Estimating blood oxygenation from photoacoustic images: can a simple linear spectroscopic inversion ever work?

Roman Hochuli
Lu An
Paul C. Beard
Benjamin T. Cox

SPIE.

Estimating blood oxygenation from photoacoustic images: can a simple linear spectroscopic inversion ever work?

Roman Hochuli, Lu An, Paul C. Beard, and Benjamin T. Cox*
University College London, Department of Medical Physics and Biomedical Engineering, London, United Kingdom

Abstract. Linear spectroscopic inversions, in which photoacoustic amplitudes are assumed to be directly proportional to absorption coefficients, are widely used in photoacoustic imaging to estimate blood oxygen saturation because of their simplicity. Unfortunately, they do not account for the spatially varying wavelength-dependence of the light fluence within the tissue, which introduces "spectral coloring," a potentially significant source of error. However, accurately correcting for spectral coloring is challenging, so we investigated whether there are conditions, e.g., sets of wavelengths, where it is possible to ignore the spectral coloring and still obtain accurate oxygenation measurements using linear inversions. Accurate estimates of oxygenation can be obtained when the wavelengths are chosen to (i) minimize spectral coloring, (ii) avoid ill-conditioning, and (iii) maintain a sufficiently high signal-to-noise ratio (SNR) for the estimates to be meaningful. It is not obvious which wavelengths will satisfy these conditions, and they are very likely to vary for different imaging scenarios, making it difficult to find general rules. Through the use of numerical simulations, we isolated the effect of spectral coloring from sources of experimental error. It was shown that using wavelengths between 500 nm and 1000 nm yields inaccurate estimates of oxygenation and that careful selection of wavelengths in the 620- to 920-nm range can yield more accurate oxygenation values. However, this is only achievable with a good prior estimate of the true oxygenation. Even in this idealized case, it was shown that considerable care must be exercised over the choice of wavelengths when using linear spectroscopic inversions to obtain accurate estimates of blood oxygenation. This suggests that for a particular imaging scenario, obtaining accurate and reliable oxygenation estimates using linear spectroscopic inversions requires careful modeling or experimental studies of that scenario, taking account of the instrumentation, tissue anatomy, likely sO_2 range, and image formation process. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JBO.24.12.121914]

1 Introduction

Blood oxygen saturation, sO_2, defined as the ratio of oxygenated to total hemoglobin concentration, is an important physiological indicator of tissue function and pathology and therefore has many clinical and preclinical applications. Examples of such applications include locating regions of brain activation,[1–3] and characterizing and staging tumors.[4,5] Optical methods for estimating blood oxygenation are also in common use, with pulse-oximetry[6] routinely used to monitor a patient's systemic blood oxygenation, and near-infrared (NIR) spectroscopy[7] and diffuse optical tomography[8] used, for example, for stroke characterization[9] and functional studies of brain activation in neonates.[10] Optical methods are appealing because the absorption spectrum of blood in the "NIR window," where water and blood absorption are low so the light can penetrate the tissue, is strongly dependent on the level of oxygenation. However, purely optical approaches are only able to provide low-resolution images of tissue oxygenation, beyond a superficial region, due to the highly scattering nature of most biological tissues.

Photoacoustic imaging does not suffer from this limitation. The basis of the photoacoustic contrast is optical absorption, so it retains the specificity of optical imaging techniques, but the image resolution can be much higher, as the information about the contrast is carried out of the tissue by ultrasonic waves, which undergo little or no scattering. Photoacoustic imaging can, therefore, provide high-resolution images of subcutaneous vasculature.[11–14] Now, because the source of photoacoustic contrast is optical absorption, there is a strong temptation to assume that the photoacoustic image is proportional to the absorption coefficient and thus, the photoacoustic spectrum is proportional to the linear sum of the absorption spectra of the absorbers present. If this were the case, then a simple linear inversion of the type used in conventional optical transmission tissue spectroscopy[15] could straightforwardly be employed to recover sO_2 from photoacoustic images acquired at different wavelengths. However, this is not the case because, while the photoacoustic image depends on the optical absorption coefficient for contrast, it is not proportional to it: a photoacoustic image depends on both the absorption coefficient spectrum and the local light fluence spectrum. Since the fluence spectrum in tissue is not generally flat, it is unreasonable to ignore this dependence and assume, without further analysis, that absorption coefficients or related quantities such as blood sO_2 can be obtained from photoacoustic images via a simple linear spectroscopic inversion. To give a practical example, the wavelength dependence of the photoacoustic source at a blood vessel will depend not only on the absorption spectra of the oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb) present in the vessel, but,

via the fluence, on the spectra of the absorbers and scatterers in the surrounding illuminated region; for example, the Hb and HbO_2 in the nearby microvasculature, water and lipids in the tissue, and melanin in the skin. This has the effect of distorting the photoacoustic spectrum, commonly known as *spectral coloring*, which is a well-known confounding factor that compromises sO_2 measurement accuracy, and there is already a large amount of literature describing attempts to ameliorate it.[16–23] Despite this, an increasing number of publications are appearing that do not take spectral coloring into account, and yet assume that the estimates of sO_2 that are obtained are trustworthy. It is particularly concerning that articles are appearing in which these questionable estimates of sO_2 are being used to make judgments in preclinical[24–36] and even clinical studies.[37–41] This trend has been exacerbated by the appearance of commercially available photoacoustic imaging platforms that appear to provide accurate estimates of sO_2, even though the justification for these claims and experimental validation is weak at best. It is therefore of paramount importance that the accuracy with which blood sO_2 can be estimated from photoacoustic images without making an explicit correction for the wavelength-dependent fluence is studied carefully. The current paper explores this by investigating whether sO_2 can be accurately estimated in the presence of spectral coloring using linear inversions in a simple numerical phantom, with specific emphasis on examining the impact of the choice of optical wavelength.

The structure of this paper is as follows. Section 2 presents the theory of how blood oxygen saturation is linked to photoacoustic images. Section 3 describes two requirements, which, if satisfied, ensure that the spectral coloring can be safely ignored for the purposes of estimating sO_2. In Sec. 4, the simple case of a single vessel embedded in a homogeneous absorbing and scattering medium is examined, with a series of modeling studies using a Monte Carlo (MC) model of light transport, to gain insight into the effects of spectral coloring. The necessary requirements for accurate sO_2 estimates are then applied to this situation to determine whether, even in this simple best case scenario, there are choices of wavelengths for which accurate sO_2 estimates are achievable without correcting for spectral coloring.

2 Estimation of Blood Oxygen Saturation from Photoacoustic Images

Photoacoustic tomography involves the generation of broadband acoustic pulses through rapid deposition of energy from visible or NIR nanosecond pulses of laser light. In tomography mode, the target tissue is flooded with light using wide-field illumination, and the resulting acoustic pulses are recorded at different spatial points on the tissue surface. There are then two steps to the image reconstruction. In the first step, the initial acoustic pressure distribution, p_0, is estimated from the measured pressure time series, for a number of optical wavelengths, λ. In the second step, tissue properties, such as optical absorption coefficients, or chromophore concentrations, or oxygen saturation, are extracted from the $p_0(\lambda)$ images. This paper is concerned with the effect of the light fluence distribution in the tissue on the estimation of oxygen saturation, sO_2, which is step 2, and so, it will be assumed that step 1 has been performed perfectly. This paper is therefore focused on the task of obtaining accurate estimates of sO_2 from a set of perfectly reconstructed images of $p_0(\lambda)$.

