COHERENCE-BASED BEAMFORMING INCREASES THE DIAGNOSTIC CERTAINTY OF DISTINGUISHING FLUID FROM SOLID MASSES IN BREAST ULTRASOUND EXAMS

ALYCEN WIACEK,* ENIOLA OLUYEMI,† KELLY MYERS,† LISA MULLEN,‡ and MUYINATU A. LEDIJU BELL*†,‡§

* Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, Maryland, USA; † Department of Radiology and Radiological Science, Johns Hopkins Medicine, Baltimore, Maryland, USA; ‡ Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland, USA; and § Department of Computer Science, John Hopkins University, Baltimore, Maryland, USA

(Received 24 September 2019; revised 16 January 2020; in final from 20 January 2020)

Abstract—Ultrasound is often used as a supplement for mammography to detect breast cancer. However, one known limitation is the high false-positive rates associated with breast ultrasound. We investigated the use of coherence-based beamforming (which directly displays spatial coherence) as a supplement to standard ultrasound B-mode images in 25 patients recommended for biopsy (26 masses in total), with the eventual goal of decreasing false-positive rates. Because of the coherent signal present within solid masses, coherence-based beamforming methods allow solid and fluid-filled masses to appear significantly different \((p < 0.001)\). When presented to five board-certified radiologists, the inclusion of robust short-lag spatial coherence (R-SLSC) images in the diagnostic pipeline reduced the uncertainty of fluid-filled mass contents from 47.5% to 15.8% and reduced the percentage of fluid-filled masses unnecessarily recommended for biopsy from 43.3% to 13.3%. These results are promising for the potential introduction of R-SLSC (and related coherence-based beamforming methods) into the breast clinic to improve diagnostic certainty and reduce the number of unnecessary biopsies. (E-mail: awiacek1@jhu.edu)

Key Words: Breast ultrasound, Coherence-based beamforming, Breast cancer.

INTRODUCTION

According to the American Cancer Society, one in eight women will be diagnosed with breast cancer during their lifetimes (Siegel et al. 2019), making early diagnosis increasingly important to improve prognosis. Currently, mammography is the most widely used screening modality, and it is considered one of the most important imaging methods to reduce the likelihood of death from breast cancer (DeSantis et al. 2011). However, mammographic sensitivity \((i.e., the ability of a mammogram to detect cancer)\) is significantly lower in patients with dense breast tissue (Mandelson et al. 2000). When breast tissue is dense, the surrounding parenchyma can often obscure diagnostic features in the mammogram, necessitating a supplement to improve detection.

Various studies have examined the impact of using other imaging modalities to supplement screening mammography, such as magnetic resonance imaging (MRI), molecular breast imaging (MBI), contrast-enhanced mammography (CEM) and ultrasound. A study by Kuhl et al. (2005) found the sensitivity of MRI to be 90.7% alone and 93.0% when combined with mammography for women at high familial risk for breast cancer. However, breast MRI requires the use of gadolinium-based contrast agents, which carry potential risks of retention in the brain that are currently being investigated (Kartamihardja et al. 2016; Murata et al. 2016; Food and Drug Administration [FDA] 2017; McDonald et al. 2017). MBI has been reported to find an additional 7.7 cancers per 1000 women, but there are concerns over breast radiation as well as patient comfort (Shermis et al. 2016;
unnecessary biopsies, Athanasiou et al. (2010) used SSI biopsy of complicated cysts. With the goal of reducing preference to avoiding ultrasound for use as a breast screening modality is its high false-positive rate, sometimes resulting in patients having to undergo unnecessary biopsies (Berg et al. 2016). In the American College of Radiology Imaging Network (ACRIN) 6666 trial, more false positives were seen with increasing breast density in ultrasound (e.g., 10% false positives with <25% density steadily increasing to 14.4% false positives with breast density >80%) (Berg et al. 2016). This trend is indicative of an interaction between dense breast tissue and the effectiveness of ultrasound. When tissue is dense, the speed of sound changes between tissue layers, causing a phenomenon called acoustic clutter (Lediju et al. 2008), resulting in diagnostic uncertainty.

To combat this high false-positive rate and reduce the number of unnecessary biopsies, investigators have utilized ultrasound to gain stiffness information on the tissue in question, including techniques such as acoustic streaming (Nightingale et al. 1999), strain elastography (Garra et al. 1997), acoustic radiation force impulse imaging (Nightingale et al. 2002; Sharma et al. 2004), shear wave elastography (SWE) (Evans et al. 2010; Berg et al. 2012; Han et al. 2019) and supersonic shear wave imaging (SSI) (Bercoff et al. 2004; Athanasiou et al. 2010), which each aim to measure the elastic response or properties of tissue by inducing small displacements. Specifically, for a preliminary study population of 21 indeterminate lesions, Soo et al. (2006) reported 50% sensitivity and 100% specificity for cyst detection with acoustic streaming. In addition, Berg et al. (2012) reported improved specificity of 78.5% using SWE compared with B-mode, with implications for avoiding biopsy of complicated cysts. With the goal of reducing unnecessary biopsies, Athanasiou et al. (2010) used SSI to modify the Breast Imaging Reporting and Data System (BI-RADS) score, demonstrating the potential avoidance of 46% of unnecessary biopsies.

