

Cross-Imaging Platform Comparison of Ultrasonic Backscatter Coefficient Measurements of Live Rat Tumors

Lauren A. Wirtzfeld, PhD, Goutam Ghoshal, PhD, Zachary T. Hafez, MS, Kibo Nam, MS, Yassin Labyed, MS, Janelle J. Anderson, MS, Maria-Teresa Herd, PhD, Alexander Haak, MS, Zhi He, MS, Rita J. Miller, DVM, Sandhya Sarwate, MD, Douglas G. Simpson, PhD, James A. Zagzebski, PhD, Timothy A. Bigelow, PhD, Michael L. Oelze, PhD, Timothy J. Hall, PhD, William D. O'Brien, Jr, PhD

Abbreviations

ANOVA, analysis of variance; BSC, backscatter coefficient; QUS, quantitative ultrasound; RMS, root mean squared; ROI, region of interest

Received December 16, 2009, from the Departments of Electrical and Computer Engineering (L.A.W., G.G., Z.T.H., A.H., R.J.M., S.S., M.L.O., W.D.O.) and Statistics (Z.H., D.G.S.), University of Illinois at Urbana-Champaign, Urbana, Illinois USA; Department of Medical Physics, University of Wisconsin, Madison, Wisconsin USA (K.N., J.J.A., M.-T.H., J.A.Z., T.J.H.); and Department of Electrical and Computer Engineering, Iowa State University, Ames, Iowa USA (Y.L., T.A.B.). Revision requested January 19, 2010. Revised manuscript accepted for publication March 2, 2010.

The University of Illinois at Urbana-Champaign thanks Mike King and David Hruska for their work on the analysis code, Ellora Sen-Gupta for analysis assistance, and Michael Tu for imaging work. This work was supported by National Institutes of Health grant R01CA111289.

Address correspondence to William D. O'Brien, Jr, PhD, Bioacoustics Research Laboratory, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 405 N Mathews, Urbana, IL 61801 USA.

E-mail: wdo@uiuc.edu

Objective. To translate quantitative ultrasound (QUS) from the laboratory into the clinic, it is necessary to demonstrate that the measurements are platform independent. Because the backscatter coefficient (BSC) is the fundamental estimate from which additional QUS estimates are calculated, agreement between BSC results using different systems must be demonstrated. This study was an intercomparison of BSCs from in vivo spontaneous rat mammary tumors acquired by different groups using 3 clinical array systems and a single-element laboratory scanner system. **Methods.** Radio frequency data spanning the 1- to 14-MHz frequency range were acquired in 3 dimensions from all animals using each system. Each group processed their radio frequency data independently, and the resulting BSCs were compared. The rat tumors were diagnosed as either carcinoma or fibroadenoma. **Results.** Carcinoma BSC results exhibited small variations between the multiple slices acquired with each transducer, with similar slopes of BSC versus frequency for all systems. Somewhat larger variations were observed in fibroadenomas, although BSC variations between slices of the same tumor were of comparable magnitude to variations between transducers and systems. The root mean squared (RMS) errors between different transducers and imaging platforms were highly variable. The lowest RMS errors were observed for the fibroadenomas between 4 and 5 MHz, with an average RMS error of $4 \times 10^{-5} \text{ cm}^{-1}\text{Sr}^{-1}$ and an average BSC value of $7.1 \times 10^{-4} \text{ cm}^{-1}\text{Sr}^{-1}$, or approximately 5% error. The highest errors were observed for the carcinoma between 7 and 8 MHz, with an RMS error of $1.1 \times 10^{-1} \text{ cm}^{-1}\text{Sr}^{-1}$ and an average BSC value of $3.5 \times 10^{-2} \text{ cm}^{-1}\text{Sr}^{-1}$, or approximately 300% error. **Conclusions.** This technical advance shows the potential for QUS technology to function with different imaging platforms. **Key words:** backscatter coefficient; quantitative ultrasound; spontaneous mammary tumors.