The initial acoustic pressure distribution p_0 is related to the spatial distributions of the light fluence Φ and tissue chromophore concentrations c_k by

$$p_0(\mathbf{x}, \lambda) = \kappa \Phi(\mathbf{x}, \lambda; c_k, \mu_s, g) \mu_a(\mathbf{x}, \lambda), \tag{1}$$

$$p_0(\mathbf{x}, \lambda) = \kappa \Phi(\mathbf{x}, \lambda; c_k, \mu_s, g) \sum_{k=1}^{K} \alpha_k(\lambda) c_k(\mathbf{x}), \tag{2}$$

where κ is a factor that takes into account the efficiency of photoacoustic generation and the detection sensitivity of the imaging system, $\mathbf{x}$ is the position, c_k is the concentration of the k'th chromophore, μ_s is the scattering coefficient, g is the anisotropy factor, μ_a is the absorption coefficient, and α_k is the molar absorption coefficient for the k'th chromophore. For a single voxel, and assuming the only chromophores that are optically absorbing at wavelengths $\lambda_1 \ldots \lambda_N$ are oxy- and deoxyhemoglobin, the wavelength dependence of the photoacoustic image amplitude can be represented in matrix form as follows:

$$\begin{bmatrix} p_0(\lambda_1) \\ \vdots \\ p_0(\lambda_N) \end{bmatrix} = \kappa \begin{bmatrix} \Phi(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Phi(\lambda_N) \end{bmatrix}$$
$$\times \begin{bmatrix} \alpha_{HbO2}(\lambda_1) & \alpha_{Hb}(\lambda_1) \\ \vdots & \vdots \\ \alpha_{HbO2}(\lambda_N) & \alpha_{Hb}(\lambda_N) \end{bmatrix} \begin{bmatrix} c_{HbO_2} \\ c_{Hb} \end{bmatrix}. \tag{3}$$

Blood oxygen saturation, sO_2, is defined as

$$sO_2 = \frac{c_{HbO_2}}{c_{HbO_2} + c_{Hb}} = \left(1 + \frac{c_{Hb}}{c_{HbO_2}}\right)^{-1}, \tag{4}$$

where c_{HbO_2} and c_{Hb} are the concentrations of oxy- and deoxyhemoglobin, respectively, and the second form makes it clear that sO_2 depends only on their ratio. The concentrations can formally be found from the photoacoustic amplitudes by inverting Eq. (3):

$$\begin{bmatrix} c_{HbO_2} \\ c_{Hb} \end{bmatrix} = \frac{1}{\kappa} \begin{bmatrix} \alpha_{HbO2}(\lambda_1) & \alpha_{Hb}(\lambda_1) \\ \vdots & \vdots \\ \alpha_{HbO2}(\lambda_N) & \alpha_{Hb}(\lambda_N) \end{bmatrix}^{\dagger}$$
$$\times \begin{bmatrix} 1/\Phi(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1/\Phi(\lambda_N) \end{bmatrix} \begin{bmatrix} p_0(\lambda_1) \\ \vdots \\ p_0(\lambda_N) \end{bmatrix}, \tag{5}$$

where $\dagger$ is used to indicate the pseudoinverse. [The constant κ has no bearing on the sO_2 estimates as it cancels out in Eq. (4).] Note that the matrix of wavelength-dependent fluence, $\Phi(\lambda_n)$, is unknown and must be estimated, which, in general, is a nontrivial task.

We might be tempted to assume, in the first instance, that spectral coloring does not exist so the fluence is not wavelength-dependent and $\Phi(\lambda_n)$ can be replaced by the identity matrix. The photoacoustic signal would then be proportional to the absorption coefficient, μ_a, and the problem becomes akin to transmission optical spectroscopy, in which the concentrations of Hb and HbO_2 can be calculated straightforwardly from the following simple linear inversion:

$$\begin{bmatrix} c_{HbO_2} \\ c_{Hb} \end{bmatrix} = \begin{bmatrix} \alpha_{HbO_2}(\lambda_1) & \alpha_{Hb}(\lambda_1) \\ \vdots & \vdots \\ \alpha_{HbO_2}(\lambda_N) & \alpha_{Hb}(\lambda_N) \end{bmatrix}^{\dagger} \begin{bmatrix} \mu_a(\lambda_1) \\ \vdots \\ \mu_a(\lambda_N) \end{bmatrix}. \tag{6}$$

Approximating the fluence matrix by a scaled identity matrix like this cannot, in general, be relied upon to give good estimates of sO$_2$ when using photoacoustic measurements. However, there may be specific circumstances, particular wavelengths, and imaging scenarios, in which this approximation does give usefully accurate estimates. The next section examines what these circumstances might be.

3 Requirements for Accurate Blood Oxygen Saturation Estimation

In order to gain some insight into the effect of the fluence on photoacoustic sO$_2$ estimation, the two-wavelength case will be examined in detail. The same general principles will apply to the use of multiple wavelengths, but focusing on the two-wavelength case allows a clearer exposition of the essential ideas.

3.1 Without Fluence Correction

The effect of fluence can be understood by noting that a change in amplitude at a point in a photoacoustic image, δp_0, due to a change in absorption coefficient there, $\delta \mu_a$ (in this case due to a change in wavelength) is given by

$$\delta p_0 = \kappa(\delta \mu_a \Phi + \mu_a \delta \Phi + \delta \mu_a \delta \Phi), \tag{7}$$

where $\delta \Phi$ is the change in fluence due to the change in absorption. Dividing through by p_0 and using Eq. (1) yields

$$\frac{\delta p_0}{p_0} = \frac{\delta \mu_a}{\mu_a} + \frac{\delta \Phi}{\Phi}, \tag{8}$$

where the small second-order term $(\delta \mu_a / \mu_a)(\delta \Phi / \Phi)$ has been neglected. When the sO$_2$ estimation proceeds by ignoring the effect of fluence, it will only be accurate when the relative change in the photoacoustic image is equal to the relative change in the absorption coefficient. Equation (6) can then be used for the inversion. From Eq. (8), it is clear that for this to be the case, i.e., for $\delta p_0 / p_0 \approx \delta \mu_a / \mu_a$, the following condition must hold.

- Requirement 1: Fluence difference is small. At every point of interest, the relative difference in the fluence at the two wavelengths should be much smaller than the relative change in absorption coefficient:

$$\left| \frac{\delta \mu_a}{\mu_a} \right| \gg \left| \frac{\delta \Phi}{\Phi} \right|. \tag{9}$$

At first glance, Eq. (9) might suggest that a good way to satisfy requirement 1 would be to choose wavelengths for which $|\delta \mu_a / \mu_a|$ is large. However, as will be demonstrated in the example in Sec. 4.2, this would be wrong because the two terms in Eq. (9) are not independent and a large change in absorption coefficient will result in a large change in fluence. It turns out that the best way to satisfy Eq. (9) is to choose wavelengths at which the absorption coefficient is similar, i.e., $\delta \mu_a$ is small. While such a choice minimizes spectral coloring, it introduces another problem. When there is only a small change in the

absorption coefficient and a consequent small change in fluence, there will be only a small change in the photoacoustic amplitude. It is therefore important to consider the effect of noise, specifically the noise in the difference image $\delta p(\mathbf{x})$. This leads to requirement 2.

- Requirement 2: Low noise. The noise in the difference image must be sufficiently low that the uncertainty that it gives rise to in the estimates of sO$_2$ is sufficiently small for the application under consideration.

Requirement 2 can be subdivided into two related requirements.

- Requirement 2a: High sensitivity. The sensitivity of the imaging device must be sufficiently high (i.e., the instrument noise level must be sufficiently low) that the noise in the two images $p_0(\lambda_1)$ and $p_0(\lambda_2)$, and therefore, the noise in the difference image δp is low enough that the level of the resulting uncertainty in sO$_2$ due to the noise is sufficiently small.

- Requirement 2b: Good conditioning. The wavelengths must be chosen so that the molar absorption coefficient matrix in Eq. (6) is sufficiently well-conditioned that the level of the uncertainty in sO$_2$ due to the noise is sufficiently small.

