Original Research

A Novel Marker, Based on Ultrasound Tomography, for Monitoring Early Response to Neoadjuvant Chemotherapy

Neb Duric, PhD,1,2,*, Peter Littrup, MD,1,2 Mark Sak, PhD,1 Cuiping Li, PhD,1 Di Chen, MD,1 Olivier Roy, PhD,1 Lisa Bey-Knight,1,2 Rachel Brem, MD3

1Delphinus Medical Technologies, Inc., Novi, MI; 2Wayne State University, Barbara Ann Karmanos Cancer Institute, Department of Oncology, Detroit, MI; 3George Washington University, Department of Radiology, Washington, DC

*Address correspondence to N.D. (e-mail: nduric@delphinusmt.com)

Abstract

Objective: To evaluate the combination of tumor volume and sound speed as a potential imaging marker for assessing neoadjuvant chemotherapy (NAC) response.

Methods: This study was carried out under an IRB-approved protocol (written consent required). Fourteen patients undergoing NAC for invasive breast cancer were examined with ultrasound tomography (UST) throughout their treatment. The volume (V) and the volume-averaged sound speed (VASS) of the tumors and their changes were measured for each patient. Time-dependent response curves of V and VASS were constructed individually for each patient and then as averages for the complete versus partial response groups in order to characterize differences between the two groups. Differences in group means were assessed for statistical significance using t-tests. Differences in shapes of group curves were evaluated with Kolmogorov–Smirnoff tests.

Results: On average, tumor volume and sound speed in the partial response group showed a gradual decline in the first 60 days of treatment, while the complete response group showed a much steeper decline (P < 0.05). The shapes of the response curves of the two groups, corresponding to the entire treatment period, were also found to be significantly different (P < 0.05). Furthermore, large simultaneous drops in volume and sound speed in the first 3 weeks of treatment were characteristic only of the complete responders (P < 0.05).

Conclusions: This study demonstrates the feasibility of using UST to monitor NAC response, warranting future studies to better define the potential of UST for noninvasive, rapid identification of partial versus complete responders in women undergoing NAC.

Key words: ultrasound tomography; breast cancer; monitoring treatment; sound speed; neoadjuvant chemotherapy.

Introduction

Locally advanced breast cancer represents a difficult clinical problem. Many patients with locally advanced disease experience relapse and eventual death from the disease (1). Data from the National Cancer Institute’s Surveillance, Epidemiology, and End Results program indicate that approximately 30 000 women are diagnosed with locally advanced breast cancer annually in the U.S. (2). The 5-year relative survival rate for women with stage III breast cancer is about 55%.

Neoadjuvant chemotherapy (NAC) increases the ability to control locally advanced breast carcinomas and promotes...
breast-conserving surgery (BCS) (3–6). NAC has become the standard of care for patients who have locally advanced and inflammatory breast cancer or who wish to pursue BCS in the U.S. (1). Achieving pathologic complete response (pCR) is a desirable outcome that can lead to improved survival. However, not all patients respond to NAC; if they do, their responses are highly variable. Thus, identifying poorly responding patients earlier would allow a timely switch to a different, possibly more effective, regimen, and/or would advance surgery. Patients in these categories would benefit by stabilizing or potentially reversing their disease, thereby reducing morbidity and mortality rates (3–6). Furthermore, predicting pCR would be highly beneficial for breast cancer drug development given the Food and Drug Administration’s acceptance of pCR as an endpoint to support accelerated approval.

Clinically, there has not been a universal, cost-effective adoption of any technology or technique that helps accurately assess, monitor, and predict individual patient response to NAC, particularly early on in the course of treatment (7).

Imaging data to support clinical decision making is limited and not routinely used in a standardized manner. Although handheld ultrasound (HHUS) is often used to monitor tumor size, this approach does not measure intrinsic tumor properties and is therefore not a sensitive measure of response in the early stages of treatment (8). MRI has been used to effectively quantify the clinical response of breast cancer to NAC (9–15). Studies have shown that diffusion-weighted imaging (9,10), dynamic contrast-enhanced MRI (11,12) and the apparent diffusion coefficient (13,14), for example, are effective at predicting pCR. Similarly, positron emission tomography (PET) imaging has shown great promise in predicting early response to chemotherapy and may have direct correlates to the higher tumor blood flow seen on MRI (16–19). Both imaging modalities have led to biomarkers showing correlation with surgical pathology findings to assess concordance and enhance the potential for preoperative planning. However, the high costs associated with both MRI and PET (7) have impeded widespread clinical acceptance of these imaging modalities.