Overall, these elasticity-based methods have been successful at distinguishing solid from fluid-filled lesions; however, they are challenging for cysts with very thick fluid and small masses deep in tissue (Nightingale et al. 1999; Garra 2007). In addition, acoustic radiation force-based and shear wave imaging methods require altered, high-energy transmit pulse sequences that are typically unable to be supported by miniaturized ultrasound systems, and they also depend on loading conditions introduced by the probe pressure required to maintain acoustic coupling and contact (Barr 2012; Barr and Zhang 2012; BaIleyguier et al. 2013), which complicates quantitative inter- and intra-patient longitudinal studies, as well as quantitative comparisons of results from different ultrasound system manufacturers (Bell et al. 2016).

Similar to elastography, quantitative ultrasound aims to discover information about the underlying microstructure of tissue by using spectral-based parameters, such as the backscatter coefficient, effective scatterer diameter and effective acoustic concentration, as well as envelope statistics through the fitting of models, such as the homodyned-K distribution (Oelze and Mamou 2016). When combined, these methods have been successful in characterizing breast lesions (Shankar et al. 2001; Oelze et al. 2007; Nam et al. 2013; Sadeghi-Naini et al. 2013; Trop et al. 2014; Samnachi et al. 2015); however, they require careful calibration using reference phantoms, are computationally complex, and are often model based, which can cause gross estimation errors (Oelze and Mamou 2016).

Rather than inducing external forces or requiring custom calibrations, advanced beamforming methods aimed at reducing acoustic clutter are particularly well suited to improve image quality and offer the promise of improving the diagnostic power of ultrasound, as seen with spatial compounding (Entrekin et al. 1999; Huber et al. 2002). In addition, tissue harmonic imaging (THI), which displays higher-frequency harmonics of the transmitted ultrasound wave, is known to assist with characterization of simple cysts. One reason for this assistance is the removal of reverberation echoes in fluid masses when these echoes have the same frequency content as the transmitted ultrasound wave (Hooley et al. 2013). However, existing literature is mixed on the overall benefit of THI for all patients. For example, Mesurolle et al. (2006) cites difficulty visualizing masses in tissue that is not purely fatty, and Cha et al. (2007) reports shadowing from normal structure in dense breasts, more conspicuous acoustic shadowing and decreased penetration depth with THI.

Short-lag spatial coherence (SLSC) and robust short-lag spatial coherence (R-SLSC) beamforming can both improve ultrasound image quality by displaying the spatial coherence of backscattered pressure waves instead of traditional brightness information. Since its creation as a clutter reduction technique, SLSC has been applied to a variety of in vivo tissue types including thyroid (Lediju et al. 2011), cardiac (Bell et al. 2013), fetal (Kakkad et al. 2015; Long et al. 2018) and liver.
Coherence-based beamforming data acquisition.

In this article, we present the clinical advantages of R-SLSC beamforming applied to breast data and assess its ability to increase diagnostic certainty and reduce the number of unnecessary biopsies.

METHODS

Study population

Twenty-five patients containing 26 masses in total were enrolled in our ongoing study after receiving informed consent and approval from the Johns Hopkins Medicine Institutional Review Board (Protocol No. IRB00127110). Patients ranged from 18–80 y in age, with a mean age of 55 y. Patients were identified as having at least one suspicious mass with diagnostic ultrasound and were scheduled for an ultrasound-guided core-needle biopsy. Patients were imaged prior to the biopsy, and pathology results were used as the ground truth for statistical calculations. In the case of surgical excision after core-needle biopsy, the ground truth pathology was updated to the surgical pathology found at excision. Masses that were successfully aspirated during the biopsy procedure were designated as fluid filled. Solid masses underwent core-needle biopsy and were further separated based on specific pathology.

Data acquisition

Each patient was scanned using an Alpinion ECUBE12R, which is a dual research and clinical ultrasound scanner (Alpinion, Seoul, Korea). This scanner was connected to an Alpinion L8-17 or Alpinion L3-8 ultrasound transducer with a transmit frequency of 12.5 or 8 MHz, respectively. We performed fundamental imaging rather than THI to prioritize higher transmit frequencies within the −6-dB transducer bandwidths (i.e., 8–17 or 3–8 MHz, respectively). Each lesion was imaged using 256 focused transmissions with a focus size that is characteristic of spatial coherence functions at higher lag values (when compared with the errors observed at lower lag values) (Nair et al. 2018). After implementation of a series of simplification steps and a convex relaxation, this RPCA problem statement is formulated as a Lagrangian and solved by minimizing the Lagrangian objective function $L(A, E, Y, p) = \| A \|_* + \frac{\lambda}{2} \| Y - D - A - E \|_F^2$ (Lin et al. 2010):

$$L(A, E, Y, \mu) = \| A \|_* + \frac{\lambda}{2} \| E \|_1 + \langle Y, D - A - E \rangle + \frac{\mu}{2} \| D - A - E \|_F^2$$

where $Y$ is a matrix of Lagrange multipliers, $\lambda$ is a positive scalar, $\lambda$ is the sparsity penalty parameter that is varied to

arrival differences, normalized correlation measurements are calculated between equally spaced elements, or lags, resulting in the normalized spatial correlation

$$\hat{R}(m) = \frac{1}{N - m} \sum_{i=1}^{N-m} \frac{\sum_{n=n_i}^{n_i+m-1} s(n) s(n+m)(n)}{\sqrt{\sum_{n=n_i}^{n_i+m-1} s_i^2(n) \sum_{n=n_i}^{n_i+m-1} s_i^2(n)}}$$

where $N$ is the number of receive elements in the transducer, $m$ is the lag and $s(n)$ is a time-delayed, zero-mean signal received at element $i$ from depth $n$.