Requirement 2b puts a bound on the condition number of the molar absorption coefficient matrix given the noise level and the accuracy required in the sO$_2$ estimates. In an extreme case, the molar absorption coefficient matrix will be singular (have no inverse). For example, when the wavelengths have been chosen at two isosbestic points for hemoglobin (e.g., 586 nm and 808 nm; see Fig. 4), δp_0 will be the same for all oxygenation levels, so there is no way to determine, given a measured δp_0, what the corresponding value of sO$_2$ should be. There is no sensitivity to sO$_2$ for this choice of wavelengths. The requirement that the matrix is not singular is another way of saying that δp_0 must map uniquely onto sO$_2$. In a less extreme case, but one in which the condition number is large, the inversion will be very sensitive to the noise in the images, so even a low level of noise could lead to uncertainties in the sO$_2$ estimates that are greater than its full range (0% to 100%), rendering them meaningless. An example of such a case would be where the molar absorption coefficient of Hb is almost proportional to the molar absorption coefficient of HbO$_2$ at the chosen wavelengths, i.e., $\alpha_{Hb}(\lambda_n) \approx K \alpha_{HbO_2}(\lambda_n)$, for the chosen wavelengths λ_n and some constant K. This could arise, for example, if the two wavelengths were chosen to lie between 875 and 925 nm, where the spectral characteristics of Hb and HbO$_2$ are similar (see Fig. 4). In this wavelength range, the measured δp_0 is only weakly dependent on blood oxygenation making accurate sO$_2$ estimates very challenging in the presence of noise. In Ref. 42, maximizing the condition number was used as the sole criterion for choosing the optimal wavelengths. The risk of considering only condition number (which is to say the sensitivity to noise) when choosing the wavelengths, and failing also to take into account the effect of spectral coloring, is that the recommended wavelengths could well be those at which spectral coloring has the greatest deleterious effect on the accuracy.

In summary, requirement 1 must be satisfied to avoid the estimate of sO$_2$ suffering from spectral coloring. Requirement 2 must be satisfied to avoid the estimate of sO$_2$ being dominated

by noise. Both requirements 1 and 2 apply to every image voxel independently and must be satisfied if sO_2 estimates from photoacoustic images are to be considered accurate, when no attempt has been made to correct for the unknown fluence spectrum and a linear inversion [Eq. (6)] used.

3.2 With Fluence Correction

There may be situations in which an approximate correction for the spectral coloring can be incorporated into the routine to estimate sO_2. For example, an estimate of the fluence spectrum from a measurement or a model can be divided out of the measured photoacoustic amplitudes to give a measured photoacoustic spectrum closer to the absorption coefficient spectrum. To obtain accurate estimates of sO_2 when using a fluence correction, similar conditions to those above apply.

- Requirement 1: The condition in Eq. (9) still applies if the fluence difference $\delta\Phi$ is replaced by $\delta\Phi'$, the corrected fluence difference:

$$\delta\Phi' := \frac{\Phi(\lambda_1)}{\Phi^{estimated}(\lambda_1)} - \frac{\Phi(\lambda_2)}{\Phi^{estimated}(\lambda_2)}. \quad (10)$$

If the fluence correction is accurate at both wavelengths, then $\delta\Phi' = 0$ and this requirement is always satisfied.

- Requirement 2: This requirement still applies, although the relevant matrix in Requirement 2b is now the "fluence-weighted" molar absorption coefficient matrix (the molar absorption coefficient matrix multiplied by the diagonal matrix containing the estimated fluence spectrum). Also, the relevant noise level is now that of the difference between the "fluence-corrected" p_0 images, which will likely be different from the uncorrected case.

3.3 Multiwavelength Case

The two-wavelength case was studied above so that the essential requirements for accurate estimation of sO_2 could be clearly identified. When multiple wavelengths are used, equivalent requirements apply: it is still necessary that the fluence variation be minimized between the wavelengths, it is still necessary that the molar absorption coefficient matrix be well-conditioned, and it is still necessary that the noise level in the measurements are sufficiently low.

4 Example: Estimating Blood Oxygen Saturation in a Single Isolated Vessel

To demonstrate the above general principles, we will describe a simple numerical example consisting of a blood-filled tube immersed in a homogeneous, absorbing and scattering background. The use of an unadorned phantom such as this permits close investigation of the effect of spectral coloring from both inside and outside the tube region, from which insights into the effect may be drawn. As the focus of this paper is on the effect of fluence, it is important to isolate its effect from any other possible causes of error. With experimental data, there may be artifacts caused by the choice of image reconstruction algorithm, limited detection aperture, acoustic band-limiting, positional uncertainty of sources and sensor, sensor directivity, heterogeneous tissue acoustic properties, and so on. The effects of

these uncertainties on sO_2 estimates are not considered in this paper, which should, therefore, be taken as a best case scenario.

In Sec. 4.2, images obtained at 26 equally spaced wavelengths in the range 500 to 1000 nm were used simultaneously in the inversion for sO_2. Section 4.3 explores whether the sO_2 estimates could be improved using pairs of these wavelengths rather than all at once, as only two wavelengths are needed to make Eq. (5) invertible, and often the use of just two wavelengths is a practical requirement.

Two cases will be examined in this paper: (1) the effect on estimates of sO_2 ignoring spectral coloring, i.e., setting $\Phi(\lambda) \propto \mathcal{I}$, the identity matrix, and (2) the effect on estimates of sO_2 approximately correcting for spectral coloring using an expression for the depth-dependent decay of the form,

$$\Phi(\lambda, z) = \exp[-\mu^{fit}(\lambda)z], \quad (11)$$

where z is the depth from the tissue surface and $\mu^{fit}(\lambda)$ is obtained by fitting Eq. (11) to the function of z obtained by taking the projection of p_0 in the lateral directions. The reason for examining this latter case is that such an approximate correction is expected to improve the sO_2 estimates, is simple and easy to implement, and is quite commonly used in one form or another (if only as a form of time-gain compensation to improve image visualization with depth). The former approach will be referred to in shorthand as an "uncorrected" and the latter as "corrected," although, of course, the correction is not perfect.

4.1 Digital Phantom

The cubic $5 \times 5 \times 5$ mm domain had cubic 25 μm voxels and index-matched, optically absorbing boundaries. A blood-filled tube, 250 μm in diameter, was embedded in the domain at a depth of 1 mm parallel to the illumination surface. The domain was illuminated with a collimated 2.5 mm diameter column of light (top-hat profile). The MC model of light transport used 10^{12} photons, which was determined to be sufficient to ensure sO_2 estimates were not impacted by MC noise (see Sec. 6). The arrangement is shown in Fig. 1.

The optical properties were chosen to be representative of soft biological tissue.[43] The blood within the tube had a total hemoglobin concentration $c_{HbT} = 150$ g l^{-1} and oxygen saturation of 90% ($c_{HbO2} = 135$ g l^{-1} and $c_{Hb} = 15$ g l^{-1}). Two cases
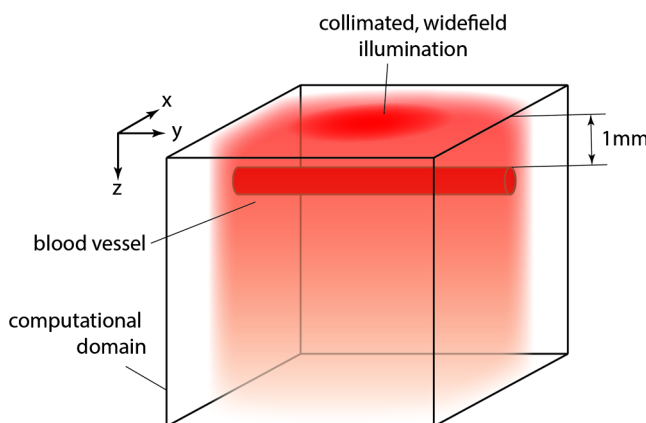
Fig. 1 Digital vessel phantom: $5 \times 5 \times 5$ mm computational domain with a blood vessel 1-mm beneath the surface and collimated widefield illumination beam of 2.5 mm diameter. The coordinate system used throughout the paper is also shown.