While studies of cost-effective technologies, such as quantitative ultrasound spectroscopy and ultrasound elastography, have shown differentiation of responders from nonresponders as early as 4 weeks after treatment initiation (20,21), these technologies are also subject to limitations on practicality, namely their operator dependence and inability to capture full volumetric views of tumors.

Ultrasound tomography (UST) (22–30) is an emerging technology that, like MRI, provides automated, operator independent, full volumetric views of the breast but, unlike MRI, is fast (2–4 minute exams) and does not require contrast. Previous studies have shown that the speed of sound, determined from UST, can accurately quantify the density of breast tissue (31–33) and tumor volume (29,30). Such quantification would represent an advance in treatment monitoring beyond the current standard of care where tumor size changes are observed by standard imaging (e.g. HHUS) and tumor firmness is deduced from manual palpation. The purpose of this study was to evaluate UST as a potentially more practical imaging tool for assessing NAC response, particularly early in therapy.

**Methods**

**Patient Recruitment**

This study was carried out under an Institutional Review Board-approved protocol, in compliance with the Health Insurance Portability and Accountability Act. Informed consent was obtained from all patients. Patients were first identified as prospective NAC patients through inspection of electronic medical records. Patients were considered eligible if they: (1) were ≥18 years old; (2) had newly diagnosed locally advanced breast cancer (defined mainly as T2 or T3, but also smaller tumors that are close to the edge of the breast that might be more difficult to resect); (3) were eligible for but not yet receiving NAC; (4) were able to read and write in English; and (5) had no physical or mental condition preventing the ability to lie down on the UST device.

A baseline measurement was defined as the first study visit where a UST scan was performed. If the exam took place on the same day and just before the first chemotherapy cycle, the baseline measurement occurred on day 0 (0 days before the first cycle). If a patient was scanned earlier than that, at their initial physician meeting, for example, they were assigned a negative day number (e.g. –10 days means they were scanned 10 days before their first chemotherapy cycle).

**Data Collection**

The images and data were collected during the period March 23, 2007 to June 17, 2011. An early UST prototype (27) of the SoftVue system (Delphinus Medical Technologies, Novi, MI) was used in the original data collection, which included 21 patients (29). The current study represents a new analysis of the data using a new high-resolution reconstruction algorithm (30).

The patients enrolled in this study were scanned with UST over the course of their NAC therapy, yielding multiple time points for each patient. The scanning was performed

---

**Key Messages**

- Tumor volume (V) and sound speed (VASS) changes, derived from ultrasound tomography, can be used to monitor response to neoadjuvant chemotherapy.
- The combination of V and VASS can be used to characterize differences between patients who achieve a pathologic complete response and those who do not.
- Differences are apparent in the first 3 weeks of treatment, justifying future studies to develop models for early prediction of pathologic response.
by a research nurse. The UST exams were timed to coincide with each treatment visit and were therefore dictated by the oncologist prescribing the treatment. The intervals for such visits depended on the treatment regimen, meaning that they were either weekly or biweekly.

A tomographic reconstruction method (30) was used to generate stacks of cross-sectional sound speed images in order to quantify the volume and sound speed of the primary tumor in 3D and determine changes over time. A board-certified radiologist (P.L.) with 25 years of experience in breast imaging inspected the images and reports to determine the location and extent of the cancer. The primary tumor was the largest tumor and the one that underwent biopsy in all 14 cases. Secondary tumors and lymph nodes were not studied.

At the completion of the data collection, the data were compiled for each patient into a time series of sound speed image stacks.

### Endpoint Definitions

The study subjects were grouped according to whether they achieved complete response or not. Surgical pathology reports were used to assess whether patients achieved pCR. Complete response is defined as the disappearance of the lesion with pathological confirmation (no cancer cells upon resection). Partial response is defined as any tumor that did not achieve pCR. This definition included partial tumor shrinkage as well as stable and progressive disease.

### Sound Speed Measurements

The measurements were made according to the following sequence.

1. The volume of the tumor was calculated automatically through a pixel count of the segmented tumor images.
2. A peritumoral volume was calculated by defining a thin annulus (1 cm wide) using elliptical regions of interest in each image slice to define an annular region whose inner boundary enclosed the tumor and whose outer boundary defined the extent of the peritumoral region (Figure 1).
3. The average tumor sound speed was calculated automatically by summing pixel sound speed values in the segmented tumor and dividing by the tumor volume.
4. The peritumoral region was assessed similarly.
5. The difference between the volume-averaged tumor sound speed and its peritumoral region was used to determine the relative sound speed of the tumor and then normalized by its initial (baseline) value.

