Phantom-based image quality test methods for photoacoustic imaging systems

William C. Vogt
Congxian Jia
Keith A. Wear
Brian S. Garra
T. Joshua Pfefer

Phantom-based image quality test methods for photoacoustic imaging systems

William C. Vogt,* Congxian Jia, Keith A. Wear, Brian S. Garra, and T. Joshua Pfefer
U.S. Food and Drug Administration, Center for Devices and Radiological Health, 10903 New Hampshire Avenue, Silver Spring, Maryland, United States

Abstract. As photoacoustic imaging (PAI) technologies advance and applications arise, there is increasing need for standardized approaches to provide objective, quantitative performance assessment at various stages of the product development and clinical translation process. We have developed a set of performance test methods for PAI systems based on breast-mimicking tissue phantoms containing embedded inclusions. Performance standards for mature imaging modalities [magnetic resonance imaging (MRI), computed tomography (CT), and ultrasound] were used to guide selection of critical PAI image quality characteristics and experimental methods. Specifically, the tests were designed to address axial, lateral, and elevational spatial resolution, signal uniformity, penetration depth, sensitivity, spatial measurement accuracy, and PAI-ultrasound coregistration. As an initial demonstration of the utility of these test methods, we characterized the performance of a modular, bimodal PAI-ultrasound system using four clinical ultrasound transducers with varying design specifications. Results helped to inform optimization of acquisition and data processing procedures while providing quantitative elucidation of transducer-dependent differences in image quality. Comparison of solid, tissue-mimicking polymer phantoms with those based on Intralipid indicated the superiority of the former approach in simulating real-world conditions for PAI. This work provides a critical foundation for the establishment of well-validated test methods that will facilitate the maturation of PAI as a medical imaging technology. © *2017 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JBO.22.9.095002]

Keywords: image quality; ultrasound; mammography; tissue phantoms; standards.

Paper 170200PR received Mar. 29, 2017; accepted for publication Aug. 15, 2017; published online Sep. 12, 2017.

1 Introduction

1.1 Photoacoustic Imaging

Photoacoustic Imaging (PAI) is a rapidly maturing imaging modality that harnesses the photoacoustic effect, where pulsed optical irradiation is absorbed by tissue and converted to thermoacoustic waves, to enable imaging with spatial resolution similar to ultrasound imaging, contrast based on optical absorption, and penetration to several centimeters.[1] Due to the ability to detect both endogenous and exogenous chromophores, PAI is a promising technique for many applications including cancer detection,[2] mammography,[3,4] oximetry,[5] and molecular imaging.[6]

A wide variety of PAI system design is observed in the literature,[4,7–15] indicating significant potential for variation in image quality. However, there are currently no well-standardized approaches for evaluating PAI system image quality. Such consensus test methods are needed to enable objective, quantitative, and consistent evaluation of PAI device performance. The availability of these tools would be beneficial for many steps of the product life cycle, including product development (optimization of instrument design, image processing algorithms, and standardization of clinical trials), regulatory clearance (quantitative evidence of effectiveness and substantial equivalence to a predicate device), manufacturing (quality control and calibration), and clinical implementation (quality assurance, constancy testing, and recalibration).

Phantom-based test methods are a mainstay of image quality test methods for well-established clinical imaging modalities, such as magnetic resonance imaging (MRI), computed tomography (CT), and ultrasound, and many commercial phantoms are available for quality control, user training, and research. These test methods have also been incorporated into international standards,[16,17] consensus documents,[18] and accreditation programs.[19,20] A recent review of these documents identified a set of commonly used performance characteristics for medical imaging.[21] Given the technological similarities between ultrasound and PAI, consensus standards for ultrasound image quality represent potential foundations for development of consensus PAI test methods.[16–18,22] Several of the essential performance characteristics mentioned in these prior documents include: spatial resolution (axial, lateral, and elevational), signal uniformity, spatial measurement accuracy, sensitivity, and penetration depth. Such characteristics are generally measured using tissue phantoms with simple, idealized target inclusions, e.g., wires for spatial resolution or filled cylindrical inclusions for sensitivity.[16,17] Generally, the goal of these standards is to present robust, comprehensive test methods that incorporate phantoms with biologically relevant properties and shape, not biomimetic replicas of specific tissue properties and morphology.

1.2 Literature Review of Photoacoustic Image Quality Testing

Many prior PAI studies have used phantom-based testing to assess image quality. However, there has been a wide variation in phantom geometric configurations, target inclusions, and the intrinsic optical and acoustic properties of the phantom media.

This variation leads to inconsistency in comparing performance results from different studies. One of the most commonly tested performance characteristics is spatial resolution, which may include axial, lateral, and elevational components. For resolution measurements, many materials have been used for generating high-contrast, subresolution targets, including human hairs,[23] metal wires,[24] and polymer ribbons.[11] These targets are often suspended in a liquid such as water or Intralipid, which have acoustic properties that are not biologically realistic. While convenient, liquid phantoms such as Intralipid are often insufficiently stable for long-term calibration or constancy testing.[25]

Other commonly evaluated characteristics include sensitivity and penetration depth. For sensitivity testing, various concentrations of a chromophore solution are injected into tubes or mixed into solid inclusions in *ex vivo* tissue or a biologically relevant tissue-mimicking phantom containing blood or contrast agents.[26–30] An absorption coefficient or concentration threshold for detectability can then be determined. In addition to quantifying PAI system sensitivity in terms of intrinsic absorption coefficient, this test method can be used to compare PAI system sensitivity using different contrast agents[31] or to assess the linearity of photoacoustic signal amplitude versus target absorption or chromophore concentration. Maximum penetration depth, another performance characteristic that is related to sensitivity, may be determined using an array of light-absorbing inclusions at increasing depths. A detection threshold based on signal-to-noise ratio or contrast-to-noise ratio is chosen and can be qualitatively compared with visual detectability of targets by an observer. This approach has been used in phantoms and *ex vivo* tissues including fluid-filled tubes embedded in a phantom,[32] ink-doped solid phantom inclusions,[33] or a single tube/capsule filled with blood or contrast agent, which is repositioned several times within layers of chicken muscle tissue.[26,29] However, the use of *ex vivo* tissue is incompatible with standardized testing due to biological variability and lack of *a priori* knowledge of ground truth in sample geometry and tissue properties.

In previous studies, imaging targets or tissue phantoms have been used for performance comparison of acoustic transducers used in a PAI system.[29,32,34–36] However, these studies employed simplified approaches, such as (1) phantoms with low optical and/or acoustic biological relevance, (2) target inclusion materials with uncharacterized or unknown optical or acoustic properties, (3) use of qualitative or subjective image quality metrics, (4) use of *in vivo* measurements (which are poorly reproducible), and (5) implementing dissimilar tissue-mimicking materials for different tests. Furthermore, these studies tended to focus on a small number of performance characteristics, whereas to adequately elucidate trade-offs in performance due to device design, a more comprehensive performance characterization must be applied. While these phantom tests were adequate for the purpose of each individual study, there remains a critical need for a set of robust, comprehensive test methods that incorporate well-characterized, biologically relevant phantoms to guide future standardization of image quality test methods for PAI.

Given the limitations of prior performance testing approaches, the purpose of this study was to develop a comprehensive suite of objective, quantitative, test methods based on stable, reproducible, and biologically relevant phantoms for evaluating image quality of PAI systems. To this end, we constructed a series of test-specific phantoms based on a previously developed photoacoustic breast-mimicking material.[37] Several

potential tissue-mimicking materials (TMMs) have been proposed for use in PAI phantoms, including hydrogels,[38] polyvinyl alcohol cryogels,[39] and polyvinyl chloride plastisol (PVCP).[40] We refer the reader to our previous work for a more detailed discussion of the properties of these TMMs and their suitability for PAI.[37] While test methods may utilize any TMM with adequate biological relevance for the intended application(s), we chose to construct phantoms using our custom breast-mimicking PVCP-based TMM due to the prevalence of mammography-related PAI research and the temporal stability of this material. Phantoms contained inclusions suitable for assessing a wide array of image quality performance characteristics, including 3-D spatial resolution, uniformity, spatial measurement accuracy, PAI-ultrasound coregistration accuracy, sensitivity, and penetration depth. We then demonstrated the utility of these phantoms for quantitative comparison of different PAI systems by swapping several acoustic transducers into a custom modular PAI system, thus simulating a comparison of different imaging systems.

2 Methods

2.1 Modular Photoacoustic Imaging System

To demonstrate the utility of phantom-based image quality testing for comparing the performance of different devices, we characterized the performance of a custom bimodal photoacoustic-ultrasound imaging system described previously.[37] Briefly, this system is comprised of a tunable near-infrared optical parametric oscillator (Phocus Mobile, Opotek, Inc., Carlsbad, California) and a 128 channel, research-grade ultrasound system (Vantage 128, Verasonics, Inc., Kirkland, Washington). In this study, all imaging was performed at a wavelength of 800 nm with a radiant exposure of 20 mJ/cm^2 (below the maximum permissible exposure for skin of $\sim$31.7 mJ/cm^2 at 800 nm[41]) over a $\sim$5-mm $\times$ 35-mm elliptical beam. The beam was positioned $\sim$5 mm from each transducer to reduce near-field photoacoustic signal generation resulting in image artifacts. Image reconstruction was performed using a proprietary Verasonics beamforming method, which performs in-phase/quadrature (IQ) demodulation and takes the absolute value of the IQ data to produce nonnegative image intensities. A plane-wave transmit beamforming algorithm was used, and therefore, lateral resolution was determined completely by receive beamforming, which was the same for ultrasound and PAI. For imaging in both liquid and solid tissue phantoms, speed of sound for the reconstruction algorithm was set to 1480 m/s. All phantoms, with the exception of the elevational resolution phantom (Sec. 2.3.2), were imaged at 10 distinct spatial locations along the elevational (out of plane) direction.