Quantitative ultrasound (QUS) has the potential to provide additional information of clinical relevance for diagnosing disease. Much QUS analysis depends on the accurate estimation of backscatter coefficients (BSCs). For QUS to be widely used and aid in diagnosis, it is imperative that BSC estimates be reproducible across different transducers and different equipment. Previous studies have eval-

uated physical phantoms with known properties to demonstrate agreement between laboratories.¹ However, in vivo animal studies in which the tissue properties are unknown are a required next step toward clinical applications. This study examined in vivo spontaneous rat mammary tumors using 3 clinical ultrasound systems and a laboratory ultrasound system with a total of 9 different transducers. Each participating laboratory analyzed the data from their ultrasound system independently, and the BSC data were compared.

Materials and Methods

Animal Model

Six Sprague Dawley rats (Harlan Laboratories, Inc, Indianapolis, IN) with spontaneous mammary tumors were imaged. Most of the tumors were fibroadenomas and one was a carcinoma. Because these tumors were spontaneous, the location, rate of growth, and histologic category varied from animal to animal. The experimental protocol was approved by the Institutional Animal Care and Use Committee of the University of Illinois at Urbana-Champaign and satisfied all university and National Institutes of Health rules for the humane use of laboratory animals.

Ultrasound Systems

Three clinical systems and a single-element laboratory system each were used to image the same rat tumors. The 3 clinical systems were an Ultrasonix RP (Ultrasonix Medical Corporation, Richmond, British Columbia, Canada) with L9-4/38 and L14-5/38 linear arrays with nominal center frequencies of 5 and 7.5 MHz, respectively; a Zonare z.one scan engine diagnostic system (Zonare Medical Systems, Inc, Mountain View, CA) with L8-3 and L14-5sp linear arrays with nominal center frequencies of 7 and 10 MHz, respectively; and an Acuson S2000 (Siemens Medical Solutions USA, Inc, Malvern, PA) with a 4V1 phased array operated as a linear array and a 9L4 linear array with nominal center frequencies of 4 and 9 MHz, respectively. Three spherically focused transducers with nominal center frequencies of 3.5, 7.5, and 13 MHz were used with the laboratory system.

Imaging Methods

The study was performed at the University of Illinois at Urbana-Champaign to enable sequential data collection from each system. In an experiment, the rat was anesthetized with 87 mg/kg of ketamine hydrochloride and 13 mg/kg of xylazine intraperitoneally. The hair over the tumor was shaved and depilated. A transverse line was drawn on the skin over the tumor to mark the area to be imaged. The rat was mounted in a custom holder and placed in a tank of room temperature ($\approx 22^\circ\text{C}$) degassed water. The temperature of the water, being below physiologic temperatures, would tend to decrease overall blood flow. Data were acquired from each scanner without moving the rat relative to the holder. Depending on the tumor size, 5 to 9 parallel planes of data were acquired with each system. For the 3 clinical systems, reference scans were acquired from well-characterized reference phantoms that each group was familiar with for processing their data. For the Ultrasonix system, a phantom with 41- μm glass beads imbedded in a homogenous background of agar with an attenuation slope of 0.19 dB/cm-MHz was imaged. For the Siemens and Zonare systems, a phantom with 7- to 20- μm glass beads imbedded in a homogenous background of microscopic oil droplets in gelatin with an attenuation slope of 0.7 dB/cm-MHz was imaged.² For the laboratory system, a reference scan was acquired from a flat acrylic surface.

After imaging, the animals were euthanized, and the tumors were excised and sent to pathology for diagnosis. The total scanning time for an individual tumor with all 4 systems was approximately 4 hours.