580nm illumination 600nm illumination
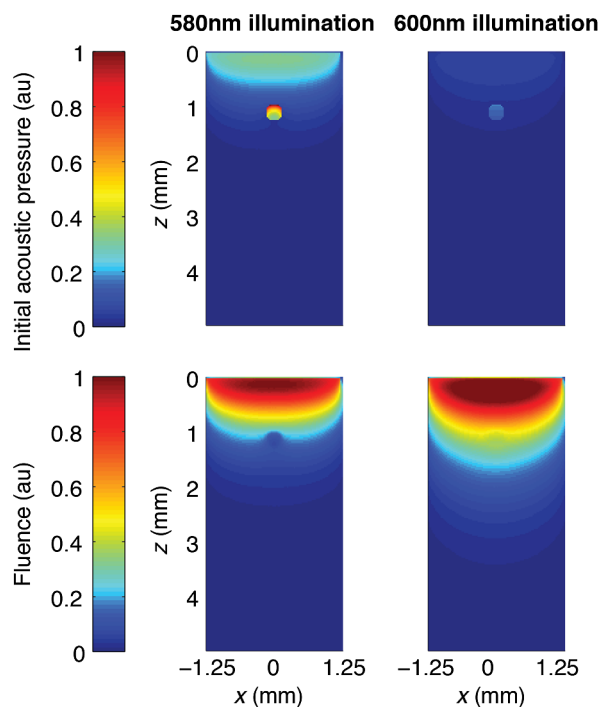
Fig. 2 Spectral coloring. Top row: MIP of initial acoustic pressure in tube phantom with spectrally varying background illuminated with top hat profile at 580 and 600 nm, normalized by maximum acoustic pressure value. Bottom row: MIP of fluence in tube phantom with spectrally varying background illuminated with top hat profile at 580 and 600 nm, normalized by maximum fluence value.

were considered for the background optical properties: (a) no spectral dependence and (b) spectral dependence typical of a capillary bed, modeled as 55% water and 45% blood with a total hemoglobin concentration, $c_{\mathrm{HbT}}^{\mathrm{bg}} = 5.63$ g l^{-1} and 60.7% oxygenation ($c_{\mathrm{HbO2}} = 3.42$ g l^{-1} and $c_{\mathrm{Hb}} = 2.11$ g l^{-1}). The scattering coefficient throughout the domain followed the relationship 21.3 mm$^{-1}\left(\frac{\lambda}{500\ \mathrm{nm}}\right)^{-1.2}$, and the Henyey–Greenstein phase function was used, with an anisotropy factor of $g = 0.9$ for the background and $g = 0.9945$ for blood. The initial pressure distribution was calculated using Eq. (1) with κ set to 1 with no loss of generality.

Figure 2 shows examples of the initial acoustic pressure and fluence distributions at wavelengths of 580 and 600 nm with a spectrally varying background [case (b)]. Both are maximum intensity projections in the y direction and are normalized by the maximum value of acoustic pressure or fluence, respectively. Spectral coloring is due to the wavelength dependence of the fluence distribution, which can be clearly seen in this example, as the fluence distributions at 580 and 600 nm are notably different, especially in the region of the tube. This difference is due to the absorption coefficient of blood being an order of magnitude greater at 580 nm than at 600 nm, and it would not be surprising if it were to impact on the ability to accurately estimate oxygen saturation using these wavelengths.

4.2 *Estimates using 26 Wavelengths over the 500- to 1000-nm Range*

As a starting point for this investigation, images were obtained at 26 wavelengths evenly spaced at 20-nm intervals between 500 and 1000 nm. These were used simultaneously to estimate sO$_2$

using Eq. (5) with $\Phi \propto \mathcal{I}$. Both the cases of (a) no spectral dependence in the background and (b) spectral dependence typical of tissue were studied.

Figure 3 shows a slice in the $x - z$ plane, halfway along the y axis, of the error in the resulting sO$_2$ estimates in the tube for both the spectrally flat (upper plot) and spectrally varying (lower plot) backgrounds. (Throughout the paper, error will be calculated as $|\mathrm{sO}_2^{\mathrm{true}} - \mathrm{sO}_2|$.) The estimates in the tube differ voxel-to-voxel, and the histograms show how the distribution of the oxygenation estimates varies with depth. The voxels included in the histograms are those within the tube in the $x - y$ slices at four depths (a, b, c, d), where each $x - y$ slice is orthogonal to the $x - z$ slice shown. The sO$_2^{\mathrm{mean}}$ values are the means of the associated histograms. The global mean sO$_2^{\mathrm{g.mean}}$ is the mean of the sO$_2$ estimates in all voxels in the tube.

The histograms of the error in the sO$_2$ estimates show that, even in the case of a spectrally flat background, the mean sO$_2$ is at best 75% (compared to the true value of 90%) and the accuracy decreases rapidly with depth through the tube. As there is no spectral coloring from the background absorbers in this case, it is due wholly to the presence of the blood within the tube. These results provide a compelling illustration of the adverse influence of spectral coloring. They suggest that even under these highly benign conditions of a spectrally flat background, reasonable accuracy, without any fluence correction, is only possible if the spatial resolution of the imaging instrument is sufficient to resolve those few voxels that lie in the center of the $x - y$ plane within the uppermost region (slice "a"). Any spatial averaging, particularly in the z direction, due to the resolution limitations of a practical imaging instrument, will rapidly produce very significant errors. For the spectrally varying background (Fig. 3 lower plot), the situation is even worse to the extent that accuracy is significantly compromised even for the "best" voxels in the center of the most superficial slice "a." The error is larger because of spectral coloring by the background region as well as the blood within the tube itself.