An important feature of this modular system is the ability to substitute different ultrasound transducer arrays, each with different operating parameters (Table 1 and Fig. 1). By testing each transducer, we can simulate performance comparison of different imaging systems since ultrasound image properties are dominated by transducer characteristics (e.g., aperture, center frequency, bandwidth, and elevational focus). Four transducers were used for B-mode ultrasound imaging and PAI of test phantoms, including a linear array (L11-4v, Verasonics), a high-density multiplexed array (ATL L12-5-50 mm, Priority Medical, Inc., Greenbrier, Tennessee), a high-frequency intraoperative array (ATL CL15-7, Priority Medical), and a low-frequency-phased array (ATL P4-1, Priority Medical). For all

Table 1 Operating parameters of four clinical ultrasound transducer arrays used for PAI.

Transducer	Center frequency (MHz)	−20 dB bandwidth (MHz)	Number of elements	Pitch (mm)	Length (mm)	Elevational focus depth (mm)
L11-4v	8	8.7	128	0.300	38.4	~20
L12-5	8.7	7.6	256	0.200	51.2	~20
CL15-7	12.4	7.9	128	0.178	22.8	~15
P4-1	2.5	2.2	96	0.295	28.3	~80
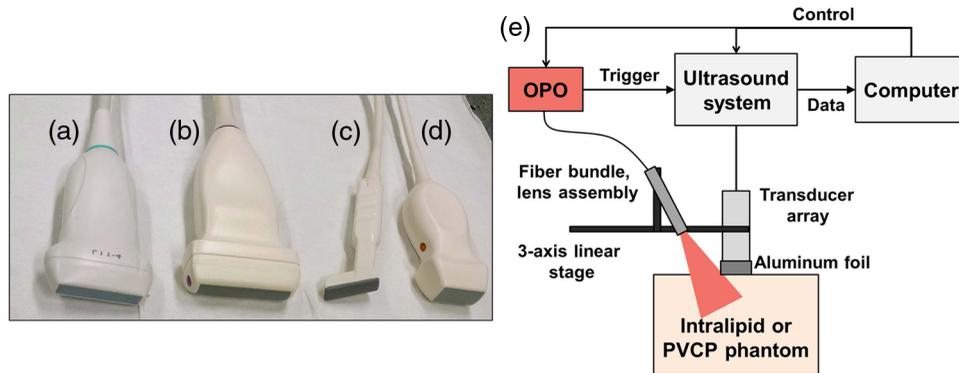
Fig. 1 Clinical ultrasound transducers compared photoacoustic image quality tests: (a) L11-4v, (b) L12-5, (c) CL15-7, (d) P4-1, and (e) schematic of the PAI system.

transducers except the L12-5, full images were acquired per laser pulse. Because the L12-5 has 256 elements, but our acquisition system only has 128 channels, three 128-element subaperture scans were acquired to generate full images (this required three laser pulses per image). All transducers were shielded with aluminum foil to reduce surface-generated photoacoustic artifacts and improve image contrast.[42] Background images were measured by averaging 30 frames with no laser output, then subtracted from target images in postprocessing. Images were normalized to the maximum intensity of the most shallow target inclusion and log-compressed, with dynamic range optimized for each transducer. Time gain compensation was not used in this study.

2.2 Tissue-Mimicking Material

Solid tissue phantoms were constructed using a previously developed breast-mimicking PVCP formulation.[37] Briefly, this material consists of 10% m/m PVC (Geon 121A, Mexichem Specialty Resins, Inc., Avon Lake, Ohio) suspended in a 3:1 mixture of benzyl butyl phthalate and di(2-ethylhexyl) adipate (308501 and 525197, Sigma-Aldrich, St. Louis, Missouri), to which 1% v/v calcium–zinc heat stabilizer was added to prevent thermal discoloration (HS16, M-F Manufacturing Co., Fort Worth, Texas). Two mg/mL anatase titanium dioxide (248576, Sigma-Aldrich) was added to impart tissue-relevant optical scattering, and 30 mg/mL soda lime glass microbeads (38- to 63-μm diameter, Spheriglass A, Potter Industries LLC, Malvern, Pennsylvania) was added to provide acoustic scattering, enabling use of the phantoms for both photoacoustic and ultrasound imaging. From acoustic through-transmission measurements performed as described previously,[37,43] this formulation has a speed of sound of 1467 ± 4 m/s, acoustic

attenuation of $(0.22 \pm 0.03) \times f^{1.76 \pm 0.06}$ dB/cm, where f is the frequency in MHz. Optical properties (Fig. 2) were determined for each PVCP phantom by performing spectrophotometry measurements in disk samples molded from each phantom's PVCP batch (38-mm diameter and 5-mm thickness), then using the inverse adding-doubling algorithm to compute absorption and reduced scattering coefficients.[37,44]

To construct phantoms, PVCP was heated in 75-mL batches in an evacuated 250-mL round-bottom flask suspended in an oil bath at 200°C. A rare-Earth stir bar was driven at 350 rpm, and the PVCP was heated for 13 min. After heating, the flask was removed and suspended over a magnetic stir plate, and was allowed to cool under stirring to ~140°C before pouring into an acrylic phantom mold. This intermediate cooling step increases the PVCP viscosity during pouring, which substantially reduces glass bead settling during the solidification phase. Phantoms required four layered pours based on mold volume. Molds contained target tubes or filaments threaded through holes in the mold walls and epoxied in place to ensure accurate positioning. An aluminum plate was used to cover the mold's top opening, resulting in a smooth surface for imaging, and pouring was performed vertically through a 4-cm × 4-cm side wall opening such that layer interfaces did not appear in photoacoustic images.

2.3 Image Quality Phantoms

This section describes four tissue phantoms used in this study, with each phantom enabling quantitative assessment of one or more image quality characteristics. These phantoms are denoted as follows: (1) filament array phantom, (2) elevational resolution phantom, (3) sensitivity phantom, and (4) penetration depth phantom (see Fig. 3).
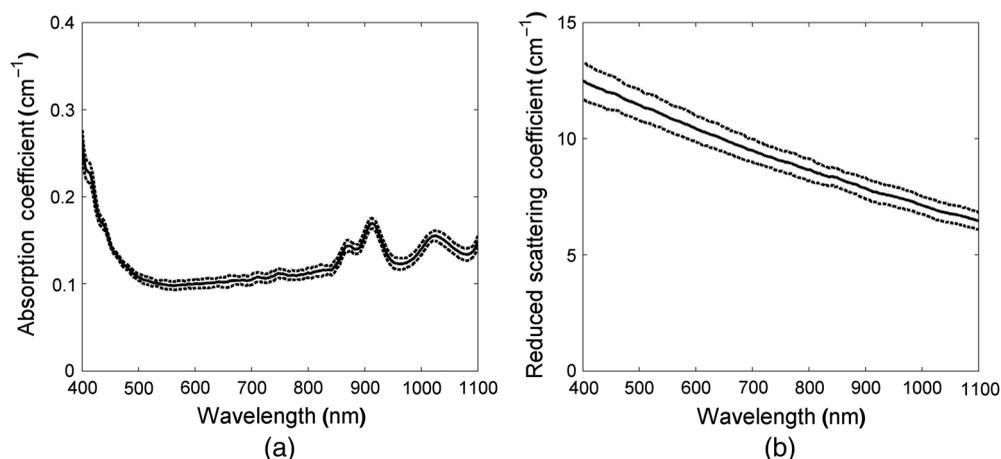
Fig. 2 PVCP TMM optical properties, including (a) absorption and (b) reduced scattering coefficients. Solid lines denote mean values, while dotted lines denote 95% confidence intervals.
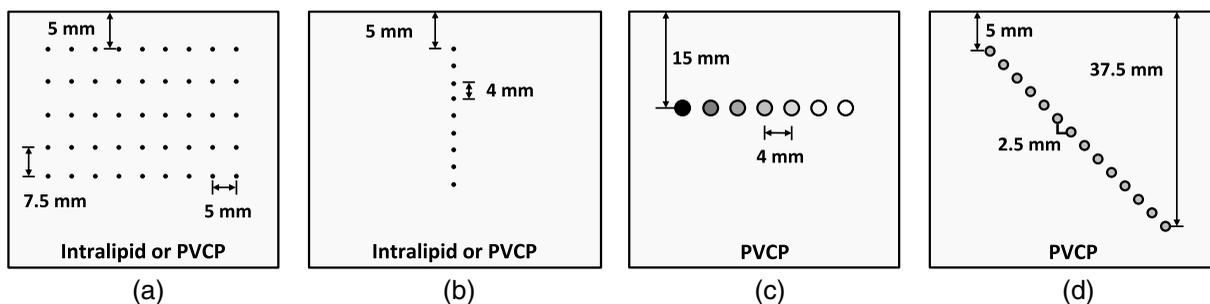
Fig. 3 Set of image quality phantoms: (a) filament array, (b) elevational resolution, (c) sensitivity, and (d) penetration depth phantoms. Imaging was performed from the top surface in each phantom.

2.3.1 Filament array phantom

Ultrasound spatial resolution test methods may use visual distinction of pairs of subresolution filaments, or more quantitative measurements, such as the full width at half-maximum (FWHM) of the line spread function (LSF), which correlates with the practical limit of visual resolution of two closely spaced sources in displayed images.[16,18,45] We sought to evaluate differences between solid and liquid phantoms by constructing spatial resolution phantoms with similar geometries but using either Intralipid or PVCP as the background medium. The Intralipid phantom consists of a 5×9 filament grid of black monofilament suture wires with 50-μm diameter and interwire spacing of 7.5-mm vertically and 5.0-mm horizontally suspended in 1% Intralipid solution prepared immediately before imaging (I141, Sigma-Aldrich). To prevent thermal and mechanical filament damage as well as improve target positioning accuracy and repeatability in PVCP gels, 50.8 μm (0.002″) diameter stainless steel wires (794600, A-M Systems, Inc., Sequim, Washington) were used, which have higher mechanical strength and better heat tolerance than suture filaments. Preliminary tests showed that steel wires suspended in Intralipid produced stronger reflection or "ghost" artifacts than nylon sutures, perhaps due to higher acoustic reflectivity of metal filaments. Since these artifacts could have overlapped with other targets and impacted resolution analysis, we chose to use nylon sutures in Intralipid phantoms. As will be shown, steel wires embedded in PVCP presented much weaker artifacts due to higher acoustic attenuation of the phantom medium.