Data Analysis

For each system, the BSCs as a function of frequency were computed for each transducer using methods developed by the respective research group. The tumor was outlined within each 2-dimensional slice and divided into regions of interest (ROIs) to be analyzed. The ROIs were $15 \times 15 \lambda$ for the laboratory and Ultrasonix systems,³ $4 \times 4 \text{ mm}$ for the Siemens system,^{4,5} and $10 \lambda \times 50$ to 100 scan lines (corresponding to 3–4 mm) for the Zonare system.^{3,6} The BSCs were calculated for each ROI and averaged across all ROIs in the slice.

The attenuation slope was estimated from the backscattered data using 3 of the systems (no estimates were made based on the Ultrasonix data) to aid in compensating for attenuation-dependent analyses. Estimates from the Zonare and laboratory systems were based on the techniques presented by Yao et al⁶ and Haak et al,⁷ respectively. Estimates from the Siemens system were made using 3 different techniques: a reference phantom method,⁸ a classic frequency shift method,⁹ and a hybrid method.¹⁰

A functional analysis of variance (ANOVA) was used to compare the frequency-dependent BSC curves from different transducers. The method was based on articles by Shen and Faraway¹¹ and Cuevas et al¹² for data in which the unit of observation is a function observed over a range of input values (frequencies) using both a large sample approximation and a bootstrap method¹³ to evaluate significance. The method required the BSC functional responses to be analyzed over the same ranges of frequencies; therefore, BSC data were subdivided into frequency subregions determined by the bandwidths of the transducers to allow all transducers to be included in the analysis in at least 1 subregion. It was necessary to run the analysis on several regions as there was no frequency range that was included in all the transducer bandwidths. $P < .05$ was considered significant.

As the slope of the BSC versus frequency data corresponds to scatterer size, the slope was estimated by fitting a line to the data from each acquired plane over the same frequency bands that were used for the functional ANOVA tests. One-way ANOVAs were performed in each frequency region for each tumor to determine if there were statistically significant differences in slope.

A root mean squared (RMS) error was also calculated to provide an estimate of the difference in values between the BSC magnitude between different systems and transducers. To calculate the RMS error, the BSC for each transducer was averaged over all slices. Within the subregion analyzed, corresponding to the same regions used in the functional ANOVA, data from all transducers were interpolated to give data at the same frequencies for all data sets. The RMS error was calculated between all pairs of transducers because there was no true value to compare against, and the average of the errors from each pair of transducers was calculated.

Results

The 6 tumors were histologically confirmed to be 5 fibroadenomas and 1 carcinoma. The anatomic locations of the tumors varied, with 3 by the right fore leg, 2 by the left fore leg, and 1 by the left hind leg. The diameters of the tumors ranged from 2 to 4 cm along the planes of image acquisition.

Attenuation slope estimates from the Zonare, Siemens, and laboratory systems, including 3 different methods with the Siemens system, are summarized in Table 1. Estimates averaged across several ROIs in the tumors ranged between 0.37 and 3.91 dB/cm-MHz in the fibroadenomas and between 0.47 and 1.23 dB/cm-MHz in the carcinoma. For the 2 tumors for which the BSC is presented below, the attenuation slope estimates were 0.67 ± 0.26 dB/cm-MHz and 1.88 ± 0.80 dB/cm-MHz for the carcinoma and fibroadenomas, respectively. An attenuation slope of 1 dB/cm-MHz was used in all the analysis and processing techniques for consistency.

Table 1. Attenuation Slope Estimates From 3 Imaging Systems, Including 3 Different Methods for the Siemens System

Ultrasound System	Fibroadenoma		Carcinoma	
	Min Estimate, dB/cm-MHz	Max Estimate, dB/cm-MHz	Min Estimate, dB/cm-MHz	Max Estimate, dB/cm-MHz
Laboratory	1.19	2.02	0.47	0.56
Siemens (classic frequency shift)	1.21	3.91	0.50	0.59
Siemens (hybrid)	0.84	3.29	0.62	0.67
Siemens (reference phantom)	0.37	2.19	0.94	1.23
Zonare	0.71	2.46	0.42 ^a	

The minimum and maximum average estimated attenuation slopes across all transducers and tumors are quoted for each system and estimation method.