The results in Fig. 3 also show that there is a tendency to underestimate sO$_2$ relative to the true value of 90%. This is the result of the inversion being weighted by wavelengths, where the change in fluence between illumination wavelengths is large. It can be seen in Fig. 4, which shows the spectra of the absorption coefficient in the tube and the fluence averaged over the tube volume that the change in fluence is most significant between 500 and 600 nm. (In a separate set of simulations, not shown here, linear inversions only using wavelengths in the 500- to 600-nm range always resulted in the underestimation of sO$_2$ relative to the true value of 90%, whereas when only wavelengths greater than 600 nm were used, sO$_2$ was overestimated.)
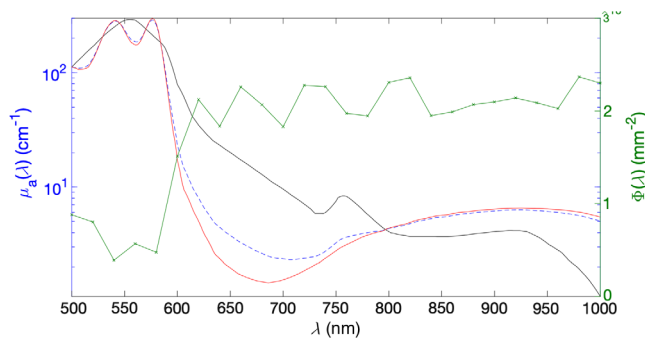
Estimating sO$_2$ over this wavelength range without accounting for the spectral coloring of the fluence is clearly not likely to give accurate values. To make a first-order correction for the fluence while retaining the simplicity of the linear inversion, the depth-dependent fluence correction, Eq. (11), was applied to the dataset obtained with a spectrally varying background. The resulting estimates are shown in the lower plots in Fig. 3. Comparing the corrected and uncorrected histograms, it is clear that the sO$_2$ estimates for this case are improved by 20% to 30% and are now comparable to the case with a spectrally flat background. Although the fluence correction provides a clear benefit, the accuracy remains, at best, only comparable to that achieved in the spectrally independent case with acceptable
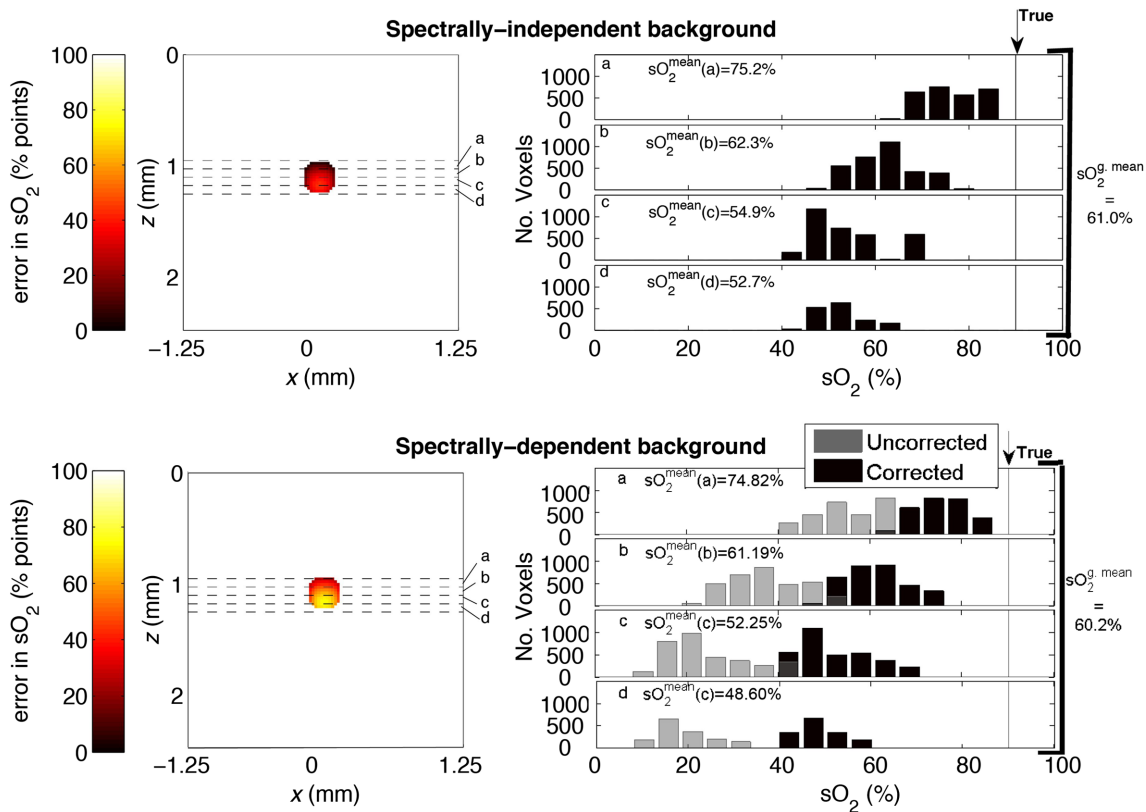
Fig. 3 sO$_2$ estimates using 26 wavelengths between 500 and 1000 nm. Top row: $x - z$ slice of error in oxygenation midway along tube's length when the background is spectrally flat and the error originates from spectral coloring of the fluence in the blood-filled tube. Dashed lines used to indicate four $x - y$ slices (a–d) from which histograms of sO$_2$ are formed (true value of 90% indicated by black line). Bottom row: $x - z$ slice of error in oxygenation midway along tube's length when the background optical properties vary with wavelength. Both the cases with and without a depth-dependent fluence correction are shown. Dashed lines used to indicate four $x - y$ slices (a–d) from which histograms of sO$_2$ are formed (true value of 90% indicated by vertical black line). (a–d) sO$_2^{mean}$ represent the mean of the sO$_2$ estimates of the voxels contained within the tube in slices a, b, c, and d, respectively, for the corrected data. The global mean sO$_2^{g.mean}$ of the corrected data, which represents sO$_2$ estimates averaged over all voxels enclosed within the entire tube, is shown to the right of the histograms.

accuracy (∼6%) only achieved for those voxels in the center of the $x - y$ plane in the most superficial slice "a."

In summary, the results in this section show that, when using 26 wavelengths in the range 500 to 1000 nm, spectral coloring

has a highly deleterious impact on the accuracy of the sO$_2$ estimates, to the extent that acceptable accuracy is unlikely to be achieved in practice using this wavelength range. However, Fig. 4 indicates that the error is likely to be most significant between 500 and 600 nm, where the wavelength dependence of the fluence is greatest. This suggests that, in contrast to conventional optical transmission spectroscopy, there may be an advantage to selecting wavelengths where the absorption of blood varies slowly with wavelength (which here is also where the absorption coefficient is low) because this will minimize spectral coloring. For this reason, in the next section, consideration is given as to whether more accurate estimates could be obtained by selecting wavelengths beyond 600 nm, where blood absorption and its spectral dependence is relatively low.[44]

4.3 Optimal Wavelength Pairs in the 500- to 1000-nm Range

In general, when estimating some quantity from experimental measurements, it is often reasonably assumed that the more data available the better the estimate will be.[45] The implicit intuition is that the effect on the estimate of the errors in the individual instances of the data will be symmetrically distributed about the

Fig. 4 Plot of absorption spectrum in the tube (dashed blue) for 90% oxygenated blood and average fluence spectrum (green) in the tube. Also shown are the spectra for fully oxygenated blood (red) and deoxygenated blood (black). Fluence simulated using MC with 10^{12} photons, collimated, illuminated 2.5 mm in diameter at wavelengths in 500- to 1000-nm range, and averaged over tube volume.

true value and so "cancel out," resulting in an estimate whose accuracy increases with the amount of data used. Here, while using more wavelengths may improve the SNR and may improve the conditioning of the multiwavelength inversion, the effect of spectral coloring will remain and could counteract both these advantages. Unfortunately, the coloring that the fluence gives the photoacoustic spectra is not guaranteed to cancel out, and so using images at more wavelengths does not guarantee a better estimate of sO_2 when spectral coloring is not corrected for. Following this reasoning, using fewer wavelengths can potentially give more accurate estimates, but only if the spectral coloring at the wavelengths chosen is low. This will be examined in this section, but rather than investigating all possible combinations of wavelengths, pairs of wavelengths will be investigated. This has two advantages: first, the practical advantage that it is a manageable number of sets of wavelengths, and second, the advantage that the simplicity of using two wavelengths allows insights to be gleaned, which might be harder to see were larger groups of wavelengths used.

In order to demonstrate that the wavelengths that give the best estimates of sO_2 are those that satisfy requirement (1), sO_2 was estimated using Eq. (5) for all possible pairs of the 26 wavelengths in the 500- to 1000-nm range, i.e., $^{26}C_2 = 325$ wavelength pairs. Three cases were examined: (a) spectrally flat background and no fluence correction ($\Phi \propto \mathcal{I}$), (b) spectrally varying background and no fluence correction, and (c) spectrally varying background and the depth-dependent correction [Eq. (11)]. Figure 5 shows histograms of the sO_2 estimates for these three cases, for just the pairs of wavelengths that gave a mean sO_2 closest to the true value. The voxels that satisfy the condition in Eq. (9) that the relative fluence change be much smaller than the relative absorption change (here interpreted as 2 orders of magnitude smaller) are highlighted in red.

Comparing Figs. 3 and 5, it is evident that these wavelength pairs provide significantly improved accuracy compared to using all 26 wavelengths across the 500- to 1000-nm wavelength range. The global mean sO_2 values in all cases are within 2% of the true value of 90% in Fig. 5, whereas the global mean sO_2 values achieved when using the 500- to 1000-nm spectral range are subject to an error of almost 30% (Fig. 3). When the

background is spectrally flat, Fig. 5(a) shows that sO_2 is accurately estimated (to within 0.5%) and that estimates made under the condition that the relative change in the absorption coefficient far exceeds the accompanying relative change in fluence are more accurate. These estimates correspond to regions at the surface of the tube and at the center of the $x - y$ slice. When the background has some spectral dependence, sO_2 estimates are less accurate with errors of 1% to 2%, and much higher variance, as shown Fig. 5(b). It can be observed that there are no voxels that satisfy Eq. (9) in Fig. 5(b) due to spectral coloring being significant at 700/820 nm. Figure 5(c) shows that the use of a depth-dependent fluence correction shifts estimates to a global mean value of 88.8% in the tube, and again the values closest to the true value are those that satisfy Eq. (9).