In-plane spatial resolution was quantified in both phantoms by selecting a rectangular region of interest (ROI) over each target LSF then locating the maximum intensity pixel. Target vertical (axial for linear arrays) and horizontal (lateral for linear arrays) intensity profiles through this maximum were used to determine resolution based on FWHM, or equivalently the -6 dB threshold distance for log-compressed images.[45] This analysis was performed in both ultrasound and photoacoustic images given that many PAI systems presented in the literature are bimodal and include coregistered ultrasound imaging.

The ability of a PAI system to accurately locate target structures, interpret their brightness, and measure their size can be an important clinical consideration. Therefore, we also utilized the spatial resolution phantom to test PAI intensity uniformity, PAI-ultrasound coregistration accuracy, and spatial measurement accuracy. The mean intensity as well as the vertical and horizontal positions of the maximum pixel of each filament target was calculated from log-compressed images. Mean intensity was calculated by selecting a 2-mm $\times$ 2-mm rectangular ROI around each target, applying a mask to select pixels above 50% of the max intensity in the ROI, then averaging intensity of this pixel subset. Intensity data versus target position were displayed as an image to visualize the intensity pattern of the filament grid. Coregistration accuracy was computed by comparing PAI-based target positions with those determined from ultrasound images while spatial measurement accuracy was determined by measuring the vertical and horizontal distances between all adjacent filament pairs for both photoacoustic and ultrasound images,

then comparing them to the known designed values for target separation distances.

2.3.2 Elevational resolution phantom

Elevational (or "out of plane") resolution is another important component of image quality for PAI system performance, particularly for systems utilizing linear array transducers, which generally have poor elevational resolution relative to axial and lateral resolution. Elevational LSFs may be generated by mechanically scanning the transducer along the short axis of a wire target, as in ultrasound resolution testing.[16] However, the previously described filament array phantom cannot be used for this test due to the high target density, which results in significant image clutter that precludes accurate elevational LSF measurement. Thus, separate liquid and PVCP phantoms were constructed consisting of a vertical column of nine filament targets (nylon sutures for liquid phantoms, steel wires for PVCP phantoms as before) with 4-mm spacing. Elevational scans were performed using a motorized stage, with ultrasound and photoacoustic images acquired every 0.25 mm. Scanning was stopped once targets were no longer visually detectable. Images of the elevational LSFs were generated as the mean of the image stack volume along the lateral dimension. Elevational resolution was calculated as the FWHM of the elevational LSF, averaging results over each vertical–horizontal slice of the stacked image volume.

2.3.3 Sensitivity phantom

A PVCP phantom was constructed to characterize PAI system sensitivity. Because preliminary acoustic characterization measurements of Intralipid showed low acoustic attenuation compared to tissue values (<0.1 dB/cm/MHz, data not shown), we expected Intralipid phantoms for sensitivity and penetration depth testing, which require biologically relevant target contrast versus depth, to be of limited biological relevance and thus did not construct such phantoms. The PVCP phantom comprised a horizontal array of seven polytetrafluoroethylene (PTFE) tubes with nominal inner diameter of 1.07 mm (STT-18, Zeus light-wall PTFE tubing, Component Supply Company, Fort Meade, Florida). To produce imaging targets with varying contrasts, six tubes were filled with a different concentration of India ink (3080-F, Chartpak, Inc., Leeds, Massachusetts) to produce absorption coefficients at 800 nm of 1, 2, 4, 6, 8, and 10 cm^{-1} while the seventh tube was filled with deionized water as a control. Because of boundary-buildup artifacts prevalent in PAI due to limited transducer bandwidth, a rectangular ROI would include the channel lumen, which has substantially lower intensity than the wall signals. If the ROI included this dark lumen, computed contrast and noise would differ significantly from the visually observed feature (strong wall signals). Thus, target ROIs were selected using a 50% of maximum intensity (−6 dB) masked subset of a 2- × 2-mm rectangular ROI placed over each target. Using this mask, the target ROI includes two separate regions containing the top and bottom wall features but produces a single contiguous region for cases in which the separate wall signals cannot be resolved. Background ROIs of the same size as the rectangular target ROI were selected immediately adjacent and to the right of each target ROI (before masking). Photoacoustic images may be presented in a linear or log-compressed form depending on whether the imaging application calls for quantitative estimation of target absorption coefficient

and thus chromophore concentration[46] or calls for enhanced contrast for qualitative image interpretation.[29,47] In this study, we computed target signal-to-noise ratio (SNR) using uncompressed image intensity data while target contrast-to-noise ratio (CNR) was computed using log-compressed data. Thus, SNR serves as a metric that is independent of image post-processing steps, such as compression and dynamic range adjustment while CNR quantifies target detectability based on the final displayed image. These quantities were computed as

$$\text{SNR} = \frac{I}{\sigma}, \qquad \text{CNR} = \frac{I - B}{\sigma}, \qquad (1)$$

where I is the average target ROI intensity, B is the average background ROI intensity, and σ is the background ROI standard deviation. The minimum detectable optical absorption coefficient was determined for each transducer by interpolating target SNR versus absorption coefficient to an SNR threshold of 2 (6 dB).

2.3.4 Penetration depth phantom

While the sensitivity phantom provides a convenient method for providing a single image containing targets with varying absorption coefficients, these results are limited because they do not account for variable signal attenuation with depth. Ultrasound visualization/penetration depth may be assessed using a homogeneous backscattering phantom region[17] or a diagonal array of cylindrical targets with known signal intensity or contrast.[18] Following the latter approach, a PVCP phantom containing a diagonal array of 0.559-mm diameter PTFE tubes (STT-24, Component Supply Company) spaced 2.5 mm vertically and horizontally was used to characterize signal attenuation and maximum penetration/visualization depth. Tubes were filled with a black ink solution with an optical absorption coefficient of 4.0 cm^{-1} at 800 nm (near the isosbestic point of oxy- and deoxy-hemoglobin), which corresponds to a hemoglobin concentration of 13.6 g/dL. Typical hemoglobin concentration in adult males ranges from 13.5 to 17.5 g/dL, whereas the range in females is from 12 to 16 g/dL.[48] ROI analysis and calculation of SNR and CNR were performed as described above for sensitivity phantom data.

3 Results

3.1 In-Plane Resolution

Representative ultrasound (Fig. 4) and photoacoustic (Fig. 5) images acquired in liquid and solid filament array phantoms are presented for each transducer. In all liquid phantom images, significant lateral streak artifacts were observed. PVCP images did not have these strong artifacts, likely due to the higher acoustic attenuation of the medium. Some near-field clutter and horizontal bands were also present, which may have been caused by the protective aluminum foil layer. The glass beads in PVCP phantoms also contribute to clutter due to enhanced acoustic scattering of photoacoustic signals generated at the transducer surface. Additionally, because the P4-1 images were reconstructed over a sector rather than a rectangular format, the targets appeared to rotate inward. This occurs because the axial direction for this scan format is actually a line from the image point to the transducer center rather than a vertical line in the image. The method we utilized for quantifying spatial resolution operates on the scan-converted image and therefore determines
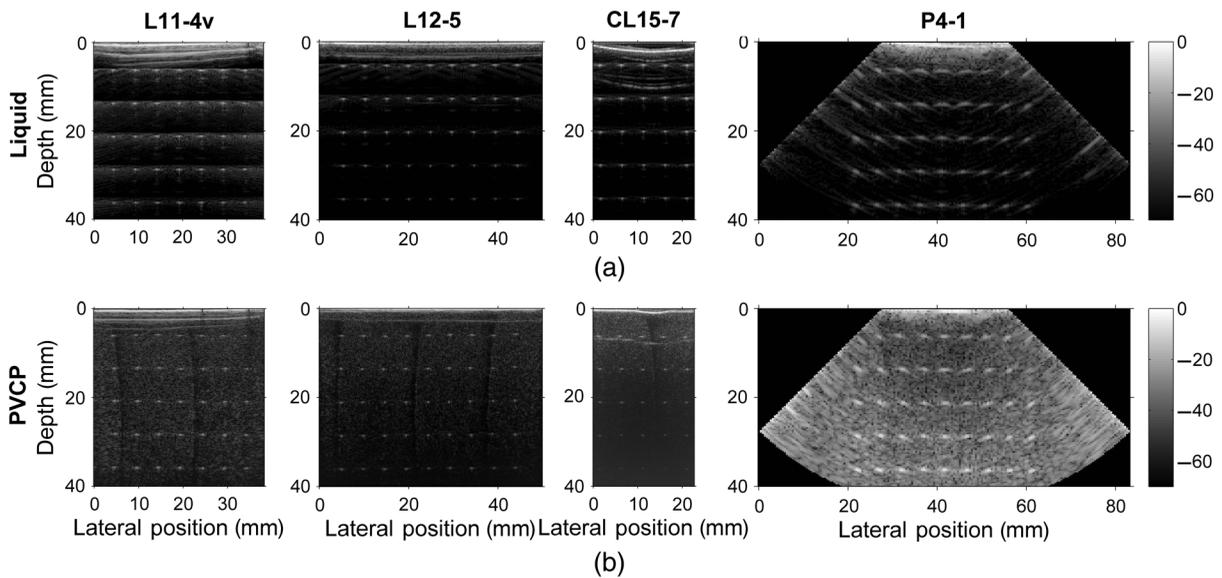
Fig. 4 Representative ultrasound images of the (a) liquid and (b) PVCP resolution phantoms, acquired using, from left to right columns, L11-4v, L12-5, CL15-7, and P4-1 transducers. Scale bars in dB.
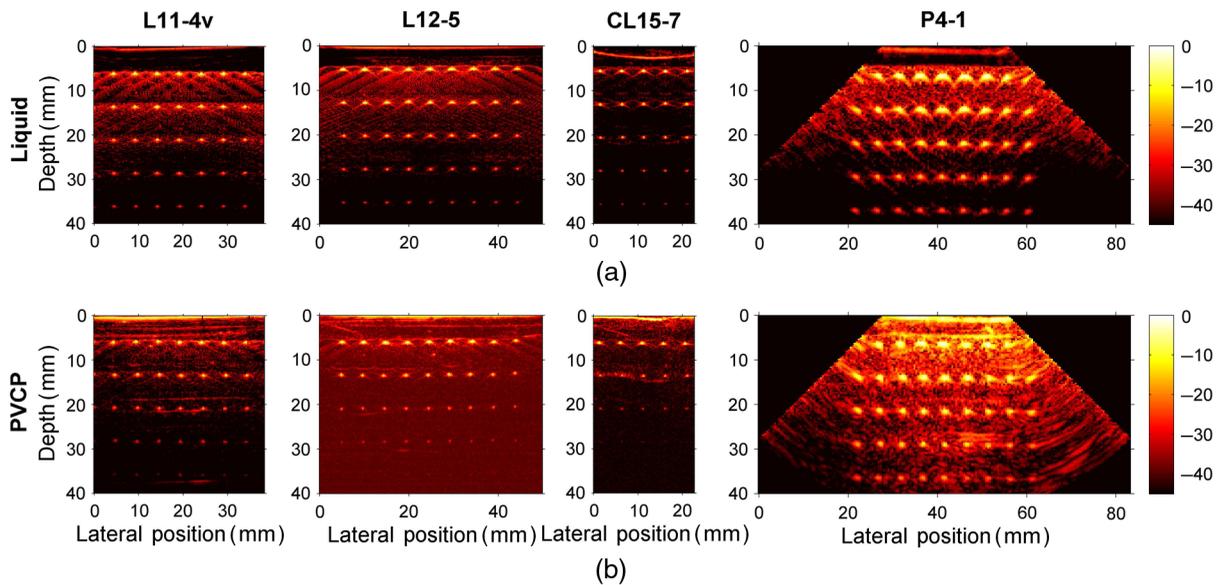
Fig. 5 Representative photoacoustic images of the (a) liquid and (b) PVCP resolution phantoms, acquired using, from left to right columns, L11-4v, L12-5, CL15-7, and P4-1 transducers. Scale bars in dB.