^aAn estimate was only available for 1 transducer.

Due to experimental errors and system limitations, some of the data series, including those in the tables and figures, were missing for particular tumors, and the remaining data were analyzed.

All transducers had considerable overlap in bandwidth with several other transducers, allowing for a direct visual comparison of the BSC magnitude and trends when displayed graphically. The carcinoma BSC results showed small variations between the multiple slices acquired with each transducer (see Figure 1A). Several of the data sets can be observed intersecting around 2 to 3 MHz (see average curves in Figure

1B), although there are some offsets in magnitude with increased frequency. Data from a single fibroadenoma are presented in Figure 2 and have substantial overlap between the different transducers and systems (Figure 2A). The variation between slices is of comparable magnitude to the variations between transducers and systems. In Figure 2B, it can be observed that the average curves from each transducer all intersect and have good agreement in magnitude. For the other 4 fibroadenomas, most data are similar to those presented in Figure 2; however, there are a couple of measurements from individual trans-

Figure 1. Backscatter coefficients from the single rat mammary carcinoma. Data are from the laboratory (blue), Ultrasonix (cyan), Siemens (red), and Zonare (magenta) systems (data from the Zonare 7-MHz system unavailable). **A**, Backscatter coefficient data from each acquired slice displayed without averaging. **B**, Backscatter coefficients averaged over all acquired 2-dimensional slices for each transducer.

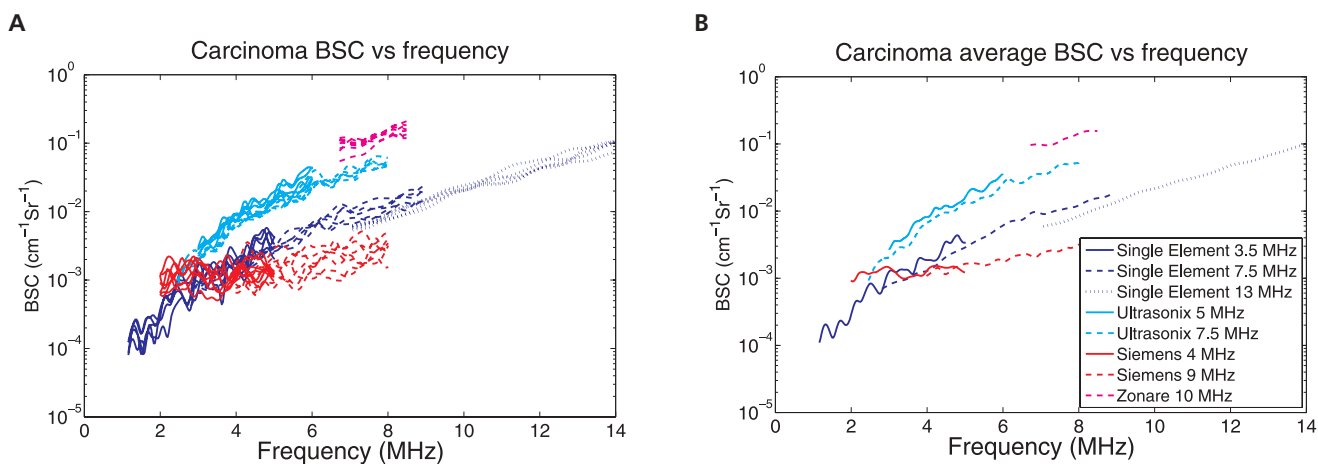
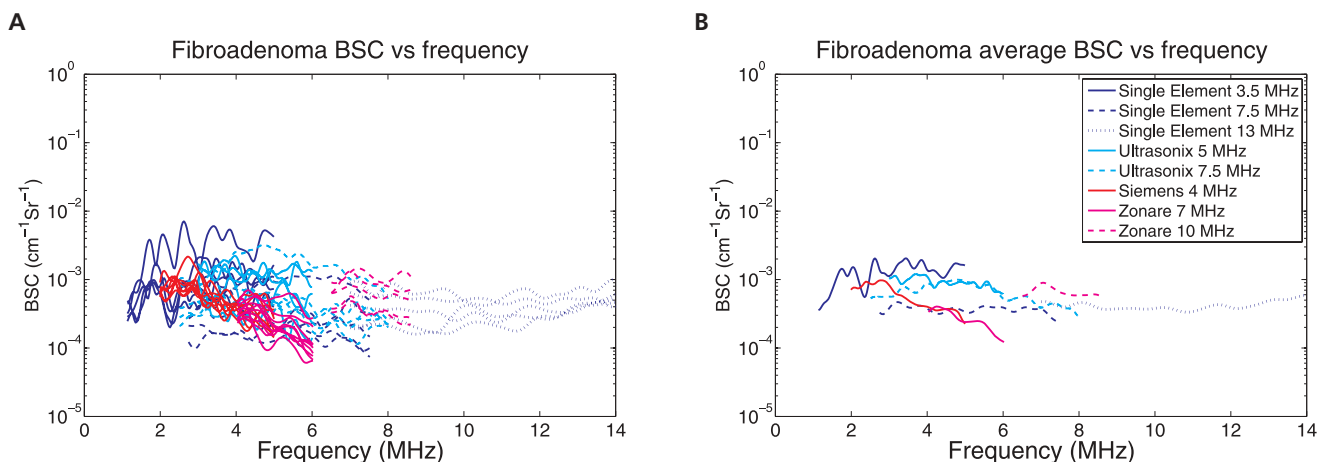