This resulting spatial coherence function is then summed up to a specific short-lag value, $M$, yielding the value of the SLSC pixel. $M$ is commonly taken to be a value of $\leq 20$ for in vivo breast images.

$$R_{sl} = \int_1^M \hat{R}(m) dm = \int_1^M R(m)$$

This process is repeated for each lateral and axial position in the image, with an axial correlation kernel of size $k = n_2 - n_1$ (Bell et al. 2015). Throughout this work, a kernel of 1.56 times the wavelength associated with the center frequency of the probe was used.

Robust SLSC beamforming. To take advantage of the sparse, high-frequency information contained at higher lag values, R-SLSC (Nair et al. 2018) denoises the coherence information by using robust principal component analysis (RPCA). To accomplish this goal, lag images generated using the original SLSC algorithm were first vectorized and stacked to form a matrix, $D$. RPCA was then performed on this matrix to solve for $A$ in the expression

$$D = A + E$$

where $A$ is the underlying low-rank matrix that estimates the ground truth, and $E$ is an error matrix that is considered to be sparse with possible high-magnitude errors, which is characteristic of spatial coherence functions at higher lag values (Nair et al. 2018). After implementation of a series of simplification steps and a convex relaxation, this RPCA problem statement is formulated as a Lagrangian and solved by minimizing the Lagrangian objective function $L(A, E, Y, p)$, which is known as the augmented Lagrange multiplier method (Lin et al. 2010):
Contrast Difference

The M and λ values were selected by matching the R-SLSC signal-to-noise ratio (SNR) to the B-mode speckle SNR to present images that are quantitatively more similar. When multiple-parameter combinations were possible, contrast was used as a secondary metric to optimize the selected parameters. These metrics are defined in the next subsection.

Performance metrics

Contrast, SNR and contrast difference were measured and compared across matched B-mode and coherence-based images created with the same channel data according to the equations

$$ \text{Contrast} = 20 \log \left( \frac{S_i}{S_0} \right) $$

$$ \text{SNR} = \frac{S_i}{\sigma_i} $$

$$ \text{Contrast Difference} = \text{Contrast}_{\text{B-mode}} - \text{Contrast}_{\text{R-SLSC}} $$

where $S_i$ and $\sigma_i$ are the mean and standard deviation, respectively, within a region of interest (ROI) inside of the target prior to log-compression, and $S_0$ and $\sigma_0$ are the mean and standard deviation, respectively, of a ROI outside of the target prior to log-compression. The contrast subscripts in eqn (7) indicate the beamforming method used for the contrast measurement.

Reader study classification system

RF data from each patient were saved and processed offline to generate matched B-mode and R-SLSC images for each mass of clinical interest. The masses were then randomized and reviewed independently by five board-certified breast radiologists, including three co-authors of this paper (E.O., K.M. and L.M.). Each reader had between 1 and 22 years of experience beyond residency. Two of the radiologists assisted with image acquisition (Readers 1 and 2). To include readers who had no prior knowledge of the specific patients enrolled in our study and no prior experience with the R-SLSC imaging technology, the remaining three readers (Readers 3, 4 and 5) did not participate in patient recruitment or imaging. Prior to reviewing the images, each radiologist was given a brief tutorial on R-SLSC image interpretation to understand the images presented.

The masses were randomized and presented to the readers under radiology reading room conditions using the graphical user interface illustrated in Figure 1. The readers were instructed to assume that no prior imaging studies were available for each mass, for us to make the assumption that variations in responses were owing solely to observed differences in the presented images. For each patient, the readers were presented with two tasks (i.e., Task 1 followed by Task 2) prior to advancing to the next patient.

For Task 1, only B-mode images were presented, including both the clinical screenshot (which includes proprietary non-linear filters) and the reconstructed B-mode image formed from the saved RF data, as illustrated in Figure 1. Both of these B-mode images were presented to avoid bias from the poorer-quality B-mode images created from saved RF data, which were processed without the non-linear filters that are present in most clinical systems. With only B-mode images, the readers were asked to select the content of the mass—(1) solid, (2) fluid, (3) mixed or (4) uncertain—and the BI-RADS (Mendelson et al. 2001) category for each mass—(2) benign finding, (3) probably benign, (4) suspicious abnormality and (5) highly suggestive of malignancy. Subcategories were not used in this classification.

For Task 2, after category selection using only the B-mode images, the readers were presented with the R-SLSC image for the same patient and asked to categorize the content and BI-RADS category of the presented mass given the original B-mode images and the added R-SLSC display. This selection using all available images is indicated as Task 2 in Figure 1. In addition, for Task 2, a duplex display (Wiacek et al. 2018) was available if desired to further analyze the R-SLSC image (labeled as Optional in Figure 1). If chosen, the reader could select a region in the B-mode image, where the R-SLSC image would be overlaid in color, with blue representing low spatial coherence and red representing high spatial coherence.

For Tasks 1 and 2, the readers had full control over tunable parameters, including dynamic range, short-lag value (i.e., $M$ in eqn [2]) and lambda value (i.e., $\lambda$ in eqn [4]) to aid decision making. The selected classifications for each mass during each task were saved for comparison and statistical analyses.

Statistical analysis

The difference between the mean contrast of 10 frames of B-mode images and 10 frames of R-SLSC images was calculated for each fluid-filled or solid mass using eqn (7). Differences between fluid and solid masses were compared using a repeated-measures
analysis of variance to determine statistical significance (p < 0.05). In addition, the reader responses for both the BI-RADS category and content assessments were compared before and after introduction of the R-SLSC image using Cohen’s $\kappa$ to determine the similarity of each assessment to the truth (Cardillo 2007a). Fleiss’ $\kappa$ for multiple observers was calculated to determine the level of inter-observer agreement among the content and BI-RADS classifications (Cardillo 2007b). Both Cohen’s and Fleiss’ $\kappa$ values <0.0 were interpreted as poor agreement; values between 0 and 0.20 were interpreted as slight agreement; values between 0.21 and 0.40 were interpreted as fair agreement; values between 0.41 and 0.60 were interpreted as moderate agreement; values between 0.61 and 0.80 were interpreted as substantial agreement; and values >0.81 were interpreted as perfect agreement, in accordance with Landis and Koch (1977).