Measurement error lies in defining the posterior and anterior edges of the tumor. In the final analysis, the error bars shown are the standard errors of the mean.

### Characterization of Time-Dependent Changes

The volume (V) and volume-averaged sound speed (VASS) of the primary tumor were determined as a function of time for each patient and imported into a spreadsheet (Microsoft Excel, Microsoft, Bellevue, WA) for subsequent graphing and analysis. These time-dependent response curves were characterized individually for each patient by fitting exponential functions of the form $e^{-t/\tau}$. The exponential response time, $\tau$, represents the time it takes for the tumor to reduce its V or VASS to $1/e$ (37%) of its initial value, which was extracted from the best fit exponential curve for the first 60 days of treatment.

In addition to characterizing the response time of individual patients, group-averaged response times were also

---

Figure 1. Sound speed images (rendered in the coronal plane). A: Images in chronologic order starting at baseline and ending 10 weeks later. Changes in the tumor at 1 o'clock (arrow) can be discerned after just one chemotherapy cycle. B: Regions of interest (ROIs) used to measure the tumor (inner circle) and the peritumoral (outer circle) regions. Changes were quantified by measuring the volume and sound speed of the tumor (inner ROI). Tumor sound speed relative to the peritumoral region (annular region) was then measured as an indicator of relative sound speed.
evaluated. Patients were grouped as partial versus complete responders. The group-averaged response times, τ, were determined for both V and VASS. The significance of any differences in group values were assessed using t-tests.

Similarly, the group-averaged response curves for V and VASS were calculated from the individual time-dependent curves. The individual curves were first aligned by interpolating individual patient data into equal increments of 10 days on the time axis. The V and VASS values, corresponding to each 10-day increment, were then averaged from all patients at those time points for the complete and partial responder groups, respectively. The net result was a single response curve for each of the two groups. The error bars were calculated as standard errors of the mean from the average of all patient data at a particular time point within each group. To determine any differences in the response curves, a Kolmogorov–Smirnov (KS) test was used. Differences were deemed significant if they formed two distinct trends with \( P < 0.05 \).

Finally, to better visualize very early changes in the complete and partial response groups, a new parameter was defined as:

\[
\gamma = \frac{V \times \text{VASS}}{V_0 \times \text{VASS}_0}
\]

where \( \gamma \) represents a normalized multiplication of V \( \times \) VASS (\( V_0 \) and \( \text{VASS}_0 \) are baseline values). The above procedure was repeated for this parameter and the KS test was similarly applied.

To compare the performance of UST volume measurements against standard of care imaging, the above procedure was also performed for HHUS and MRI volume measurements. Mass sizes based on HHUS and MRI were collected. Sizes were presented as triaxial for HHUS and biaxial for MRI. In the former case, the tumor volume was estimated using the ellipsoidal formula \( d_1 \times d_2 \times d_3 \times \pi/6 \), where \( d_1 \), \( d_2 \), and \( d_3 \) are the three diameters of the HHUS measurement. In the case of MRI, many of the original studies were not available so, in those cases, only two axes of measurement were used since they could be obtained from the clinical reports. In such cases, the formula, \( d_1 \times d_2 \times (d_1 + d_2)/2 \times \pi/6 \) was used to estimate the volume.

The basic steps followed in this study, from data acquisition to characterization of response curves, are illustrated schematically in Figure 2.

**Results**

Although the original study included 21 patients, 7 datasets did not yield usable images because of poor signal quality (the waveform reconstruction method is more sensitive to signal quality because the smaller pixels it reconstructs contain less signal energy). Consequently, this study proceeded with a total of 14 patient data sets corresponding to 182 UST patient exams (average of 13 exams per patient). Of these 14 patients, 13 had HHUS studies and 8 had MRI examinations for comparison.

Patients ranged in age from 29 to 60 years. Patient height, weight, and body mass index were in the range of 61–68 inches, 118–230 lbs and 21–41 kg/m², respectively.

Most tumors were poorly differentiated (8/14, 57%) and either triple-negative (ER−PR−HER2−)/4/14, 29%) or HER2+ (5/14, 36%). Four patients (29%) achieved pCR while the remaining 10 (71%) did not. These numbers were too small to pursue a statistical analysis of differences between complete responders and partial responders in relation to tumor grade, receptor status, or type of NAC treatment.