Figure 2. Backscatter coefficients from a rat fibroadenoma show substantial overlap between transducers and ultrasound systems. Data are from the laboratory (blue), Ultrasonix (cyan), Siemens (red), and Zonare (magenta) systems (data from the Siemens 9-MHz system unavailable). **A**, Backscatter coefficient data from each acquired slice displayed without averaging. **B**, Backscatter coefficients averaged over all acquired 2-dimensional slices for each transducer.



ducers that are offset in magnitude from the rest of the transducers that are not observed in the other tumors. It is possible that experimental error may have resulted in an offset in magnitude.

Both the functional ANOVA and RMS errors were calculated over the frequency ranges of 4 to 5 and 7 to 8 MHz for both the carcinoma and the single fibroadenoma presented in Figures 1 and 2, respectively. Results of the functional ANOVA indicate that for the carcinoma (Figure 1), there were significant differences in curves for the 2 frequency bands analyzed. For the fibroadenoma presented in Figure 2, there were no significant differences in BSC for the higher frequency range analyzed (7–8 MHz); however, for the lower range (4–5 MHz), there was a significant difference. As there were statistically significant differences observed, the RMS errors between pairs of curves over these ranges were calculated to aid in determining the magnitude of the discrepancies. Tables 2 through 5 summarize the RMS errors between pairs of data sets over the specified frequency range, with Tables 2 and 3 outlining the RMS errors for the carcinoma and Tables 4 and 5 outlining the RMS errors for the fibroadenoma. The RMS errors were highly variable, both in terms of the absolute value and also the percentage of BSC magnitude. The highest errors were observed for the carcinoma between 7 and 8 MHz (Table 3), with an RMS error of $1.1 \times 10^{-1} \text{ cm}^{-1}\text{Sr}^{-1}$ and an average BSC value of $3.5 \times 10^{-2} \text{ cm}^{-1}\text{Sr}^{-1}$, or approximately 300% error. The lowest RMS errors were observed for the fibroadenoma between 4 and 5 MHz (Table 4), with an RMS error of $4 \times 10^{-5} \text{ cm}^{-1}\text{Sr}^{-1}$ and an average BSC value of $7.1 \times 10^{-4} \text{ cm}^{-1}\text{Sr}^{-1}$, or approximately 5% error.