Sensitivity and specificity for correctly diagnosing cysts were measured as follows:

Sensitivity = \[
\frac{TP}{TP + FN}
\]  \hspace{1cm} (8)

Specificity = \[
\frac{TN}{TN + FP}
\]  \hspace{1cm} (9)

Here a true positive (TP) was defined as a cyst that was not recommended for biopsy in Task 2 of the reader study (i.e., BI-RADS 2 or 3), a false negative (FN) was defined as a cyst that was recommended for biopsy in Task 2 (i.e., BI-RADS 4 or 5), a true negative (TN) was defined as a solid mass that was recommended for biopsy in Task 2 (i.e., BI-RADS 4 or 5) and a false positive (FP) was defined as a solid mass that was not recommended for biopsy in Task 2 (i.e., BI-RADS 2 or 3). The masses with a ground truth classified as mixed were omitted from this analysis, considering that they technically contain a mixture of TP and TN.

We compared the sensitivity and specificity of our results with those of results from two of the few breast studies with data available and with a similar goal of distinguishing cysts from solid masses (Soo et al. 2006) or evaluating cyst detection (Chiorean et al. 2012). To compare our results with the acoustic streaming results reported by Soo et al. (2006), TP was defined as a cyst in which acoustic streaming was generated, FN was defined as a cyst with no streaming detected, TP was defined as a solid mass with no streaming detected, and FN was defined as a solid mass with streaming. This analysis was performed for the study and control groups combined, resulting in 39 masses in total.

To compare with the performance of B-mode alone from the same study (Soo et al. 2006), any cyst in the study group was considered a FP and any solid mass in this group was considered a TN (considering that the study group consisted of patients with discrete breast lesions that were considered indeterminate for a cyst or solid mass with initial diagnostic sonography). Otherwise, TP for B-mode alone was defined as the control group cysts, and TN was defined as the control group solid masses.
To compare our results with the elastography results reported by Chiorean et al. (2012), TP was defined as all simple and complicated cysts that were downgraded after elastography (i.e., a final diagnosis of BI-RADS 2), and FN was defined as complicated cysts that were not altered in diagnosis as a result of elastography. Because no solid masses were included in this study, there were no TNs or FPs, and therefore, specificity could not be calculated. The results for complex masses (i.e., masses with mixed content according to our definition) were not included in this analysis, and only the results for simple and complicated cysts were included, resulting in a total of 92 masses.

To compare with the performance of B-mode alone from the same study (Chiorean et al. 2012), any simple cyst was considered a TP and any complicated cyst was considered a FN. Because no solid masses were included in this study, there were no TNs or FPs, and therefore, specificity could not be calculated.

RESULTS

Contrast difference across multiple masses

Our previous work (Wiacek et al. 2019) revealed improved contrast in R-SLSC images compared with B-mode images for fluid-filled regions such as simple and complicated cysts. Conversely, we also found that hypoechoic solid masses contain low-amplitude internal echoes caused by coherent scatterers, which results in coherent signal in R-SLSC images. The difference in coherence within the fluid-filled and solid masses allows them to appear distinct from one another.

Extending this analysis from 9 to 26 in vivo masses, Figure 2 illustrates the contrast difference between the respective B-mode image and R-SLSC image for all masses included in this study. In blue are the fluid-filled masses, which display a positive contrast difference, indicating an improvement in contrast. In gray are the solid masses, which display a negative contrast difference, indicating coherent content, and therefore confirming the presence of a solid mass. This difference in contrast differences between fluid-filled and solid masses is statistically significant ($p < 0.001$). Specific examples of masses are available in Figure 2, and contrast results are summarized in Table 1. For any masses with mixed solid and fluid content in this study, the ROI for contrast measurements was selected as the solid component of the mass. Therefore, the masses in Table 1 are categorized as fluid-filled, benign and malignant, where benign encompasses both benign solid and benign mixed masses.

Fluid-filled masses

In Figure 3 (a, f, k) is an example of a simple cyst. Figure 3a is the clinical screenshot representing the additional information that was accessible to the radiologists when making decisions during the reader study. Figure 3f was beamformed using standard delay-and-sum beamforming from the saved RF channel data, which were
used for quantitative comparisons. Figure 3k was generated using R-SLSC beamforming where the $M$ and $\lambda$ values were selected to match the B-mode speckle SNR. Qualitatively, the R-SLSC image shows improved contrast and reduced acoustic clutter compared with the B-mode image created from raw data. The heterogeneous tissue texture of the B-mode image is also largely removed in the corresponding R-SLSC image. In addition, the R-SLSC image has a dark-field artifact present because of the use of focused transmits, which may be removed with synthetic transmit aperture imaging (Bottenus et al. 2013). Quantitatively, the contrast measurements match the qualitative observations. The contrast in the B-mode image in Figure 3f is $-19.8$ dB. For the same ROI, the contrast in the R-SLSC image in Figure 3k is improved to $-22.0$ dB. The average contrast improvement is $3$ dB across all simple cysts in the study population.