V and VASS time curves were constructed for all 14 patients; an example is shown in Figure 3. Henceforth, V and VASS response times are referred to as shrinkage and softening times, respectively. Analysis of the time curves showed differences in the average shrinkage and softening times for the partial and complete responders (Table 1).

Some cases in the partial response group exhibited negative shrinkage/softening times, when the tumor grew and/or increased in sound speed (negative \( \tau \) values). Negative values presented a challenge for calculating average values of \( \tau \). Instead, a lower bound on the average \( \tau \) was calculated by averaging only the positive \( \tau \) values among the partial responders. The tumor volume change was thus characterized by an average shrinkage time of \( >183 \pm 61 \) days for the partial responders and \( 100 \pm 39 \) days for the complete responders, a significant difference (\( P = 0.028 \)). The average softening time for sound speed was \( >127 \pm 37 \) days for partial responders and \( 51 \pm 23 \) days for the complete responders, also a statistically significant difference (\( P = 0.003 \)).

Two of the 10 partial responders showed an increase in tumor volume in the first 60 days, while none of the complete responders did. In the case of sound speed, only 5 of the 10 partial responders exhibited decreasing sound speed while all 4 complete responders showed a decrease (Table 2).

Inspection of Table 1 and Table 2 suggests that sound speed...
offers greater discriminatory power than volume and that, when both decline, the complete responders are well differentiated from the partial responders.

The group-averaged response curves for volume data and VASS are shown in Figure 4. The complete and partial response groups were determined to be statistically different. The V and VASS distributions were distinct at a significance level of $P = 0.047$ and $P = 0.003$, respectively. The results for γ are shown in Figure 5. The overall responses of the two groups were found to be significantly different (KS test, $P = 0.012$), particularly in the first 3 weeks of treatment, where the difference between the two groups was greater compared to using V or VASS alone.

A similar analysis was performed for HHUS and MRI tumor volume measurements. For HHUS, no significant difference was found in the average tumor shrinkage time of the partial versus complete responders ($P = 0.799$). This analysis was not possible for MRI because only baseline scans were available for the complete responders.

**Discussion**

The results of this study demonstrate that almost all patients exhibited some degree of NAC response, as marked by changes in V and VASS. The data also suggest that patients with initial declines in $\gamma$ (both V and VASS) were more likely to achieve pCR. With this parameter, differences between the two groups arise more rapidly in the early phases of treatment (i.e. the first 2 weeks).

This study builds on previous work that was based on lower-quality sound speed images reconstructed from a simpler algorithm (29). While the earlier study also demonstrated the ability to monitor changes in and V and VASS, it did not yield statistically significant differences between the complete and partial responders. By contrast, this study demonstrates differences in the V, VASS, and $\gamma$ parameters between the two groups. The improvement is most likely attributable to the higher-resolution imaging used in this study.

While the partial responders exhibited some shrinkage of the tumor volume, sound speed changes within that reduced volume were relatively stable. The lack of softening in the partial response group may represent a surrogate endpoint biomarker for viable tumor in the remaining volume, as it shrinks. Recent studies suggest that early and greater presence of tumor-infiltrating lymphocytes (TILs), along with reestablishment of more normal vasculature, are also prognostic indicators of complete response (34). This may also reflect molecular levels of programmed death protein-1 on the TILs in response to programmed death ligand 1 expressed on the breast cancer cells. In contrast, the tumors among the complete responders are more likely to become softer, which
could be facilitated by these processes to produce greater tumor necrosis and/or less viable residual tumor early in the process. It is possible that the early softening measured by sound speed reflects underlying cellular and molecular processes that facilitate more thorough destruction of cancer cells in the residual tumor by the end of treatment, thereby leading to pCR.

**Predicting Response**

The lower limits for the shrinkage and softening times for partial responders, as determined above, represent possible cutoffs for separating the two groups. Thus, partial responders can be identified on the basis of response times being $\tau_V > 183$ days for volume and $\tau_S > 127$ days for sound speed (i.e. $\tau_V \land \tau_S > 183 \land 127$ days), where “$\land$” represents the logical “and” symbol. Applying this criterion to the partial responder data yields 2 false-positives and 0 false-negatives.

Predicting response would, however, benefit from characterizing even earlier changes. One way to better visualize such changes is through the parameter $\gamma$, as described earlier. With this parameter, differences between the two groups arise more rapidly in the early phases of treatment (i.e. the first 2 weeks), reaching a maximum difference at 30–50 days. As noted earlier, the overall responses of the two groups are significantly different. The fact that much of the differentiating effect occurs within the first 2 weeks highlights the need for frequent measurements immediately after the initiation of NAC.

