitioners had between 7 and 22 years of breast ultrasound experience and had at least 12 months of experience performing SWE of solid breast lesions.

Four SWE images in two orthogonal planes were obtained. The ROI utilized in all cases was 2 mm in diameter (Fig. 1). The maximum diameter was obtained from the US images. The mean stiffness in kPa was taken as the average of the values taken from four SWE images acquired on two orthogonal planes. Imaging data was acquired prospectively, i.e., without knowledge of the final pathological outcome. If the lesion was not seen on US, the US diameter was taken as 1 mm, while SWE readings in the region of the previous abnormality were taken. This was identified by looking at the body mark on the previous US scan and viewing the pre-treatment MRI. Percentage reductions in tumor diameter and mean stiffness were calculated. Cut-off values for percentage...
Results

The mean age of the 64 patients was 52yrs (range: 25 – 79yrs). 25 patients had triple negative tumors, 21 had HER2 +ve cancers and 18 had luminal cancers. The first 3 cycles of chemotherapy in all patients consisted of 5-fluorouracil, epirubicin and cyclophosphamide (FEC). The mean pre-NACT stiffness was 127.4 kPa (range: 55 – 289, SD = 54) and the mean pre-treatment US size was 27 mm (range: 8 – 45, SD: 9). pCR occurred in 17 of 64 (27 %) women (Fig. 2), while 47 patients had residual disease (Fig. 3). At interim scanning, a residual mass visible on US was seen in 62 of 64 (97 %) women. Of the 2 women with no mass on interim US scanning, one had a focal area of residual stiffness. Both of these patients had a pCR. The mean interim stiffness was 90.3 kPa (range: 17 – 218, SD: 55) and the mean interim US size was 21 mm (range: 1 – 43, SD: 10).

pCR was seen in 10 of 14 (71 %) women where masses had an interim stiffness value of ≥ 50kPa, compared to 7 of 50 (14 %) women whose masses had an interim stiffness value of < 50kPa, p < 0.0001. Using this 50kPa cut-off yielded a sensitivity of 59 %, specificity of 91 %, PPV of 71 %, NPV of 86 % and a diagnostic accuracy of 83 % for pCR (Table 1). Increasing or lowering this threshold did not improve overall performance.

The cut-off value for the percentage reduction in stiffness giving the same proportion of patients (17/64, 27 %) as those showing a pCR was 50 %. The cut-off value for the percentage reduction in US diameter giving the same proportion of patients (17/64, 27 %) as those showing a pCR was 61 %. The cut-off value for the percentage reduction in stiffness and US diameter combined and giving the same proportion of patients (27 %) as those showing a pathological pCR was 114 %. The associations between pCR and changes in US and SWE parameters based on these cut-off values at interim scanning are shown in Table 1. The diagnostic accuracy using percentage reduction in stiffness, percentage reduction in US diameter and a combination of percentage reduction in stiffness and US diameter were 78 %, 72 % and 75 % respectively. The percentage reduction in stiffness and percentage reduction in US diameter gave statistically significant results (p < 0.0004 and p = 0.03, respectively). The combination of percentage reduction in stiffness and US diameter did not improve performance compared to that achieved by percentage reduction in stiffness alone (Table 1).

Baseline and interim MRIs were available for 59 of 64 (92 %) patients. Using a > 30 % reduction in maximum diameter as an indicator of response, the sensitivity, specificity and diagnostic accuracy of MRI were 50 %, 79 % and 71 %, respectively. The percentage reduction in MRI diameter and its association with pCR was statistically significant (p = 0.03) (Table 1).

ROC analysis yielded the following AUC measurements for predicting pathological complete response at interim assessment: change in MRI diameter: 0.68, percentage change in stiffness: 0.82, stiffness value at interim: 0.78, percentage change in US diameter: 0.67 and percentage change in a combination of change in US diameter and reduction in stiffness: 0.83. The ROC curves are shown in Fig. 4 – 8. All of the ROC curves shown on one graph are shown in Fig. 9. The only statistically significant difference in AUC measurements following a pairwise comparison.

Means and associated standard deviations were calculated for lesion size and stiffness before treatment and at interim. Receiver operator characteristic (ROC) curves were produced and the area under the curve (AUC) determined using Medcalc software. A pairwise statistical comparison of the ROC curves was also performed.
was the difference between the percentage change in US diameter and the combination of US diameter change and stiffness change (p = 0.02). Images where different imaging modalities yielded differing interim estimates of response are shown in ▶ Fig. 10.
Discussion

We have shown that interim SWE after 3 cycles of NACT using a simple threshold of 50 kPa for mean elasticity is strongly associated with the chance of pCR after 6 cycles of NACT. This simple threshold assessment has a stronger association with pCR than estimating the percentage reduction in stiffness. Both of these parameters appear to outperform grayscale US assessment using uni-dimensional diameter and MRI using RECIST criteria when analysis is performed using cut-off values. Analysis using ROC curves suggests that the percentage change in stiffness and percentage change in a combination of US diameter change and stiffness change are the best parameters at interim scanning to
predict pCR at the end of treatment. Regardless of the assessment method, the best parameters include SWE assessment.

A common problem when using grayscale US to assess tumor response to NACT is that many cancers when treated change from an ovoid shape to having a flatter, plaque-like appearance. When this happens, the uni-dimensional diameter can remain unchanged even if the tumor volume has reduced considerably. This problem can be overcome by using volumetric assessments using a 3D probe. One study has shown the use of a 3D probe to be helpful in this clinical setting [15]. However, 3D probes only have a footprint of 4 cm so they are not able to measure the volume of large tumors. Many tumors also cause considerable posterior acoustic shadowing which can obscure the posterior border of the tumor and thus interfere with volumetric assessment. Another problem when using US is that even when a pathological complete response has been achieved, a residual mass of fibrous tissue may be seen, measured and assumed to be viable tumor. One advantage of using SWE is that these residual fibrous masses with no residual cancer tend to be soft on SWE, so that assessment on SWE is less prone to errors of this nature. SWE benefits from being easy to perform and from the fact that the results have good reproducibility with the intra-class coefficient (ICCC).