Given that in most practical settings the ratio $(\delta\Phi/\Phi)/(\delta\mu_a/\mu_a)$ is unknown, and therefore cannot be used directly as an indicator of the accuracy of the sO_2 estimate, it is desirable to obtain a result that is of more general applicability. One could, for example, determine the wavelength pairs (λ_1, λ_2) that yield the most accurate sO_2 estimates on average over the tube, thereby providing some indication of which wavelengths are best-suited to recovering sO_2. Figure 6 illustrates this by showing the error in oxygenation, averaged over the tube, obtained at a range of wavelengths λ_1 and λ_2 between 500 and 1000 nm when the background was spectrally varying. Four different tube sO_2 values were considered. The bottom-left triangles show estimates obtained without the depth-dependent fluence correction, whereas the top-right triangles contain values obtained when the data were corrected using Eq. (11).

Several observations emerge from the results in Fig. 6. First, it is noticeable that there are two distinct spectral regions "a" and "b," indicated by white-colored pixels, where the error is particularly high ($>100\%$). The error is large for region "a" because it encompasses the spectral range of 500 to 600 nm, where blood absorption is strongly wavelength-dependent, resulting in significant spectral coloring. Region "b" represents the spectral range of 820 to 900 nm and exhibits poor accuracy for a different reason, namely, poor conditioning of the molar absorption coefficient matrix as the spectra of Hb and HbO_2 vary almost in proportion, i.e., the ratio $\alpha_{Hb}/\alpha_{HbO_2}$ is approximately constant in
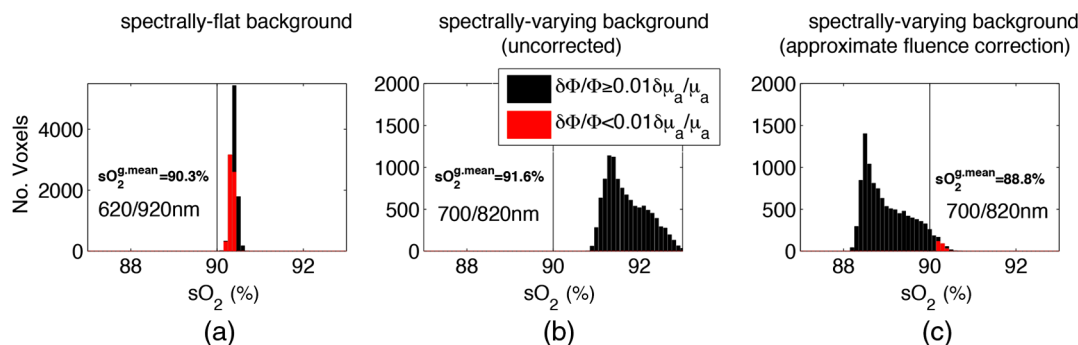
Fig. 5 Histograms of sO_2 estimates in each voxel within the tube for (a) spectrally flat with mean sO_2 of 90.3%, (b) spectrally varying background with mean sO_2 of 91.6%, and (c) spectrally varying with depth-dependent fluence correction with mean sO_2 of 88.8% obtained using linear inversion at 620/920 nm, 700/820 nm, and 700/820 nm, respectively; sO_2 estimates for which the ratio $(\delta\Phi/\Phi)/(\delta\mu_a/\mu_a) < 0.01$ shown in red and $(\delta\Phi/\Phi)/(\delta\mu_a/\mu_a) \geq 0.01$ in black. Notes: vertical line indicates true oxygenation of 90%; there are no sO_2 estimates for which $(\delta\Phi/\Phi)/(\delta\mu_a/\mu_a) < 0.01$ for the spectrally varying background (uncorrected) due to severity of spectral coloring of fluence; note the scale change on vertical axis between flat and various backgrounds. These wavelength pairs yielded the most accurate sO_2 estimates of all pairs examined. $sO_2^{g.mean}$ represents the mean of the sO_2 estimates all voxels enclosed within the entire tube and is termed the global mean sO_2.
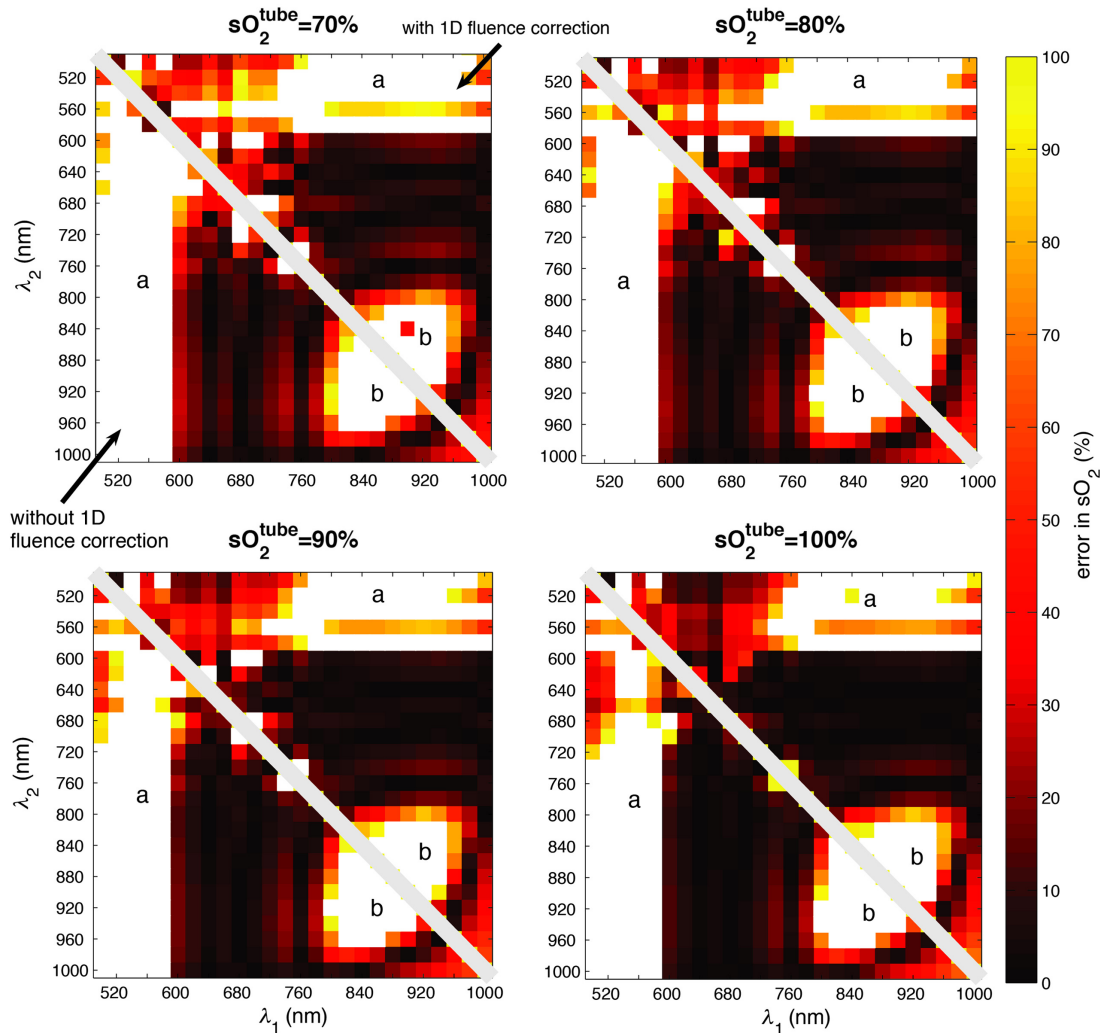
Fig. 6 Matrix of average error in oxygenation for 26 wavelength pair combinations in 500- to 1000-nm range for different values of oxygenation in the tube, sO_2^{tube}. In each figure, the bottom-left triangle represents sO_2 estimates made with a spectrally varying background and the top-right triangle represent sO_2 estimates made with a spectrally varying background using a depth-dependent fluence correction. The white pixels represent estimates with an error >100%. In the regions marked "a," this error is due to spectral coloring; in the regions marked "b," it is due to poor conditioning. Top left: $sO_2^{tube} = 70\%$. Top right: $sO_2^{tube} = 80\%$. Bottom left: $sO_2^{tube} = 90\%$. Bottom right: $sO_2^{tube} = 100\%$.