image vertical and horizontal resolutions, which is equivalent to axial and lateral resolution for the three linear array transducers but not the P4-1-phased array.

For liquid phantoms, the transducer was immersed at approximately the same location for each measurement, but there was some variability among the four transducers as can be seen in the slight variations of target depths (Figs. 6 and 7). This variability was reduced in PVCP phantoms, which provide a consistent planar surface on which to place the transducer. This demonstrates another advantage of using a solid phantom over a liquid medium.

Vertical resolution was found to not depend significantly on target depth in the liquid phantom (Figs. 6 and 7). Additionally, three transducers (L11-4v, L12-5, and CL15-7) show similar resolution values due to their similar acoustic bandwidths

(see Table 1), while the P4-1 array has lower bandwidth and thus worse axial resolution. However, vertical resolution spot size increased with depth in the PVCP phantom for the L11-4v, L12-5, and CL15-7 transducers. This effect is caused by the phantom's nonlinear (with frequency) acoustic attenuation coefficient, which decreases bandwidth with propagation depth.[49] This effect was not observed with the P4-1 array due to its lower center frequency and the lower phantom attenuation nonlinearity in this frequency range.

Comparison of Figs. 6(b) and 7(b) shows that lateral resolution was similar for ultrasound and photoacoustic images. This is because plane-wave transmit was used for ultrasound imaging, and therefore lateral resolution was completely determined by receive beamforming, which was true for both ultrasound and photoacoustic images. In the liquid phantom, the L11-4v and
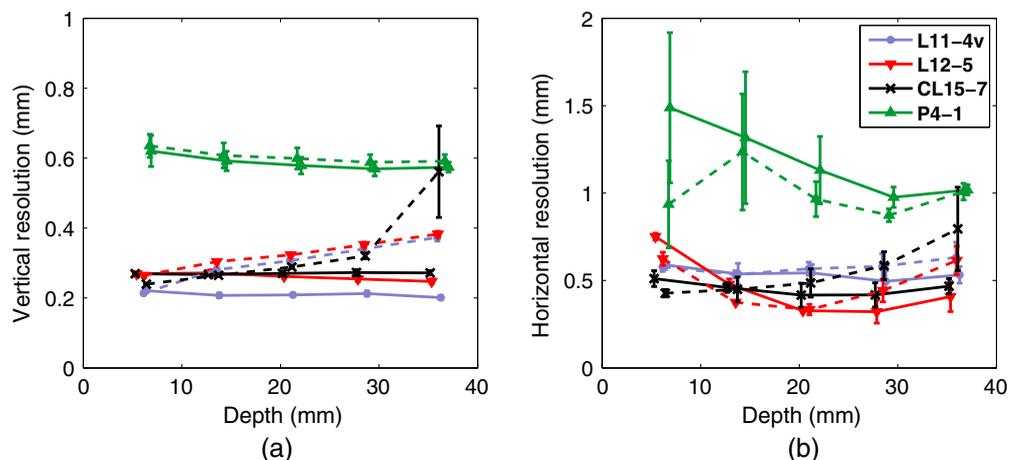
Fig. 6 Ultrasound (a) vertical and (b) horizontal resolution measurements (FWHM) versus depth in liquid (solid lines) and PVCP phantoms (dashed lines). Error bars denote 95% confidence.
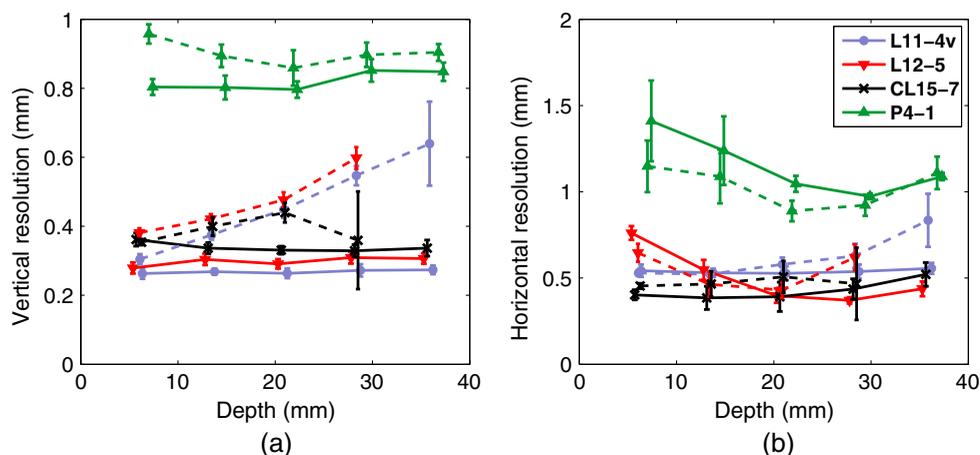
Fig. 7 Photoacoustic (a) vertical and (b) horizontal resolution measurements versus depth (FWHM) in liquid (solid lines) and PVCP phantoms (dashed lines). Error bars denote 95% confidence.

CL15-7 transducers had relatively constant lateral resolution, whereas the L12-5 and P4-1 transducers showed measureable lateral focusing effects. For the L12-5, this was due to the wide lateral transducer aperture, whereas the P4-1 focusing may be a product of its convex elevational element geometry, as it is optimized for deep tissue ultrasound. Ultrasound resolution data followed similar trends with depth as PAI data, although vertical resolution was generally better in ultrasound images than in photoacoustic images. Worse PAI vertical resolution may be caused by out-of-plane photoacoustic signal contributions from the filament targets, which would blur the reconstructed target, or targets may have been blurred by a combination of frame averaging and trigger jitter observed in PAI mode (1 to 2 vertical pixels).

We also evaluated the effect of speed of sound mismatch between the reconstruction algorithm and actual phantom values. Mismatches in this study did not exceed 1%; by acquiring images in the liquid filament phantom at 1470, 1486, and 1500 m/s ($\pm 1\%$ of the measured speed of sound in 1% Intralipid), we quantified variation in performance metrics with speed of sound mismatch. Axial resolution degraded by <10% for all transducers while lateral resolution degraded by <30%. Lateral resolution was more strongly affected due to

reconstruction defocusing, but observed trends and comparisons between transducers did not significantly change.