In Figure 3 (b, g, l) is an example of a complicated cyst. Figure 3b is the clinical screenshot for qualitative comparisons. Figure 3g is the B-mode image generated with the saved RF channel data for quantitative comparisons. Figure 3l was generated from the same raw data using R-SLSC beamforming. Similar to the simple cysts, the R-SLSC image shows improved contrast compared with its B-mode counterpart; however, the R-SLSC

### Table 1. Summary of the 26 masses included in the reader study

<table>
<thead>
<tr>
<th>Mass type</th>
<th>No. of masses</th>
<th>Mean contrast (dB)</th>
<th>B-mode</th>
<th>R-SLSC</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complicated cyst</td>
<td>4</td>
<td>$-13.56$</td>
<td>$-16.40$</td>
<td>2.85</td>
<td></td>
</tr>
<tr>
<td>Simple cyst</td>
<td>3</td>
<td>$-20.69$</td>
<td>$-23.79$</td>
<td>3.09</td>
<td></td>
</tr>
<tr>
<td>Benign</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apocrine metaplasia</td>
<td>3</td>
<td>$-8.52$</td>
<td>$-4.31$</td>
<td>$-4.21$</td>
<td></td>
</tr>
<tr>
<td>Fibroadenolipoma</td>
<td>1</td>
<td>$-2.23$</td>
<td>$0.44$</td>
<td>$-2.67$</td>
<td></td>
</tr>
<tr>
<td>Fibroadenoma</td>
<td>4</td>
<td>$-9.79$</td>
<td>$-1.52$</td>
<td>$-8.27$</td>
<td></td>
</tr>
<tr>
<td>Fibrocystic change</td>
<td>2</td>
<td>$-11.64$</td>
<td>$-4.93$</td>
<td>$-6.71$</td>
<td></td>
</tr>
<tr>
<td>Fibroepithelial lesion</td>
<td>1</td>
<td>$-3.01$</td>
<td>$8.43$</td>
<td>$11.43$</td>
<td></td>
</tr>
<tr>
<td>Malignant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ductal carcinoma in situ</td>
<td>3</td>
<td>$-8.92$</td>
<td>$-1.43$</td>
<td>$-7.49$</td>
<td></td>
</tr>
<tr>
<td>Invasive ductal carcinoma</td>
<td>4</td>
<td>$-14.56$</td>
<td>$-2.24$</td>
<td>$12.32$</td>
<td></td>
</tr>
</tbody>
</table>

BI-RADS = Breast Imaging Reporting and Data System; R-SLSC = robust short-lag spatial coherence.

Fig. 3. (a–e) Clinical screenshots. (f–j) B-Mode images created from raw data. (k–o) Corresponding robust short-lag spatial coherence images of five different masses in the female breast. These masses appear hypo-echoic in the B-mode images, and the solid or fluid content is more readily distinguishable in the robust short-lag spatial coherence images.
image shows some scattered internal coherence representative of proteinaceous material characteristic of complicated cysts. The contrast in the B-mode and R-SLSC images is $-21.8$ and $-24.1$ dB, respectively, reflecting improved contrast in the R-SLSC image. Overall, despite the minimal coherence within the mass in the R-SLSC image, the average contrast improvement is 2.8 dB across all complicated cysts in the study population. These results are summarized in Table 1.

Masses with solid content

In Figure 3 (c, h, m) is an example of a benign solid mass, specifically a fibroadenoma. Both Figure 3c, the clinical screenshot, and Figure 3h, the B-mode image generated using saved RF channel data, show the mass appearing as hypo-echoic with some internal echoes. Figure 3m, the corresponding R-SLSC image, appears to have isocoherent structures and begins to blend in with the background. This coherence is present because of the solid content of the mass. The contrast in the B-mode image created from raw data is $-11.7$ dB, representative of its hypo-echoic appearance. The R-SLSC image shows significantly worse contrast at $-3.6$ dB, which indicates the presence of a solid mass. The average contrast difference between B-mode images and R-SLSC images for the fibroadenomas in this study was $-8.3$ dB. These results and the results for other types of solid masses in this patient population are summarized in Table 1.

In Figure 3 (d, i, n) is an example of a malignant solid mass, specifically an invasive ductal carcinoma. Similar to the benign solid mass, In Figure 3 (d, i), this malignant solid mass appears largely hypo-echoic in the B-mode image. Figure 3n is the respective R-SLSC image which begins to blend in with the background because of the internal coherence, indicative of the solid content. The contrast for the B-mode image is $-7.1$ dB. Because of the coherent tissue content, the R-SLSC image has significantly worse contrast at $-1.5$ dB, indicating the presence of solid tissue. These results and the results for the other malignant masses are summarized in Table 1.

Finally, in Figure 3 (e, j, o) is an example of a benign mixed mass, that is, a mass containing both solid and fluid tissue components, with biopsy pathology revealing fibrocystic changes. As illustrated in Figure 3 (e, j), this mass appears hypo-echoic, with some internal echoes in the B-mode image. Figure 3o is the R-SLSC image, which appears to have some coherent components in the middle of the mass, with less coherence on the right and left sides of the mass, indicative of both solid and fluid components. This mass is considered to have “mixed” content. The contrast of the B-mode image is $-13.6$ dB, which is reduced to $-5.3$ dB in the R-SLSC image using an ROI containing the solid components of the mass.

Reader study results

The results of the reader content classification are summarized in Figure 4. The leftmost bar of each grouping shows the ground truth for each type of mass (i.e., simple cyst, complicated cyst, benign solid, malignant solid and benign mixed). This ground truth is based on pathology results. The results from each reader appear to the right of the ground truth, and the left bar in each result pair indicates the reader decision using only B-mode images (i.e., Task 1), while the right bar of each result pair indicates the reader decision when combining both B-mode and R-SLSC for decision making (i.e., Task 2). Each color represents the percentage of the masses classified in each category.