No significant differences in slope among the data sets for the fibroadenoma were observed using the 1-way ANOVA analysis. However, there were considerable differences in the slope of the data sets for the carcinoma.

Discussion

The agreement between all systems and transducers varied between the individual tumors, with some systems showing a much greater discrepancy for a particular tumor. However, the overall agreement of the BSC for this number of

Tables 2 and 3. Root Mean Squared Errors Between Averaged BSC Data Sets Acquired With Different Transducers for the Carcinoma (Figure 1)

$\times 10^{-2} \text{ cm}^{-1}\text{Sr}^{-1}$	1	2	4	5	6	7
1		0.1	0.9	0.7	0.2	0.2
2			1	0.8	0.08	0.08
4				0.2	1	1
5					0.8	0.8
6						0.02
7						

$\times 10^{-2} \text{ cm}^{-1}\text{Sr}^{-1}$	2	3	5	7	9
2		0.4	4	0.8	10
3			4	0.4	11
5				4	6.8
7					11
9					

Multiplicative factors are in the top left corner for each subtable. The transducers are numbered as follows: 1, single element 3.5 MHz; 2, single element 7.5 MHz; 3, single element 13 MHz; 4, Ultrasonix 5 MHz; 5, Ultrasonix 7.5 MHz; 6, Siemens 4 MHz; 7, Siemens 9 MHz; 8, Zonare 7 MHz; and 9, Zonare 10 MHz. For example, for the carcinoma in the 7- to 8-MHz frequency range, there was an RMS error of $4 \times 10^{-2} \text{ cm}^{-1}\text{Sr}^{-1}$ between the Ultrasonix 7.5-MHz transducer data and the Siemens 9-MHz transducer data. Bold values are within 20% relative RMS error (RMS error divided by mean BSC).

Tables 4 and 5. Root Mean Squared Errors Between Averaged BSC Data Sets Acquired With Different Transducers for the Fibroadenoma (Figure 2)

$\times 10^{-4} \text{ cm}^{-1}\text{Sr}^{-1}$	1	2	4	5	6	8
1		11	5.5	5.0	11	11
2			5.4	5.8	0.4	0.7
4				0.8	5.4	0.7
5					5.8	0.4
6						0.4
8						

$\times 10^{-4} \text{ cm}^{-1}\text{Sr}^{-1}$	3	5	9
3		2.9	0.6
5			3.2
9			

Multiplicative factors are in the top left corner for each subtable. The transducers are numbered as follows: 1, single element 3.5 MHz; 2, single element 7.5 MHz; 3, single element 13 MHz; 4, Ultrasonix 5 MHz; 5, Ultrasonix 7.5 MHz; 6, Siemens 4 MHz; 7, Siemens 9 MHz; 8, Zonare 7 MHz; and 9, Zonare 10 MHz. For example, for the fibroadenoma in the 4- to 5-MHz frequency range, there was an RMS error of $0.7 \times 10^{-4} \text{ cm}^{-1}\text{Sr}^{-1}$ between the Ultrasonix 5-MHz transducer data and the Zonare 7-MHz transducer data. Bold values are within 20% relative RMS error (RMS error divided by mean BSC).

systems is encouraging, especially as the first in vivo intercomparison. Despite the differences in processing techniques, including the different types of references required between the single-element and array systems, differences in ROI sizes, system architecture, and transducer frequency characteristics, it was possible to maintain good agreement. This suggests that the obtained BSC is a fairly robust estimate and not particularly sensitive to the analysis parameters. Room for improvement may lie within the experimental procedures to control for variations from one scan to the other and allow for more consistent data to be acquired.

Fibroadenomas are heterogeneous within an individual tumor and have substantial variations in structure from one tumor to the next. This heterogeneity within the tumor may account for BSC variations from slice to slice and would also result in variations between equipment due to limitations in acquiring the same plane. The carcinoma histologic evaluation showed a well-differentiated carcinoma with less tissue heterogeneity compared to the fibroadenoma, which may explain the lower variability in BSC between slices.