3.2 Elevational Resolution

Representative ultrasound and photoacoustic images of liquid and solid elevational phantoms are shown in Figs. 8 and 9 while Fig. 10 shows elevational resolution results. Ultrasound images showed focal depths that generally agreed with manufacturer-provided values. Also, PVCP results showed degradation of both ultrasound and PAI resolution with depth caused by acoustic attenuation, a tissue-relevant effect not captured in the liquid phantom. However, the reduced photoacoustic image penetration depth in the PVCP phantom limits observation of this effect. Photoacoustic images of the PVCP phantom showed nonuniform background due to near-field clutter. One of the most striking differences between photoacoustic images in liquid and PVCP phantoms was that the LSFs in liquid included strong, asymmetric lateral streaks, and thus required greater scan distances to completely capture the targets. While some streaking is expected due to the finite aperture of the acoustic lens, the prominent signal at greater distances is an artifact caused by out-of-plane sources. As the transducer and beam are scanned (from left to right in the images, illumination from the left), the
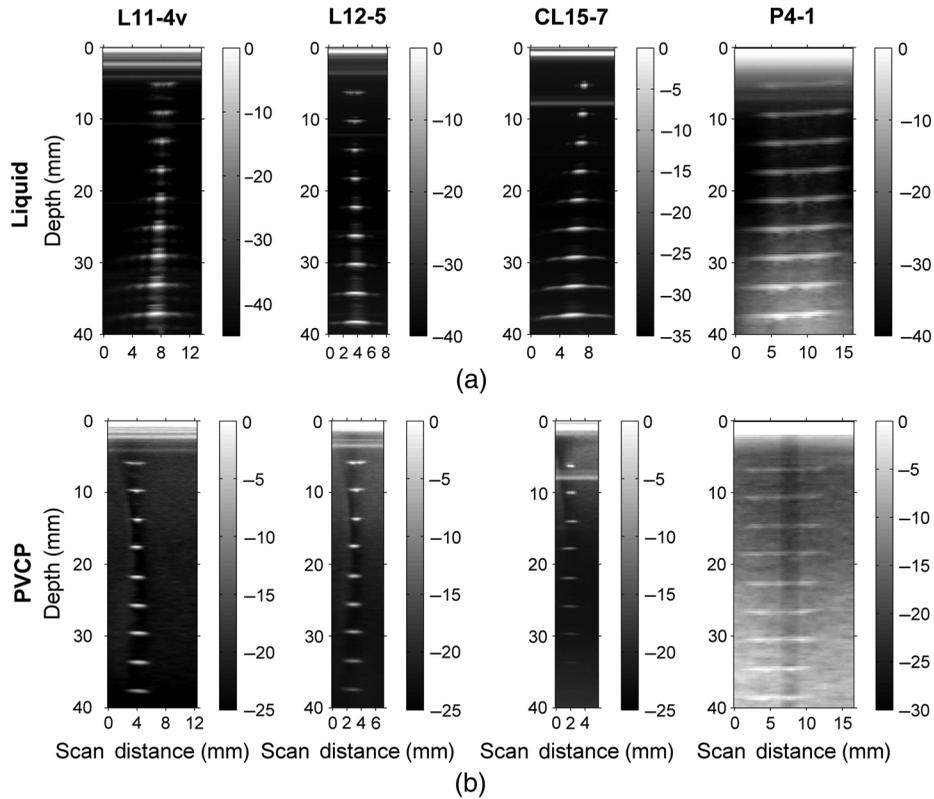
Fig. 8 Representative ultrasound images of the (a) liquid and (b) PVCP elevational resolution phantoms, acquired using, from left to right columns, L11-4v, L12-5, CL15-7, and P4-1 transducers. Scale bars in dB.
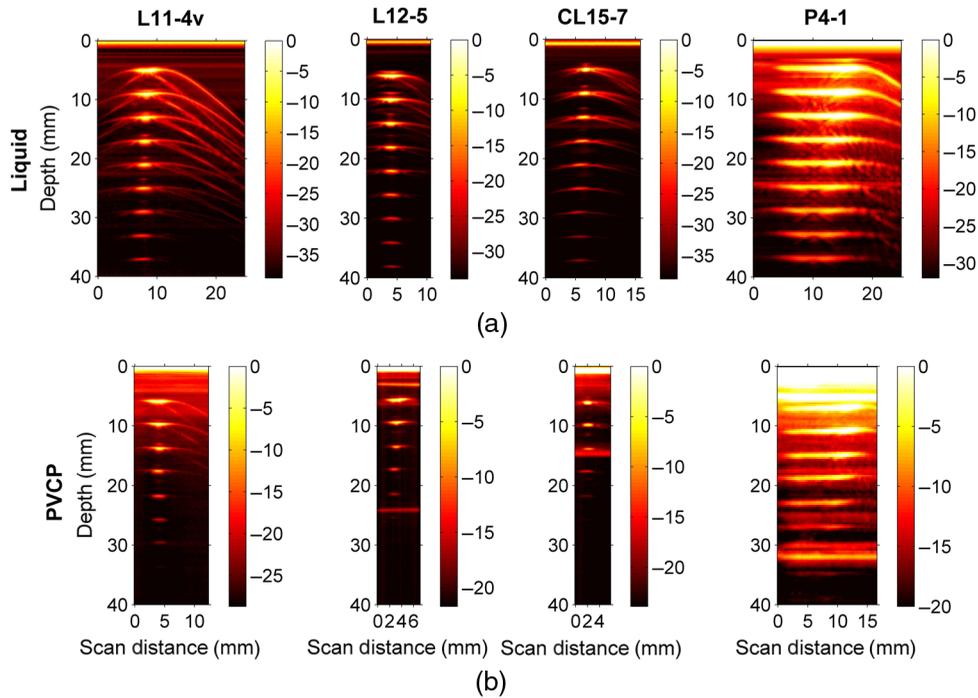
Fig. 9 Representative photoacoustic images of the (a) liquid and (b) PVCP elevational resolution phantoms, acquired using, from left to right columns, L11-4v, L12-5, CL15-7, and P4-1 transducers. Scale bars in dB.
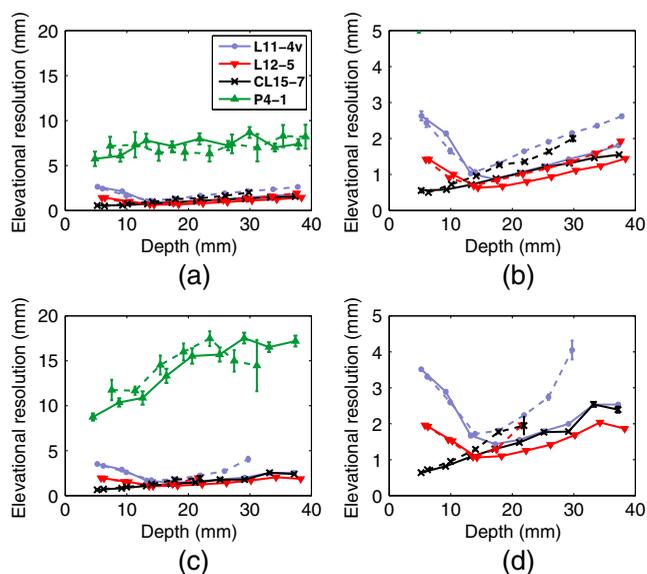
Fig. 10 (a) Ultrasound elevational resolution measurements (FWHM) versus depth in liquid (solid lines) and PVCP phantoms (dashed lines), with zoomed view shown in (b). (c) Photoacoustic elevational resolution measurements versus depth in liquid (solid lines) and PVCP phantoms (dashed lines), with zoomed view shown in (d).

targets continue to be illuminated and generate photoacoustic signals even when not located directly in the image plane. Because these out-of-plane waves travel a greater pathlength and the image reconstruction algorithm assumes a two-dimensional (2-D) imaging plane, the apparent target depth increases, resulting in the gradual descending streak artifact. These artifacts are substantially weaker in the PVCP phantom, most likely because steel wires produce much weaker photoacoustic signals than nylon sutures, although the higher acoustic attenuation of PVCP may also have influenced this effect. In a simple experiment using the L11-4v transducer, we acquired photoacoustic images of a 1% Intralipid phantom containing a nylon suture, steel wire, and 0.5-mm PTFE tube filled with India ink tuned to an absorption coefficient of 4 cm at 800 nm. The nylon and steel targets both displayed higher intensity than the tube (24.5 and 7.2 dB, data not shown), which implies their absorption is higher than biologically relevant values. Thus, the liquid phantom overestimates out-of-plane sensitivity as compared to a more biologically realistic phantom, but the PVCP-steel wire phantom may somewhat overestimate out-of-plane sensitivity compared to a real tissue environment. More extensive testing of out-of-plane sensitivity is beyond the scope of this study, but future out-of-plane sensitivity test methods should incorporate phantom inclusions with well-characterized, biologically relevant optical absorption.

3.3 Uniformity

Figure 11 shows mean intensity depth profiles for each transducer in both liquid and PVCP phantoms. Because images were log-compressed, intensity increased linearly with depth. Also, intensities were generally lower in the PVCP phantom, most likely due to weaker signal generation by steel wires and higher phantom acoustic attenuation compared with the liquid phantom. By plotting target intensity as a function of target grid position, 2-D uniformity maps were generated to evaluate uniformity within the entire target array (Fig. 12). While the
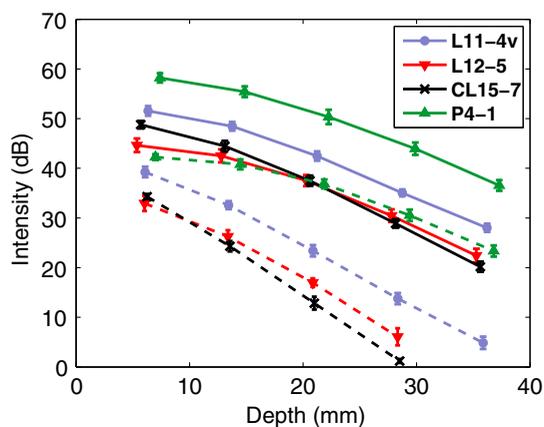
Fig. 11 Target intensity versus depth for log-compressed photoacoustic images in liquid (solid lines) and PVCP phantoms (dashed lines). Error bars denote 95% confidence intervals.

expected depth-dependent nonuniformity was observed in these maps, there was also significant lateral nonuniformity due to beam illumination geometry and light diffusion in the phantom. This lateral nonuniformity was highest for transducer arrays with greater length such as the L12-5 where the elliptical laser spot does not span the entire image plane.

3.4 Spatial Measurement and Coregistration Accuracy

Spatial measurement and PAI-ultrasound coregistration accuracy results for the PVCP filament array phantom are presented in Table 2. Results in Intralipid were not significantly different and are omitted for clarity. PAI- and ultrasound-based distance measurements were found to be in good agreement with each other and the reference target spacing values. Vertical and horizontal measurement precision was roughly correlated with transducer bandwidth. The P4-1 array exhibited the worst performance due to its low bandwidth and coarse element spacing. Spatial measurement accuracy was ultimately limited by pixel dimensions, which were large relative to the designed spacing values (1% to 4% vertical, 3% to 9% horizontal over all transducers). This resulted in average errors smaller than 1 pixel length, with low variation due to coarse sampling. Average coregistration errors were less than 1 pixel (~0.1 to 0.4 mm), but maximum coregistration errors of 1 to 2 pixels were observed, depending on the selected transducer and spatial direction (Table 2). Horizontal coregistration accuracy is limited by a discrete array element spacing, whereas vertical coregistration error may be caused by system timing jitter, resulting in temporal shifts in received pressure wave signals. From speed of sound mismatch testing as described in Sec. 3.1, mismatches of ±1% resulted in <1% variation in spatial measurement error and no significant variation in coregistration error.