The uncertainty surrounding the content of simple cysts ranged from 33%–100% with B-mode alone (specifically, 66.7%, 100%, 66.7%, 33.3% and 33.3% for Readers 1, 2, 3, 4 and 5, respectively). With R-SLSC included in the decision making, the uncertainty was removed for Readers 1–4, and all masses were correctly classified as fluid. However, Reader 5 remained uncertain about 33.3% of these masses.

For the complicated cysts, when using only B-mode images, Readers 1–4 were uncertain or misclassified 50% of the masses, and Reader 5 was uncertain about 25% of the masses. With the inclusion of R-SLSC, the uncertainty was reduced for Readers 1 and 3, misclassifications were removed for Reader 2 and classifications remained the same for Readers 4 and 5. Two of the complicated cysts were classified as mixed, and possible reasons for this are presented in the Discussion.

For the benign solid masses when using only B-mode imaging, Readers 1, 2 and 4 were uncertain or misclassified 33.3% of masses. Readers 3 and 5 were uncertain about the content of 66.7% and 55.6% of these masses, respectively. With the inclusion of R-SLSC, the uncertainty or misclassification was reduced to 11.1% for Readers 1, 2 and 3, and reduced to 0% for Reader 4. Reader 5 remained the same with 55.6% uncertainty.

For the malignant solid and benign mixed masses, there was one case where uncertainty was reduced from 75% to 0% (Reader 3, benign mixed) and two cases where uncertainty increased with R-SLSC (Reader 3, malignant solid, and Reader 5, benign mixed). Otherwise, the classifications when using only B-mode and when adding R-SLSC remained approximately similar, with some discrepancies in the definition of mixed content.

To assess possible connections between the quantitative metrics (which are reported in Fig. 2) and the uncertain or misclassified masses from the reader study (which are reported in Fig. 4), Figure 5 illustrates the contrast difference of each mass as a function of the number of readers uncertain about the content of that mass (top) and the number of misclassifications for that
mass (bottom). For each mass, the filled circle represents the B-mode observations, and the open circle represents the R-SLSC observations. A left arrow indicates a decrease in uncertainty or misclassifications with the introduction of R-SLSC, and a right arrow indicates an increase in uncertainty or misclassifications with the introduction of R-SLSC. A data point without an arrow indicates no change with the introduction of R-SLSC.

Figure 5 illustrates an overall reduction in uncertainty and misclassifications with the introduction of R-SLSC, regardless of the magnitude of the contrast difference. In particular, the summary of uncertainty for fluid-filled masses (reported in the top left of Fig. 5) reveals that two or three readers were uncertain with B-mode alone, and in each of these cases, the uncertainty was minimized to no to two readers with the introduction of R-SLSC (with the exception of one complicated cyst that caused no changes in uncertainty by two readers). For the solid masses, overall the majority of masses decreased in uncertainty with R-SLSC, with the exception of two malignant solid masses moving from no readers being uncertain with B-mode to one reader being uncertain with R-SLSC included. The uncertainty in these two cases was associated with the same reader (i.e., Reader 3), and this uncertainty did not change the reader’s recommendation for biopsy (i.e., BI-RADS 4) during Task 2. For the mixed masses, the uncertainty of the readers either remained the same or decreased.

The summary of misclassifications of fluid-filled masses (reported in the bottom left of Figure 5) reveals that one complicated cyst resulted in two fewer misclassifications (from solid to the correct classification of...
fluid) with the inclusion of R-SLSC. In addition, one complicated cyst resulted in two misclassifications as mixed with the inclusion of R-SLSC. There were no additional misclassifications nor changes in misclassifications for these fluid masses. Similarly, for the solid masses, the number of misclassifications for the majority of these masses either remained the same or decreased, with the exception of one solid mass that was classified as mixed with the introduction of R-SLSC. For the mixed masses, the number of misclassifications remained the same regardless of the imaging method.

The results of the reader BI-RADS classification for the simple cysts are summarized in the top of Figure 6. The first pair of bars show the ideal diagnosis on the left (based on the true pathology, which is BI-RADS 2, meaning the mass will neither be followed nor biopsied), and the actual diagnosis listed in the patients’ medical records is shown on the right. The following five pairs of results indicate the rating of each reader. The left bar in each pair indicates the reader decision using only B-mode images (i.e., Task 1), and the right bar indicates the reader decision when both B-mode and R-SLSC are available (i.e., Task 2).

When B-mode alone was used, the top of Figure 6 illustrates that Readers 1, 3, 4 and 5 assigned BI-RADS 4 to 33%–100% of the simple cysts, indicating the need for a biopsy in these cases. With the inclusion of R-SLSC, 0% of the masses required biopsy for Readers 1–4. In addition, Reader 1 decided to follow 33.3% of the masses, and the diagnosis provided by Reader 5 remained the same (i.e., biopsy 33.3% of these masses).

The remaining results of the reader BI-RADS classification are also summarized in Figure 6 in the same format as for the simple cysts. For the solid and mixed mass types (i.e., benign solid, malignant solid and benign mixed), the ideal diagnosis is a combination of two categories depending on additional imaging findings, and therefore, two colors are displayed in the bar representing the ideal choice for these two masses; either choice is considered correct.