Although statistically significant differences in BSC values were observed over some ranges, the value of the offsets calculated by the RMS error indicates that the magnitudes of the offsets were quite small, especially in the case of the fibroadenoma presented. Furthermore, in several instances, there was such high reproducibility of BSC results for different slices for the same transducer that even very small differences in BSC between transducers were statistically significant. The combination of the statistical methods with the RMS error calculations could provide us with metrics for comparison of future studies to determine where improvements arise.

A limitation in attempting to establish agreement between BSC values from different imaging systems was the lack of a standardized method for evaluating the agreement in BSCs across different platforms. Statistical tests for differences cannot establish equivalence. Additionally, a large number of samples increases the likelihood of detecting statistically significant differences, which may or may not be diagnostically significant. To use metrics such as the RMS error between data sets, it becomes necessary to first

decide how close together data sets need to be to constitute sufficient agreement. This study used a provisional criterion of 20% or less relative RMS error between measurements, compared with the mean value of the measurements. The slope of the BSC is tied to the effective scatterer size and therefore might be a value of more fundamental interest to compare across imaging systems. At this point, there is no clear choice of the best metric for comparing BSC values or what the necessary level of agreement is to perform the desired analyses based on the BSC. It will be an important consideration going forward to determine how to compare results between different imaging platforms.

In conclusion, despite some variability between individual tumors, these results suggest that it is possible to obtain agreement among BSC measurements from different ultrasound systems scanning a live animal, even when processing techniques vary. This type of agreement demonstrates the potential to base diagnoses on QUS parameters.

References

1. Wear KA, Stiles TA, Frank GR, et al. Interlaboratory comparison of ultrasonic backscatter coefficient measurements from 2 to 9 MHz. *J Ultrasound Med* 2005; 24:1235–1250.
2. Wilson T, Zagzebski J, Li Y. A test phantom for estimating changes in the effective frequency of an ultrasonic scanner. *J Ultrasound Med* 2002; 21:937–945.
3. Insana MR, Hall TJ. Parametric ultrasound imaging from backscatter coefficient measurements: image formation and interpretation. *Ultrason Imaging* 1990; 12:245–267.
4. Madsen EL, Insana MF, Zagzebski JA. Method of data reduction for accurate determination of acoustic backscatter coefficients. *J Acoust Soc Am* 1984; 76:913–923.
5. Chen J, Zagzebski JA, Dong F, Madsen EL. Estimating the spatial autocorrelation function for ultrasound scatterers in isotropic media. *Med Phys* 1998; 25:648–655.
6. Yao LX, Zagzebski JA, Madsen EL. Backscatter coefficient measurements using a reference phantom to extract depth-dependent instrumentation factors. *Ultrason Imaging* 1990; 12:58–70.
7. Haak A, Hafez ZT, Anderson JJ, et al. Algorithm for estimating the attenuation slope from backscattered ultrasonic signals. In: *Proceedings of the 2009 IEEE International Ultrasonics Symposium*. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2009:1946–1949.
8. Tu H, Varghese T, Madsen EL, Chen Q, Zagzebski JA. Ultrasound attenuation imaging using compound acquisition and processing. *Ultrason Imaging* 2003; 25:245–261.

9. Kuc R, Li H. Reduced-order autoregressive modeling for center-frequency estimation. *Ultrason Imaging* 1985; 7: 244–251.
10. Kim H, Varghese T. Hybrid spectral domain method for attenuation slope estimation. *Ultrasound Med Biol* 2008; 34:1808–1819.
11. Shen Q, Faraway J. An F test for linear models with functional responses. *Statistica Sinica* 2004; 14:1239–1257.
12. Cuevas A, Febrero M, Fraiman R. An ANOVA test for functional data. *Comput Stat Data Anal* 2004; 47:111–122.
13. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York, NY: Chapman & Hall; 1993.