3.5 Sensitivity

Representative photoacoustic images of the sensitivity phantom are shown in Fig. 13. There was significant clutter in all images, a result of photoacoustic signal generation at the transducer/phantom interface as well as in out-of-plane regions of the phantom. Additionally, horizontal bands were present due to reflections in the aluminum foil layer, while diagonal streaks were caused by photoacoustic signal generation near the lateral
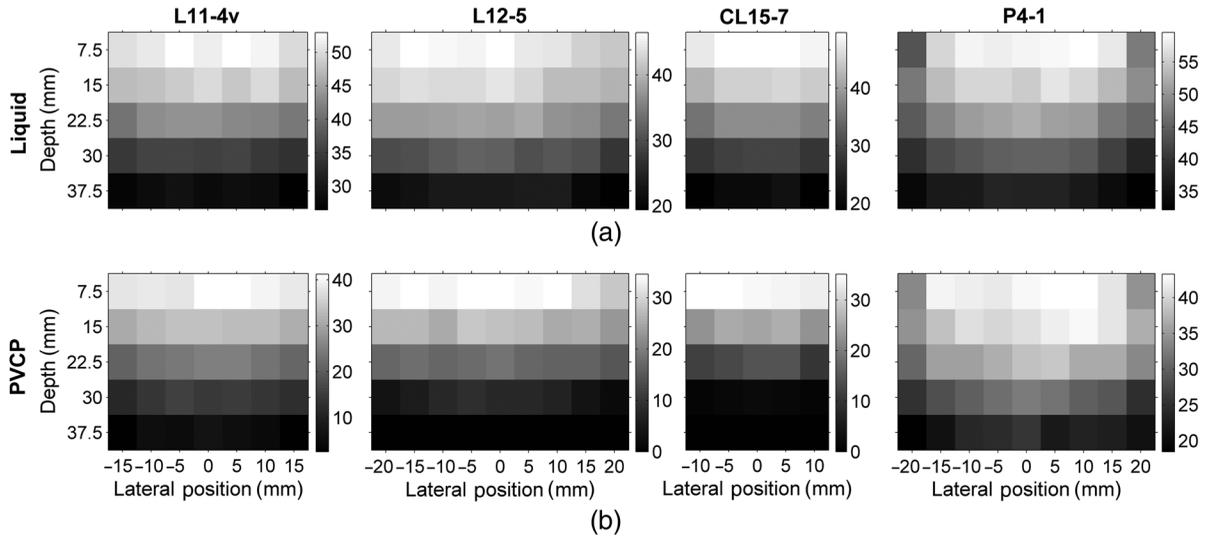
Fig. 12 Photoacoustic intensity uniformity maps for (a) liquid and (b) PVCP resolution phantom images acquired using, from left to right columns, L11-4v, L12-5, CL15-7, and P4-1 transducers. Scale bars in dB.

Table 2 PAI and ultrasound spatial measurement error and maximum PAI-ultrasound coregistration error over all targets for each transducer. Error ranges reported as 95% confidence intervals.

	L11-4v	L12-5	CL15-7	P4-1
US vertical spatial measurement error (mm)	0.07 ± 0.02	0.09 ± 0.02	0.10 ± 0.03	0.14 ± 0.04
US horizontal spatial measurement error (mm)	0.17 ± 0.03	0.16 ± 0.03	0.11 ± 0.05	0.28 ± 0.07
PA vertical spatial measurement error (mm)	0.10 ± 0.03	0.12 ± 0.03	0.14 ± 0.04	0.20 ± 0.06
PA horizontal spatial measurement error (mm)	0.16 ± 0.02	0.19 ± 0.04	0.16 ± 0.06	0.27 ± 0.08
Max vertical coregistration error (mm) (pixels)	0.095 (1)	0.189 (2)	0.142 (2)	0.592 (2)
Max horizontal coregistration error (mm) (pixels)	0.300 (1)	0.195 (1)	0.178 (1)	0.888 (2)
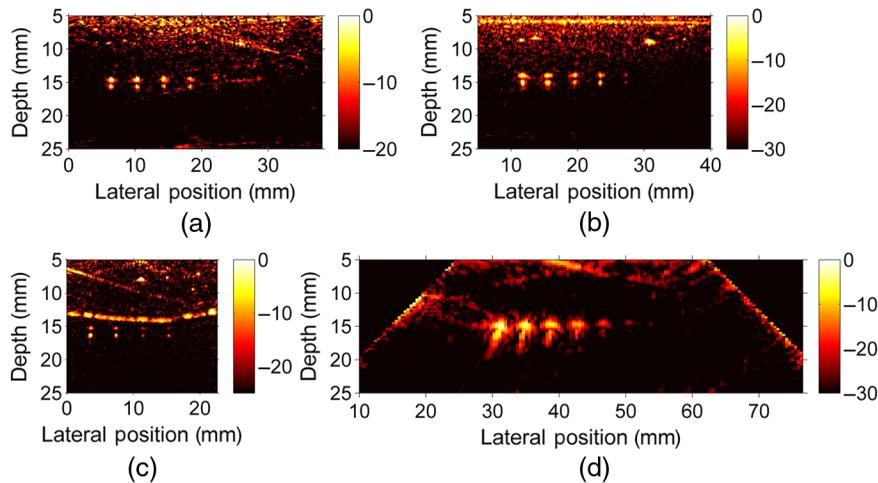
Fig. 13 Photoacoustic images of sensitivity phantom for (a) L11-4v, (b) L12-5, (c) CL15-7, and (d) P4-1 arrays. Scale bars in dB.
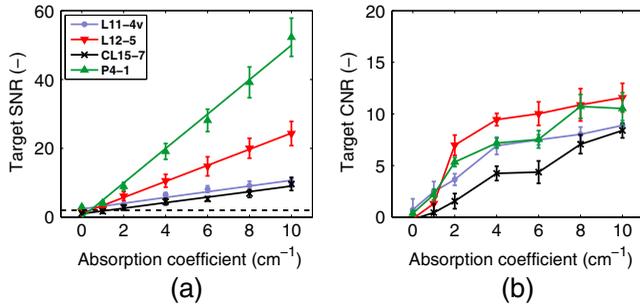
Fig. 14 Target (a) SNR and (b) CNR versus absorption coefficient for the four transducers. The black dashed line denotes SNR = 2 (6 dB). Error bars denote 95% confidence intervals.

ends of the transducer and foil layer. In P4-1 array images, ghost artifacts were seen below the targets; these are the results of photoacoustic waves being generated at the phantom surface, propagating into the phantom, and then being reflected off the tubes and back to the transducer. Since this pathlength is twice the tube depth, these waves are reconstructed below the tubes as the algorithm assumes one-way beam propagation. The presence of glass beads in PVCP phantoms increases scattering of these photoacoustic waves, which results in stronger near-field clutter signals. Removal of these artifacts is the subject of ongoing research by Singh and Steenbergen.[50]

Table 3 Minimum detectable absorption coefficients for each transducer determined using qualitative inspection or by interpolating data to an SNR threshold of 2 (6 dB).

	L11-4v	L12-5	CL15-7	P4-1
Qualitative limit (cm^{-1})	2	2	4	1
SNR threshold-based limit (cm^{-1})	0.94	1.17	2.3	0.61

Quantitative SNR and CNR results are shown in Fig. 14. An SNR threshold of 2 (6 dB) was found to correlate reasonably well to the detectability limits found by qualitative inspection, whereas CNR results suggested a CNR threshold of ~2 to 4. SNR data showed a linear dependence on target absorption coefficient, whereas CNR data showed nonlinear trends primarily resulting from log compression. Additionally, the P4-1 array showed the highest sensitivity based on SNR data, but the L12-5 produced CNR values similar to or higher than those of the P4-1 array. The L12-5 produced images with lower clutter than other arrays, which produced higher CNR values. This may imply that the L12-5 is better shielded against the laser beam or that the transducer better rejects out-of-plane signals, which is plausible due to its superior elevational resolution.

The minimum detectable absorption coefficient for each transducer was determined using either visual inspection or by interpolating SNR results for a threshold of 2 (6 dB) as shown in Table 3. Results suggested that transducers with lower center frequencies have greater sensitivity based on the minimum detectable absorption coefficient, which may be due to reduced phantom acoustic attenuation at lower frequencies and the higher low-frequency content of photoacoustic waves. However, the L11-4v showed lower sensitivity than expected based on transducer frequencies, as its SNR curve slope is lower than the L12-5 and is similar to the CL15-7. This was likely due to high clutter artifact strength, which increased background ROI variation in these images. Intensity nonuniformity over the tube array may also have affected the results; although uniformity data (Fig. 12) showed lateral variation of <3 dB for all transducers at a depth of 15 mm.