For the complicated cysts, using B-mode alone, Reader 1 required biopsy for 75% of the masses (i.e., BI-RADS 4), and Readers 2 and 3 required biopsy for 50% of the masses. Reader 5 recommended biopsy for 25% of the masses, and Reader 4 did not recommend biopsy for any of the masses. When R-SLSC was included, the number of biopsies was reduced for Readers 1, 2 and 3, requiring biopsies for 50%, 0% and 25% of the masses, respectively. The recommendations of Readers 4 and 5 remained the same for these complicated cysts.

For the masses with solid content (i.e., the benign, malignant and mixed categories), the overall diagnostic decision remained similar when using B-mode only and when including R-SLSC for each reader, with the exception of two cases in which mixed diagnoses of BI-RADS 3 and 4 were modified to all BI-RADS 4 with the inclusion of R-SLSC (Reader 4, benign solid, and Reader 3, benign mixed).
Figure 7 illustrates the results of the statistical analyses. For each reader, Cohen’s kappa ($\kappa$) was calculated between the truth (or ideal) diagnosis and either the B-mode selections or the selections when R-SLSC was added. When classifying the content of the mass, the $\kappa$ value with B-mode only was 0.42, 0.50, 0.25, 0.48 and 0.43 for Readers 1, 2, 3, 4 and 5, respectively, as reported in Figure 7a. These results indicate fair agreement for Reader 3 and moderate agreement for the remaining readers. When R-SLSC was added, the $\kappa$ value improved to the range 0.65—0.80 for Readers 1—4, indicating substantial agreement with the truth. For Reader 5, the $\kappa$ value was reduced to 0.38 when R-SLSC was added, indicating fair agreement with truth.

When deciding the clinical path through BI-RADS, the $\kappa$ values with B-mode alone were 0.17, 0.36, 0.28, 0.53 and 0.57 for Readers 1, 2, 3, 4 and 5, respectively, as reported in Figure 7b. These results indicate slight, fair and moderate agreement between the ideal diagnosis and the selections made with only B-mode image information by Readers 1, 2/3 and 4/5, respectively. When R-SLSC was included in the decision process, the $\kappa$ values improved to the range 0.65—0.80 for Readers 1—4, indicating substantial agreement between the ideal diagnosis and the selections made with R-SLSC included in the diagnostic decision. In comparison, the $\kappa$ value remained the same for Reader 5 (i.e., 0.57, which represents moderate agreement). In calculating these Cohen’s
values, BI-RADS 4 was used as the ideal diagnosis for any mass containing solid content (i.e., benign solid, malignant solid and benign mixed), as this is the standard diagnosis for baseline imaging at our clinic.

Finally, the inter-reader agreement was evaluated using Fleiss’ $\kappa$. Figure 7c summarizes these results. For content classification, the $\kappa$ value when using only B-mode was 0.33, indicating fair agreement; however, when R-SLSC is included, the $\kappa$ value increases to 0.44, indicating moderate agreement. Similarly, for the BI-RADS classification, the $\kappa$ value when using only B-mode was 0.30 and increased to 0.50 when R-SLSC was included, indicating fair and moderate agreement, respectively.

Table 2 outlines the sensitivity and specificity of correctly diagnosing cysts for each reader. Sensitivity was improved when using R-SLSC combined with B-mode, compared with B-mode alone, for each reader except Reader 5. The specificity was improved when using R-SLSC combined with B-mode, compared with B-mode alone, for Readers 3 and 4. Otherwise, specificity remained the same for Readers 1, 2 and 5. The means ± standard deviation of sensitivity and specificity results for the five readers were 57% ± 29% and 90% ± 7%, respectively, with B-mode alone, and 86% ± 14% and 95% ± 5%, respectively, when using R-SLSC combined with B-mode, representing an overall increase in sensitivity and specificity with R-SLSC included.

When compared with the acoustic streaming results, for which the sensitivity was 61% (Soo et al. 2006), inclusion of R-SLSC resulted in greater sensitivity for each reader (and greater mean sensitivity for all readers). Similarly, when compared with the free-hand elastography sensitivity of 87% (Chiorean et al. 2012), the inclusion of R-SLSC had a similar mean sensitivity, with two readers achieving a sensitivity higher than that for free-hand elastography (i.e., Readers 2 and 4). The mean specificity with R-SLSC included was 5% lower than the comparative specificity of 100% for acoustic streaming (Soo et al. 2006), with two readers achieving a specificity of 100% (i.e., Readers 1 and 4).

**DISCUSSION**

**Clinical implications**

The work described in this article indicates that the inclusion of R-SLSC images in clinical decision making has promising potential to improve both the certainty of mass contents and the BI-RADS classification accuracy, particularly for fluid-filled simple cysts, as illustrated in Figures 4, 5 and 6. Fluid-filled regions that appear hypoechoic typically contain low-amplitude, incoherent echoes, which can be attributed to the presence of acoustic clutter. Because R-SLSC images display spatial coherence rather than traditional amplitude information, the signal caused by acoustic clutter is reduced with R-SLSC imaging, resulting in the positive contrast differences reported in Figure 2 ($p < 0.001$). This difference is suspected to be the reason...
for the improved certainty and the reduced need for biopsy or follow-up of the simple cysts; the top left plot of Figure 5 indicates that these benefits are achieved for contrast improvements as minimal as 1.2 dB.

The improved contrast in the simple cysts allowed readers to be more confident in a fluid-filled diagnosis, resulting in no uncertainty and no recommendations for biopsies from four of the five readers. We also emphasize here that the recruited patients were recommended for biopsy. Thus, although our study population included a total of three simple cysts, these simple cysts are expected to represent a few of the most challenging simple cyst cases, which is also demonstrated by a minimum of three readers being uncertain of the content of each cyst in B-mode images (as observed in the top left plot of Figure 5). It is notable that R-SLSC reduced uncertainty in these challenging cases.