The relative sparsity of HHUS and MRI datapoints in this study is the result of routine clinical practice at this institution, as resources were not available to support additional HHUS and MRI exams. This sparsity limited the ability to explore the early NAC response with these modalities, and also may explain the lack of observed differentiation between the complete and partial responders in this study. Any exponential fitting method would benefit from a high frequency of datapoints leading to more accurate predictive modeling. The need for frequent measurements has also been discussed in some MRI studies (9,10). Given the nonlinear nature of the tumor changes and the higher frequency of UST measurements, this study better captured the early changes, particularly in cases where V and VASS drop rapidly.

In previously published PET and MRI studies, where frequent measurements were available, differences in response could also be discerned within the first 1–3 weeks of treatment (9–19). Unfortunately, frequent use of PET and MRI imaging is limited by cost and logistics, while HHUS is limited because it only measures volume and does so with more operator dependence, and less accuracy, than the other modalities (8).

A unique aspect of this study is its ability to image patients safely and many times during their treatment. The exam is quick, well tolerated, and can be easily integrated into the clinical workflow and performed at each infusion visit. Such a high frequency of observation is unprecedented and opens the

---

**Figure 4.** A: The change in volume (V) between the two response groups, showing a steady decline for the partial responders and a steeper decline for the complete responders. B: The change in volume-adjusted sound speed (VASS) between the two groups, showing a steady decline for the complete responders but not for the partial responders.

**Figure 5.** The change in $\gamma$ (the product of volume $\times$ volume-adjusted sound speed) as a function of time (days), showing a steady decline for the partial responders (orange) but a much steeper decline for complete responders (blue), particularly within the first few weeks.
door to the long-term goal of providing a safe, cost-effective, and comfortable imaging strategy to measure locally advanced breast tumor response to NAC, allowing prediction of clinical and pathologic response early in the treatment process.

Study Weaknesses
A comparison with more frequent MRI measurements would have provided a better benchmark for the relative performance of UST. However, the high cost of adding MRI exams beyond standard of care (i.e. one scan at start and one at end of therapy) prohibited such a comparative study. This feasibility study, therefore, was limited to demonstrating that UST can measure tumor response and provide guidelines for a future comparative study to support the development of a robust predictive model.

While these results suggest that differences between the complete and partial response groups are significant, it is worth noting that this significance is driven by the sheer number of datapoints (a total of 182 patient exams). It is likely that the multiple datapoints per patient (an average of 13 per patient) is responsible for discerning the shapes of the response curves. Thus, it would be desirable to expand the patient set beyond the small numbers described here for a more definitive analysis.

Furthermore, while the results are suggestive of possible predictive parameters, this cannot be concluded at this time. First, this is a post hoc analysis, so any predictive parameter, such as response time, would have to be tested prospectively on a new set of patient data utilizing separate training and validation datasets. Second, the response time differences are averages over the two groups. They may well differentiate the two groups, but a useful predictive parameter would have to work at the individual level. Furthermore, the predictive power may vary by molecular tumor type, but the size of this study limits such an assessment at this time. A larger, prospective study is therefore needed to further investigate the predictive parameter at the individual level, using, for example, an receiver operating charateristic analysis.

Conclusions
In this feasibility study, the emerging technology of UST was used to show that the combination of tumor volume and sound speed is a sensitive marker of response to NAC. Quantification of tumor shrinkage and softening times showed that the patients who ultimately achieved pCR responded significantly more quickly than those that did not. Furthermore, it was possible to differentiate the complete and partial response groups within the first three weeks of treatment, in alignment with published results from MRI and PET studies.

These results demonstrate the feasibility of using UST to monitor NAC response, which warrants future studies to better define the potential of UST for noninvasive, rapid identification of partial versus complete responders in women undergoing NAC. Clinical decision making would improve by transitioning nonresponders to alternative treatment quickly or by demonstrating effective response to NAC. Leveraging the low cost of UST relative to MRI and PET would facilitate its translation into the clinic.

Funding
The authors acknowledge support from NIH grant R44CA165320.

Conflict of Interest Statement
Drs Duric and Littrup are co-inventors of, and have intellectual property interests in, the ultrasound technology used for the measurements described in this paper. Dr Brem sits on the board of directors of Delphinus Medical Technologies and the remaining authors are employees of Delphinus.

References