3.6 Penetration Depth

Representative photoacoustic images of the penetration depth phantom (Fig. 15) show similar clutter and ghost artifacts as in the sensitivity phantom. Also, the fundamental appearance of the target varies with transducer; due to band-limited sensing,
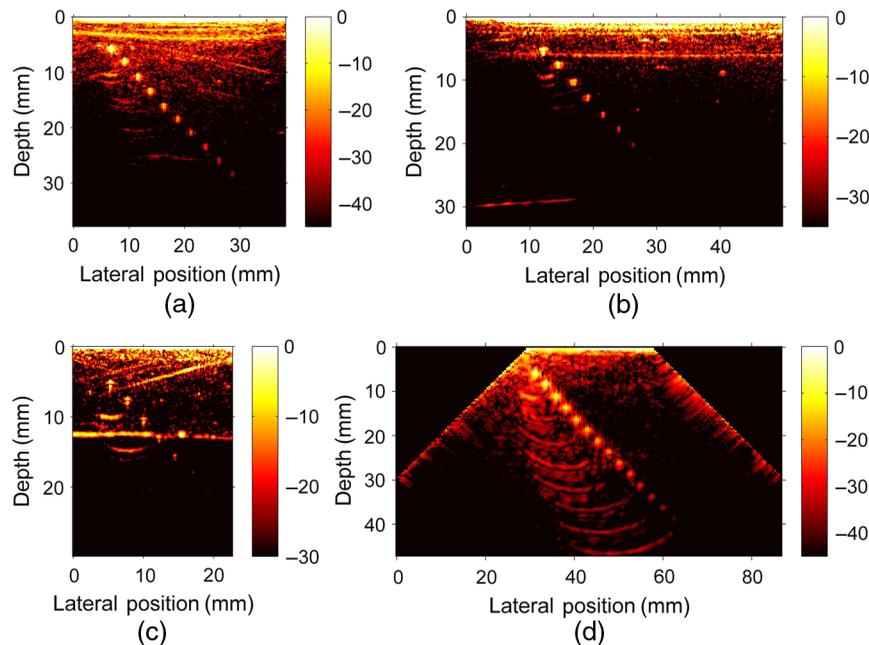
Fig. 15 Photoacoustic images of penetration depth phantom for (a) L11-4v, (b) L12-5, (c) CL15-7, and (d) P4-1 arrays. Scale bars in dB.
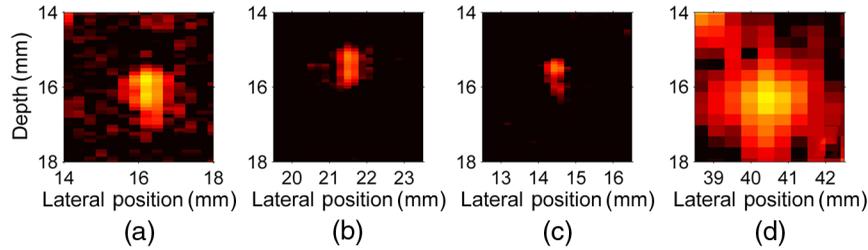
Fig. 16 Zoomed images of targets for (a) L11-4v, (b) L12-5, (c) CL15-7, and (d) P4-1 transducers, showing wall detection for all transducers except the P4-1.

all transducers except the P4-1 capture only the top and bottom walls of each target (Fig. 16). Because the P4-1 has poor axial resolution, the walls could not be resolved and these images showed a more homogeneously filled target. From Fig. 17, an SNR threshold of 2 again showed agreement with qualitative target detectability. However, SNR data exhibited a plateau effect for shallow depths resulting from increased background variation due to clutter. The highest SNR was achieved with the L12-5, but only for some shallow targets where target intensity was high but clutter was also minimal. Depth-varying clutter also strongly influenced CNR data by increasing both background ROI intensity and standard deviation. CNR data also show that while the penetration depth based on CNR does depend on transducer, the maximum CNR was consistently ~8 to 9 over all depths for each transducer. This may be the result of dynamic range optimization used to reduce clutter in log-compressed images. Table 4 shows comparison of maximum penetration depths determined using either qualitative inspection or by interpolating based on a threshold SNR of 2. Results indicated that lower frequency transducers achieve deeper penetration than high-frequency transducers due to the frequency-dependent acoustic attenuation of the phantom. However, a trade-off clearly exists between spatial resolution and penetration depth, as lower transducer frequency implies

lower bandwidth and thus worse axial resolution, as demonstrated in spatial resolution testing.

4 Discussion

Standardized image quality test methods are a critical component of product development, quality assurance, and regulatory decision-making, all of which often require direct performance comparison between design configurations or finalized device products. Our set of test methods was shown to enable objective, quantitative performance evaluation of different device configurations in a more comprehensive and rigorous manner than methods previously described in the literature. Test results illustrated well-known design trade-offs, such as differences in PAI target appearance caused by band-limited sensing artifacts or the trade-off between spatial resolution and sensitivity/penetration depth caused by frequency-dependent tissue acoustic attenuation and the broadband nature of photoacoustic signals. However, these tests provide quantification of such relationships that can better inform product design, quality assurance testing, clinical translation, and regulatory evaluation. Additionally, results showed that axial resolution was more strongly degraded with depth in PVCP phantoms as compared to Intralipid, suggesting that the use of a more realistic phantom may result in better prediction of real-world performance.

Our study builds upon and expands on previous work by others in several ways. First our test methods are designed to be broadly applicable to clinical linear-array PAI systems, whereas test methods described by others are often tailored for a specific system configuration, including tomographic preclinical animal imagers. Also, we demonstrate the importance of using phantom materials possessing biorealistic optical and acoustic properties and show how these properties influence device performance. Finally, our work highlights the important advantages of robust, temporally stable phantoms over liquid phantoms for image quality testing. As noted by others, stable phantoms are invaluable for constancy testing[33] and also enable device calibration, verification of system maintenance and repairs, and device unit comparison in multisite clinical trials.

The test methods developed in this study are meant to serve as a model for designing and implementing well-validated, rigorous approaches to photoacoustic image quality assessment, rather than to define a single set of test designs that should be applied to all PAI systems and applications. Although we chose to implement these test methods using a breast-mimicking PVCP TMM, the testing paradigm outlined here could be readily modified for other tissues of interest by substituting different application-specific TMM formulations and/or configurations of embedded inclusions with tissue-specific optical and acoustic properties. There is precedence for this in ultrasound image quality consensus documents: while some standards
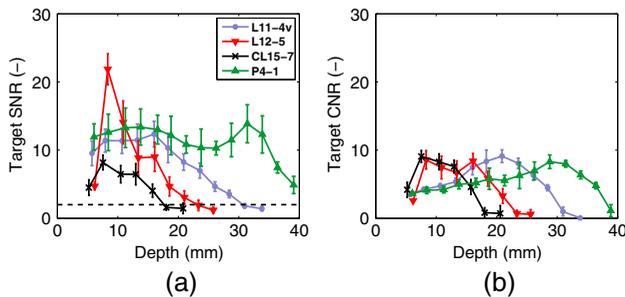
Fig. 17 Depth dependencies of target (a) SNR and (b) CNR for each transducer. The black dashed line denotes SNR = 2 (6 dB). Error bars denote 95% confidence intervals.

Table 4 Maximum penetration depth for each transducer determined using qualitative inspection or by interpolating data to a SNR threshold of 2 (6 dB).

	L11-4v	L12-5	CL15-7	P4-1
Qualitative limit (mm)	28.3	22.5	17.2	38.4
SNR threshold-based limit (mm)	28.6	21.4	16.8	>38.4

specify phantom properties that are generally representative of soft tissues,[16] others have been tailored to specific applications, such as prostate imaging.[51] Additionally, while the current study focused on clinical linear-array transducers, which are commonly used for PAI, it should be possible to adapt the testing paradigm described here for a wide variety of PAI systems, including configurations with larger or smaller fields of view, handheld probes with more complex illumination and detection geometries,[52] scanning photoacoustic tomography approaches,[53] and microscopy systems.[1] While we focused on macroscopic PAI systems, these phantoms and associated test methods could potentially be adapted for high-resolution PAI or photoacoustic microscopy systems. The high viscosity and short curing time of PVCP may pose challenges when molding small, solid inclusions, but PVCP could potentially be used as a biologically relevant background medium containing microscopic structures or fluid channels composed of other materials.

5 Conclusion

We have developed a robustly evaluated suite of phantom-based test methods for characterizing key aspects of photoacoustic image quality in an objective, quantitative, and reproducible manner. Using a modular PAI system, we demonstrated the utility of our four basic phantom designs as multifunctional tools for medical device development, specifically elucidating the effect of detection components on system performance. This study may serve as a foundation to guide future standardization of PAI system performance evaluation, which will improve public health by facilitating technological advancement, clinical translation, and regulatory evaluation of this emerging imaging modality.

Disclosures

The authors have no conflicts of interest, financial or otherwise, to disclose. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services.