Table 1 Associations between changes in US, SWE and MRI parameters at interim scanning and pCR at the end of NACT.

<table>
<thead>
<tr>
<th></th>
<th>pCR (n)</th>
<th>No pCR (n)</th>
<th>p-value</th>
<th>sensitivity</th>
<th>specificity</th>
<th>diagnostic accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>shear wave &lt; 50 kPa at interim</td>
<td>10</td>
<td>4</td>
<td>&lt;0.0002</td>
<td>59%</td>
<td>91%</td>
<td>83%</td>
</tr>
<tr>
<td>shear wave ≥ 50 kPa at interim</td>
<td>7</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stiffness &lt; 52 % of baseline</td>
<td>10</td>
<td>6</td>
<td>&lt;0.0004</td>
<td>59%</td>
<td>85%</td>
<td>78%</td>
</tr>
<tr>
<td>stiffness ≥ 52 % of baseline</td>
<td>6</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US diameter &lt; 61 % of baseline</td>
<td>8</td>
<td>9</td>
<td>0.026</td>
<td>47%</td>
<td>81%</td>
<td>72%</td>
</tr>
<tr>
<td>US diameter ≥ 61 % of baseline</td>
<td>9</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>combination reduction &lt; 114 %</td>
<td>9</td>
<td>8</td>
<td>&lt;0.004</td>
<td>53%</td>
<td>83%</td>
<td>75%</td>
</tr>
<tr>
<td>combination reduction ≥ 114 %</td>
<td>8</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRI &gt; 30 % reduction in diameter</td>
<td>8</td>
<td>9</td>
<td>0.03</td>
<td>50%</td>
<td>79%</td>
<td>71%</td>
</tr>
<tr>
<td>MRI ≤ 30 % reduction diameter</td>
<td>8</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4 ROC curve demonstrating the relationship between change in MRI diameter from baseline to interim scanning and pCR. The AUC is 0.68 (95% CIs 0.54 to 0.79).

Fig. 5 ROC curve demonstrating the relationship between percentage change in stiffness from baseline to interim scanning and pCR. The AUC is 0.82 (95% CIs 0.70 to 0.90).
for quantitative measurements between scans taken by different observers being 0.85 which is in the “near perfect” range if 4 images are taken on 2 planes. Stiffness measurements taken by different observers from the same scans have an ICC of 0.99. Qualitative assessments of SWE images (not performed in this study) are less reproducible [7, 14, 16]. The performance of SWE on breast masses has a few limitations. Lesions very deep in the breast can be difficult to assess and patients with breathing difficulties can also be impossible to scan accurately.
It is becoming more common to see pCR in breast cancer as NACT regimens become more effective, resulting in more women undergoing surgical resection of breast tissue even though no viable tumor is present [17]. In response to this, clinical study of the use of percutaneous sampling of the tumor bed using vacuum-assisted biopsy devices is becoming more attractive [18]. Before this can happen, imaging prediction of pCR has to improve. It is becoming clear that pure anatomical assessment of the tumor bed is not suitable for this purpose and that functional assessment either using MRI or other modalities such as PET/CT or
SWE offer superior prediction [3, 5, 19]. It is also becoming clear that different immunophenotypes of breast cancer have different rates and patterns of response to NACT and that this necessitates a tailored approach by the radiologist when assessing response to NACT [4].

Response of the primary tumor to NACT is also important for guiding axillary surgery, as response in the axilla is closely correlated to response in the breast. In patients who had biopsy-proven axillary disease at diagnosis but appear to have pCR prior to surgery, there is now some evidence to suggest the safety of a sentinel node biopsy rather than an axillary clearance, albeit with a higher false-negative rate than previously accepted [20]. Thus, correct and accurate assessment of response to NACT prior to surgery will influence both surgical options for the tumor bed and the axilla.

This study has a number of limitations. It comes from a single center with a special research interest in SWE but as the technique is straightforward to perform, we do not foresee difficulty replicating these results at other centers. The number of patients in the study is modest so differences in performance according to immunophenotype have not been assessed. It is possible that SWE may perform better in certain immunophenotypes in this setting as this is certainly the case when using SWE to predict response at baseline and when using MRI to predict pCR at interim scanning [4, 21].

Two previous similarly sized recent studies have addressed the topic of assessing interim response of breast cancer to NACT using SWE. Their findings are comparable to our study. What is new in the current study is the evidence that SWE interim assessment is at least as accurate as grayscale US or MRI using RECIST criteria. No studies have been performed to assess the value of SWE in evaluating response after 1 or 2 cycles of NACT. It is therefore unclear how soon SWE could be used to assess response after the commencement of NACT. A previous study has assessed the value of evaluating response to NACT with SWE at the end of NACT. They found that a threshold of 30kPa was useful in predicting response. They found this cut-off value more predictive than other higher values. They did not attempt to look at the value of percentage reduction in stiffness compared to baseline values [11].

In conclusion, SWE shows promise as a method of interim prediction of response in women with breast cancer treated with NACT and could be used to inform surgical decision making.

Conflict of interest

AE had a PhD student part funded supersonic imagine. The student did not contribute to this work.

Acknowledgments

This work was funded by Breast Cancer Now 2012ON46.

References