this wavelength range. Second, if the wavelengths that give the most accurate estimate of sO_2 are selected, a depth-dependent fluence correction may not provide a noticeable improvement over no correction in the accuracy of the sO_2 estimates. This is consistent with Figs. 5(b) and 5(c) and is likely to be due to the fact that the spectral coloring is less significant when wavelengths are well-chosen. The exception to this is where one wavelength lies in the 500- to 580-nm range. In this case, the application of the depth-dependent fluence correction provides a benefit as spectral coloring is more severe in this wavelength range. This may be important in applications that necessitate the use of shorter wavelengths, for example, in situations where high SNR is required or the excitation wavelength range of the laser source is limited.

Figure 6 also shows that in each case inversions, whether with or without the fluence correction, yield the highest accuracy at wavelengths between 620 and 1000 nm (excluding region "b"), with few wavelength pairs outside this range producing accurate sO_2 values. Moreover, the number of

wavelength pairs within this wavelength range yielding an error between 0% and 100% (the range of the color bar in Fig. 6) does not vary significantly for the four different tube sO_2 values. However, the number of wavelength pairs yielding reasonable accuracy (an error <10%) is strongly dependent on the value of sO_2^{tube} and decreases with decreasing sO_2^{tube}. This can be seen most clearly by comparing the maps for $sO_2^{tube} = 70\%$ and $sO_2^{tube} = 100\%$. These show that, while almost any pair of wavelengths between 620 and 1000 nm is likely to yield reasonable accuracy for $sO_2^{tube} = 100\%$, the same does not apply at $sO_2^{tube} = 70\%$, where there are significantly fewer wavelengths yielding accurate estimates. This is because the wavelength dependence of blood absorption increases with decreasing sO_2 for wavelengths below the isosbestic point at 800 nm.

At low sO_2^{tube} values, spectral coloring, therefore, increases for the majority of wavelengths in the 620- to 1000-nm range, so there are fewer wavelengths that provide the low levels of spectral coloring required for accurate estimates. Moreover, at low sO_2 values, those wavelength pairs that do yield accurate

estimates become increasingly sparsely distributed in wavelength space. Under the conditions of low sO_2^{tube}, the combination of these factors makes it very challenging to select the optimal wavelengths since they depend critically on the tube sO_2, which is generally unknown *a priori* in practice. (Interestingly, additional simulations, not shown, showed that, while the optimal wavelengths were dependent on the tube sO_2, they remained relatively independent of changes in the background sO_2. This may be due to the relatively low background absorption in this simulation.)

In summary, the results in this section suggest that, unlike the 500- to 600-nm wavelength range, where spectral coloring severely compromises accuracy under almost all conditions, it may be possible to achieve reasonable accuracy in some circumstances by using longer wavelengths between 620 and 1000 nm. For example, for high sO_2^{tube} values (>90%), a modest level of accuracy may be achieved by selecting any pair of wavelengths in this range. This relaxed wavelength selection criterion also has the important practical advantage that accurate estimates are not then contingent upon accurate *a priori* knowledge of sO_2^{tube}, just that it is around 90% or higher. This could be useful, e.g., when measuring sO_2 in an artery, where the physiological range of sO_2 is quite narrow. However, for lower sO_2^{tube} (e.g., 70%), wavelength selection becomes more problematic as there are only relatively few sparsely distributed wavelength pairs that provide reasonable accuracy and identifying these wavelengths requires *a priori* knowledge of sO_2^{tube}, which is generally unavailable *in vivo*. Even our seemingly encouraging conclusion that blind optimal wavelength selection within the 600- to 1000-nm range is feasible when sO_2^{tube} is high (>90%) may be unduly optimistic since it is based on the assumption of noise-free data. In practice, any imaging system will introduce noise, which may make certain wavelengths, where the absorption coefficient is low, impractical to use as the SNR would be too low, as discussed in Sec. 4.4.

4.4 Signal-to-Noise Ratio

The suggestion above to select wavelengths above 620 nm, where blood absorption changes less with wavelength than in the 500- to 600-nm range, means SNR is sacrificed in favor of the accuracy of the oxygenation estimates. This trade-off is made evident by considering that we require that the change in the initial pressure, p_0, as a result of the change in wavelength from λ_1 to λ_2, exceeds the minimum detectable change in p_0, the "image noise equivalent pressure": $|\delta p_0| > \eta$, or

$$|\delta\mu_a| > \frac{\sqrt{2}\eta}{\Gamma\Phi}, \tag{12}$$

assuming any system-dependent scaling has been calibrated out, and the wavelengths have already been chosen so that $\delta p_0 \approx \Gamma\Phi\delta\mu_a$. (The $\sqrt{2}$ factor comes from assuming that $p_0(\lambda_1)$ and $p_0(\lambda_2)$ are uncorrelated and have equal SNR such that their variance is additive.[46])

The inequality in Eq. (12) was applied to the data in Fig. 6 in order to illustrate the effect that taking SNR into account can have on the wavelength selection. This was performed for 325 wavelength pairs in the 500- to 1000-nm range and typical values for the parameters in Eq. (12) were used: $\eta = 0.2$ kPa and $\Gamma = 0.15$.[47] The fluence was computed using the MC-generated data used in the inversion and was scaled by an incident fluence of 20 mJ cm^{-2}. Using these parameters yielded

mean image SNRs in the range 26.5 to 45.7 dB, depending on wavelength and the value of the tube sO_2. These SNRs are significantly higher than those of high quality *in vivo* photoacoustic images of superficial blood vessels[48,49] and therefore on the optimistic side of what can be achieved in practice. The data in Fig. 6 were then replotted in Fig. 7 but masking the pairs of wavelengths for which the above inequality in Eq. (12) is not satisfied. In other words, those wavelengths for which the change in initial pressure from λ_1 to λ_2 was less than η.

For all the oxygenation levels studied, the number of wavelengths that yield accurate estimates in the presence of noise is considerably reduced compared to the noise-free case (shown in Fig. 6). For the remaining wavelength pairs, the depth-dependent fluence correction can provide an improvement in accuracy, particularly for low values of sO_2^{tube}. This is because for low values of sO_2^{tube}, most of the wavelength pairs that satisfy Eq. (12) comprise at least one wavelength that lies within the spectral range of 500 to 620 nm, where blood absorption is high and varies rapidly with wavelength. The resulting spectral coloring can, therefore, be mitigated by the fluence correction.