References

1. L. V. Wang and S. Hu, "Photoacoustic tomography: in vivo imaging from organelles to organs," *Science* **335**(6075), 1458–1462 (2012).
2. S. Mallidi, G. P. Luke, and S. Emelianov, "Photoacoustic imaging in cancer detection, diagnosis, and treatment guidance," *Trends Biotechnol.* **29**(5), 213–221 (2011).
3. M. Heijblom et al., "Photoacoustic image patterns of breast carcinoma and comparisons with magnetic resonance imaging and vascular stained histopathology," *Sci. Rep.* **5**, 11178 (2015).
4. R. A. Kruger et al., "Photoacoustic angiography of the breast," *Med. Phys.* **37**(11), 6096–6100 (2010).
5. X. D. Wang et al., "Noninvasive imaging of hemoglobin concentration and oxygenation in the rat brain using high-resolution photoacoustic tomography," *J. Biomed. Opt.* **11**(2), 024015 (2006).
6. K. E. Wilson, T. Y. Wang, and J. K. Willmann, "Acoustic and photoacoustic molecular imaging of cancer," *J. Nucl. Med.* **54**(11), 1851–1854 (2013).
7. A. Dima and V. Ntziachristos, "Non-invasive carotid imaging using optoacoustic tomography," *Opt. Express* **20**(22), 25044–25057 (2012).
8. C. Haisch et al., "Combined optoacoustic/ultrasound system for tomographic absorption measurements: possibilities and limitations," *Anal. Bioanal. Chem.* **397**(4), 1503–1510 (2010).
9. P. van Es et al., "Initial results of finger imaging using photoacoustic computed tomography," *J. Biomed. Opt.* **19**(6), 060501 (2014).
10. R. G. M. Kolkman et al., "Photoacoustic imaging of blood vessels with a double-ring sensor featuring a narrow angular aperture," *J. Biomed. Opt.* **9**(6), 1327–1335 (2004).
11. E. Zhang, J. Laufer, and P. Beard, "Backward-mode multiwavelength photoacoustic scanner using a planar Fabry–Perot polymer film ultrasound sensor for high-resolution three-dimensional imaging of biological tissues," *Appl. Opt.* **47**(4), 561–577 (2008).
12. S. A. Ermilov et al., "Laser optoacoustic imaging system for detection of breast cancer," *J. Biomed. Opt.* **14**(2), 024007 (2009).
13. S. R. Kothapalli et al., "Deep tissue photoacoustic imaging using a miniaturized 2-D capacitive micromachined ultrasonic transducer array," *IEEE Trans. Biomed. Eng.* **59**(5), 1199–1204 (2012).
14. S. Manohar et al., "The Twente photoacoustic mammoscope: system overview and performance," *Phys. Med. Biol.* **50**(11), 2543–2557 (2005).
15. S. Vaithilingam et al., "Three-dimensional photoacoustic imaging using a two-dimensional CMUT array," *IEEE Trans. Ultrason. Ferroelect. Freq. Control* **56**(11), 2411–2419 (2009).
16. "Ultrasonics–pulse-echo scanners–part 1: techniques for calibrating spatial measurement systems and measurement of system point-spread function response," IEC 61391.1:2006, International Electrotechnical Commission, Geneva, Switzerland (2006).
17. "Ultrasonics–pulse-echo scanners–part 2: measurement of maximum depth of penetration and local dynamic range," IEC 61391-2:2006, International Electrotechnical Commission, Geneva, Switzerland (2006).
18. M. M. Goodsitt et al., "Real-time B-mode ultrasound quality control test procedures. Report of AAPM ultrasound task group no. 1," *Med. Phys.* **25**(8), 1385–1406 (1998).
19. American College of Radiology, "Phantom test guidance for the ACR MRI accreditation program," http://www.acraccreditation.org/modalities/mri (22 December 2016).
20. C. H. McCollough et al., "The phantom portion of the American College of Radiology (ACR) computed tomography (CT) accreditation program: practical tips, artifact examples, and pitfalls to avoid," *Med. Phys.* **31**(9), 2423–2442 (2004).
21. J. Pfefer and A. Agrawal, "A review of consensus test methods for established medical imaging modalities and their implications for optical coherence tomography," *Proc. SPIE* **8215**, 82150D (2012).
22. American College of Radiology, "Ultrasound accreditation program requirements," http://www.acraccreditation.org/modalities/ultrasound (22 December 2016).
23. B. Huang et al., "Improving limited-view photoacoustic tomography with an acoustic reflector," *J. Biomed. Opt.* **18**(11), 110505 (2013).
24. X. Li et al., "Intravascular photoacoustic imaging at 35 and 80 MHz," *J. Biomed. Opt.* **17**(10), 106005 (2012).
25. B. W. Pogue and M. S. Patterson, "Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry," *J. Biomed. Opt.* **11**(4), 041102 (2006).
26. H. X. Ke et al., "Performance characterization of an integrated ultrasound, photoacoustic, and thermoacoustic imaging system," *J. Biomed. Opt.* **17**(5), 056010 (2012).
27. C. Kim et al., "Deeply penetrating in vivo photoacoustic imaging using a clinical ultrasound array system," *Biomed. Opt. Express* **1**(1), 278–284 (2010).
28. M. Fonseca et al., "Characterisation of a phantom for multiwavelength quantitative photoacoustic imaging," *Phys. Med. Biol.* **61**(13), 4950–4973 (2016).
29. J. Kim et al., "Programmable real-time clinical photoacoustic and ultrasound imaging system," *Sci. Rep.* **6**, 35137 (2016).
30. W. J. Akers et al., "Multimodal sentinel lymph node mapping with single-photon emission computed tomography (SPECT)/computed tomography (CT) and photoacoustic tomography," *Transl. Res.* **159**(3), 175–181 (2012).
31. C. J. H. Ho et al., "Multifunctional photosensitizer-based contrast agents for photoacoustic imaging," *Sci. Rep.* **4**, 5342 (2014).

32. K. Daoudi et al., "Handheld probe integrating laser diode and ultrasound transducer array for ultrasound/photoacoustic dual modality imaging," *Opt. Express* **22**(21), 26365–26374 (2014).
33. S. E. Bohndiek et al., "Development and application of stable phantoms for the evaluation of photoacoustic imaging instruments," *PLoS One* **8**(9), e75533 (2013).
34. H. Zafar et al., "Linear-array-based photoacoustic imaging of human microcirculation with a range of high frequency transducer probes," *J. Biomed. Opt.* **20**(5), 051021 (2015).
35. M. S. Singh and H. B. Jiang, "Ultrasound (US) transducer of higher operating frequency detects photoacoustic (PA) signals due to the contrast in elastic property," *AIP Adv.* **6**(2), 025210 (2016).
36. J. Rebling et al., "Optoacoustic characterization of broadband directivity patterns of capacitive micromachined ultrasonic transducers," *J. Biomed. Opt.* **22**(4), 041005 (2017).
37. W. C. Vogt et al., "Biologically relevant photoacoustic imaging phantoms with tunable optical and acoustic properties," *J. Biomed. Opt.* **21**(10), 101405 (2016).
38. J. R. Cook, R. R. Bouchard, and S. Y. Emelianov, "Tissue-mimicking phantoms for photoacoustic and ultrasonic imaging," *Biomed. Opt. Express* **2**(11), 3193–3206 (2011).
39. W. F. Xia et al., "Poly(vinyl alcohol) gels as photoacoustic breast phantoms revisited," *J. Biomed. Opt.* **16**(7), 075002 (2011).
40. G. M. Spirou et al., "Optical and acoustic properties at 1064 nm of polyvinyl chloride-plastisol for use as a tissue phantom in biomedical optoacoustics," *Phys. Med. Biol.* **50**(14), N141–N153 (2005).
41. "American National Standard for safe use of lasers," ANSI Z136.1-2007, Laser Institute of America, Orlando, Florida (2007).
42. G. Held et al., "Effect of irradiation distance on image contrast in epi-optoacoustic imaging of human volunteers," *Biomed. Opt. Express* **5**(11), 3765–3780 (2014).
43. K. A. Wear, "Cancellous bone analysis with modified least squares Prony's method and chirp filter: phantom experiments and simulation," *J. Acoust. Soc. Am.* **128**(4), 2191–2203 (2010).
44. S. A. Prahl, M. J. C. Van Gemert, and A. J. Welch, "Determining the optical-properties of turbid media by using the adding-doubling method," *Appl. Opt.* **32**(4), 559–568 (1993).
45. J. M. Thijssen, G. Weijers, and C. L. de Korte, "Objective performance testing and quality assurance of medical ultrasound equipment," *Ultrasound Med. Biol.* **33**(3), 460–471 (2007).
46. R. Ma et al., "Multispectral optoacoustic tomography (MSOT) scanner for whole-body small animal imaging," *Opt. Express* **17**(24), 21414–21426 (2009).
47. K. Jansen et al., "Photoacoustic imaging of human coronary atherosclerosis in two spectral bands," *Photoacoustics* **2**(1), 12–20 (2014).
48. H. K. Walker, W. D. Hall, and J. W. Hurst, *Clinical Methods: The History, Physical and Laboratory Examinations*, 3rd ed., Butterworth-Heinemann, St. Louis, Missouri (1990).
49. P. A. Narayana and J. Ophir, "A closed form method for the measurement of attenuation in nonlinearly dispersive media," *Ultrason. Imaging* **5**(1), 17–21 (1983).
50. M. K. A. Singh and W. Steenbergen, "Photoacoustic-guided focused ultrasound (PAFUSion) for identifying reflection artifacts in photoacoustic imaging," *Photoacoustics* **3**, 123–131 (2015).
51. D. Pfeiffer et al., "AAPM Task Group 128: quality assurance tests for prostate brachytherapy ultrasound systems," *Med. Phys.* **35**(12), 5471–5489 (2008).
52. X. L. Dean-Ben et al., "Volumetric hand-held optoacoustic angiography as a tool for real-time screening of dense breast," *J. Biophotonics* **9**(3), 253–259 (2016).
53. J. Xia et al., "Whole-body ring-shaped confocal photoacoustic computed tomography of small animals in vivo," *J. Biomed. Opt.* **17**(5), 050506 (2012).

William C. Vogt received his BS degree in mechanical engineering from the University of Massachusetts Amherst in 2009 and his PhD in biomedical engineering from Virginia Polytechnic Institute and State University in 2013. Since 2013, he has been conducting photoacoustic imaging research as a research fellow at the Food and Drug Administration (FDA) in the Office of Science and Engineering Laboratories. His research interests include photoacoustic imaging, tissue phantoms, nanoparticles, and biophotonic medical device characterization and evaluation.

Congxian Jia received her BS and MS degrees in mechanical engineering from Beijing University, China, in 1999 and 2002, respectively. She also received her MS degree in biomechanics of aerospace and mechanical engineering from Boston University in 2004 and her PhD in biomedical engineering from the University of Michigan in 2010. Currently, she works as a research fellow at the FDA, and her research interests include ultrasound elasticity imaging and photoacoustic imaging.

Keith A. Wear received his BA degree in applied physics from the University of California, San Diego and his MS and PhD degrees in applied physics from Stanford University. He is the FDA Acoustics Laboratory Leader. He is an associate editor of *J. Acoust. Soc. Am.*, *Ultrason. Imag.*, and *IEEE Trans. Ultrason., Ferroelectr., Freq. Contr.* He is a fellow of Acoustical Society of America, American Institute for Medical and Biological Engineering, and American Institute of Ultrasound in Medicine.

Brian S. Garra trained at the University of Washington and the University of Utah and currently practices radiology at the Washington DC Veterans Affairs Medical Center. He also does research and medical device evaluation at the FDA as an associate director-clinical in the Division of Imaging, Diagnostics and Software Reliability. Currently, he is working on elastography, phantoms for photoacoustic system/elastographic system evaluation, the quantitative imaging biomarker alliance (QIBA), and ultrasound outreach in Peru.

T. Joshua Pfefer received his BS degree in mechanical engineering from Northwestern University, his MS degree in mechanical engineering, and his PhD in biomedical engineering from the University of Texas at Austin, and was a research fellow at the Wellman Laboratories of Photomedicine. In 2000, he joined the FDA, where he is currently the leader of the Optical Diagnostic Devices Laboratory. His group's research focuses on safety and effectiveness in emerging clinical biophotonic spectroscopy and imaging technologies.