Complicated cysts can be difficult to diagnose because of the presence of both acoustic clutter and proteinaceous material inside of the mass, causing confusion with hypoechoic solid masses that require biopsy (Meyer et al. 1992). On examination of the R-SLSC image of a complicated cyst in Figure 3l, the fluid-filled content is evident by the black signal within the mass; however, the proteinaceous material can also be visualized within the mass. Therefore, we suspect that the complicated cysts classified as mixed by Readers 1 and 3 occurred because the content within complicated cysts was enhanced with R-SLSC imaging and because more guidelines regarding the definition of mixed content were needed. Otherwise, the majority of readers were more confident in the classification of complicated cysts in the fluid category, and assigned complicated cysts a less invasive diagnosis with R-SLSC included in the decision process.

When the percentages of uncertain and biopsy-recommended masses for each reader are considered, the added information from R-SLSC images reduced the mean uncertainty about the content of fluid-filled masses (i.e., both simple and complicated cysts) from 47.5% to 15.8%, and reduced the mean percentage of fluid-filled masses recommended for biopsy from 43.3% to 13.3%. Decreasing the number of biopsies of fluid-filled masses helps to reduce patient discomfort and anxiety at the time of initial diagnosis, as well as save resources in the health care system. These results could be increasingly important in remote areas where the resources used to perform ultrasound-guided biopsies are scarce.

The content selection and BI-RADS diagnosis provided by Reader 5 remained largely the same throughout the reader study. This reader cited difficulty in trusting the R-SLSC images without additional data or patient history, and as a result, this reader chose to keep most of the clinical decisions the same regardless of the imaging method. Therefore, our assumption that responses were based solely on observed differences in the presented images did not seem to hold true for Reader 5 because of these factors that we did not anticipate and were unable to control. Incorporation of this feedback will be the focus of future work (e.g., including readers in the patient recruitment and diagnostic retrieval process to gain more confidence and familiarity with the new technology prior to performing additional reader studies). When Reader 5 is removed from the inter-reader assessment, and only the agreement among readers who were more confident with the novel image presentation format is considered, the inter-reader agreement improved from κ = 0.44 to κ = 0.60 for the content classification; both values represent moderate agreement nonetheless.

Finally, although the reader study indicates that the masses with solid content (i.e., benign solid, malignant solid and benign mixed) would follow similar clinical paths whether B-mode alone or B-mode paired with R-SLSC was used, the confidence in deciding the content of each mass was improved. In particular, when R-SLSC was used, the errors and uncertainty in classification were largely removed for the majority of readers, as illustrated in Figure 5.

Study limitations

The study described in this article focused on 26 masses in 25 different patients. While it is promising that statistically significant differences were seen between B-mode and R-SLSC images of solid and fluid-filled masses, larger-scale studies with increased patient numbers would be beneficial to determine the full clinical impact of R-SLSC. Despite the small number of masses, the comparison of results combining B-mode and R-SLSC with results from similar early-phase studies (Soo et al. 2006; Chiorean et al. 2012) indicates improved sensitivity for all readers when compared with acoustic streaming results (Soo et al. 2006), improved sensitivity for two readers when compared with elastography results (Chiorean et al. 2012) and specificity within one standard deviation of the mean specificity achieved by all readers, as reported in Table 2.

In addition, two of our five readers (i.e., Readers 1 and 2) assisted with the acquisition of the ultrasound data. Although this could be considered a potential bias in the study design, a minimum of 154–228 days elapsed between the last imaged patient and the reader study, these two readers were not involved in the preparation of the reader study nor in analysis of the reader study results, and they both stated that they did not remember the diagnosis of the masses in question. In addition, the results from the additional three readers who did not participate in data collection were compared with the results from the remaining two readers (see Fig. 7c), resulting in moderate agreement among the five readers. This agreement indicates that potential bias was not a reason for the improvements obtained with R-SLSC images included.
Additional limitations include the offline image formation and the static presentation of the masses to each reader. In addition, readers cited less confidence in their clinical decisions without patient medical history, other imaging findings and the ability of real-time imaging capabilities. Therefore, future studies will incorporate a summary of the patient medical history and imaging findings. Given our initial success with offline processing, future work will additionally focus on implementing real-time coherence-based imaging capabilities (Hyun et al. 2017) on our dual clinical and research ultrasound system to make larger-scale studies more feasible.

**CONCLUSIONS**

This work is the first to explore the effect of coherence-based beamforming in breast ultrasound to inform clinical decision making, highlighting the potential of this novel technique to reduce the number of unnecessary biopsies of breast masses. Solid breast masses produce spatial coherence images that appear distinctly different from those of fluid-filled masses. The information from R-SLSC images reduced the uncertainty of fluid mass content from 47.5% to 15.8%, and the number of fluid-filled masses recommended for biopsy was reduced from 43.3% to 13.3%. This benefit was achieved without affecting the clinical decisions for the solid masses, which remained approximately the same with and without the R-SLSC image information. These results provide promising support for additional investigations into the potential inclusion of R-SLSC in the breast clinic as a supplement to standard B-mode images to improve diagnostic certainty and reduce the number of unnecessary biopsies of fluid-filled masses.

**Acknowledgments**—This research was supported by a Ralph E. Powe Junior Faculty Enhancement Award from Oak Ridge Associated Universities (awarded to M.A.L.B.). The authors thank breast radiologists Emily Ambinder, M.D., and Philip Di Carlo, M.D., for participating in our reader study.

**Conflict of interest disclosure**—All authors declare no competing interests.

**REFERENCES**