In terms of optimal wavelength selection, there are several spectral regions that yield reasonable accuracy. For $sO_2^{tube} = 100\%$ and $sO_2^{tube} = 90\%$, the spectral ranges defined by regions "a" and "b" in Fig. 7 provide accurate estimates with errors predominantly below ~10%. However, the number of wavelength pairs in these regions decreases with decreasing sO_2^{tube} to the extent that these regions disappear entirely for $sO_2^{tube} = 70\%$. This is because $\delta\mu_a$ becomes progressively smaller with decreasing sO_2 for wavelengths that lie in these spectral ranges until it no longer satisfies Eq. (12). Perhaps, a more useful spectral range is that defined by region "c," where one wavelength is 600 nm and the other must lie between 800 and 1000 nm. Any wavelength pair lying in region "c" will not only satisfy Eq. (12) but will also yield reasonable accuracy for all the sO_2^{tube} values represented in Fig. 7. This is because blood absorption at 600 nm is high so $\delta\mu_a$ arising from the wavelength change $\lambda_1 = 600$ nm to $\lambda_2 > 800$ nm is large enough to satisfy Eq. (12) but not so high as to introduce excessive spectral coloring. However, although accuracy appears to be reasonable for these wavelength pairs, it does decrease significantly with reducing sO_2^{tube}. For example, for the spectral range defined by region "c" and for $sO_2^{tube} = 100\%$, if $\lambda_2 = 600$ nm, then the error varies from 2.8% to 8.1% depending on the choice of λ_1. By contrast for $sO_2^{tube} = 70\%$, the corresponding error varies from 7.7% to 18% for region "c." The wavelength pair that yielded the best accuracy across all four sO_2 values in Fig. 7 was found to be 600/800 nm with an error ranging from 9.8% for $sO_2^{tube} = 70\%$ to 3.6% for $sO_2^{tube} = 100\%$.

Although this appears to suggest that reasonable accuracy may be possible using the 600-/800-nm wavelength pair, this should be treated with caution because the spectral region for one of the wavelengths is narrow, and a higher noise level or a small mismatch between the model and an experimental scenario could easily shift or eliminate it. In this study, optimistic assumptions of perfect image reconstruction and a homogeneous background have been made, so there are no artifacts, and the typical SNR achieved *in vivo* is often lower than used here. The latter will have the effect of reducing the number of wavelengths that satisfy Eq. (12) and for those remaining wavelengths that do satisfy it, they will be biased to shorter wavelengths, where spectral coloring is greater and accuracy reduced. Furthermore, the specific wavelength pairs that do
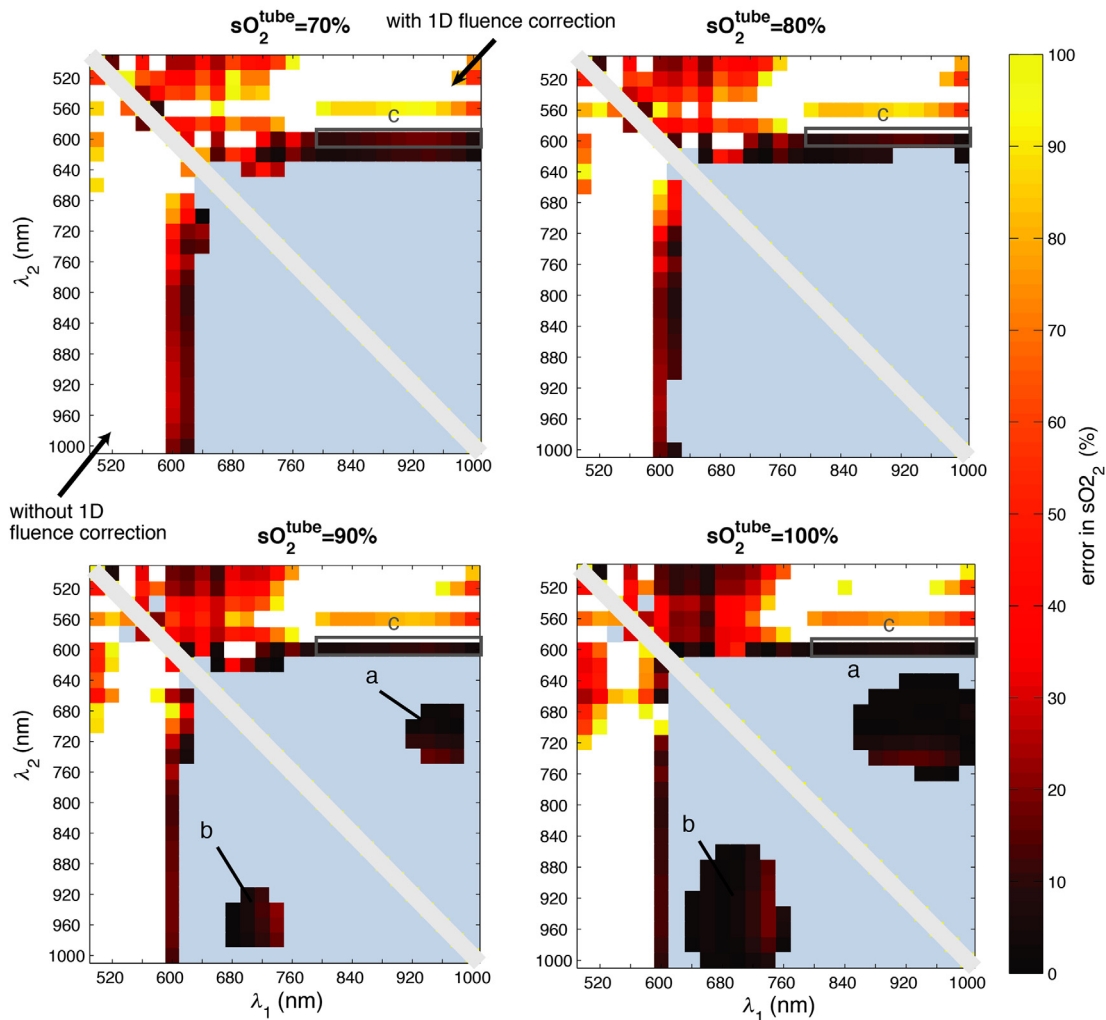
Fig. 7 Matrix of average error in oxygenation for 26 wavelength pair combinations in 500- to 1000-nm range for different values of oxygenation in the tube, sO_2^{tube}. In each figure, the bottom-left triangle represents sO_2 estimates made with a spectrally varying background and the top-right triangle represents sO_2 estimates made with a spectrally varying background using a depth-dependent fluence correction. The white pixels represent estimates with an error >100%. In the regions marked "a," this error is due to spectral coloring; in the regions marked "b," it is due to poor conditioning. Light blue regions indicate wavelengths that do not satisfy the condition in Eq. (12). (A) $sO_2^{tube} = 70\%$, (B) $sO_2^{tube} = 80\%$, (C) $sO_2^{tube} = 90\%$, and (D) $sO_2^{tube} = 100\%$.

yield accurate estimates will be increasingly dependent on the sO_2^{tube} values, which are unknown, making it hard to identify those wavelength pairs with confidence.

In summary, in the previous section it was observed that it may be possible to achieve accurate estimates for sO_2 values by selecting wavelength pairs in the 600- to 1000-nm range. However, noise has the negative impact of not only significantly reducing the number of wavelength pairs that provide accurate estimates but also making the specific values of these wavelengths increasingly dependent on the sO_2^{tube}. Indeed, it is the latter that arguably presents the greatest challenge since sO_2^{tube} is typically unknown *a priori*, determining it being the point of the measurement, so there is no certain way of identifying those few wavelength pairs that yield accurate estimates.

5 Summary and Conclusions

In this paper, we investigated whether blood oxygenation can be accurately estimated from multiwavelength photoacoustic images when using a simple linear inversion, which makes the simplifying assumption that the fluence does not change with wavelength, and spectral coloring is therefore ignored. The resulting linear inversions were tested using a digital phantom of a tube in a homogeneous background under the assumption that the acoustic properties and Grüneisen parameter within the domain were constant and that the acoustic inversion used to reconstruct p_0 provided an exact image without artifacts, distortion, or other errors.

These simulations demonstrated that spectral coloring distorts the internal fluence distribution and can introduce significant inaccuracies in the estimation of blood oxygenation. It was found that using a set of illumination wavelengths covering the NIR/visible spectrum (500 to 1000 nm) yields highly inaccurate blood oxygenation estimates. This is due to strong spectral coloring arising from the large changes in fluence that occur between 500 and 600 nm due to the strong wavelength dependence of hemoglobin absorption. Improved accuracy in sO_2

estimates was obtained when the absorption coefficient of blood changed little between illumination wavelengths resulting in reduced spectral coloring, which is generally achieved at wavelengths in the range 620 to 1000 nm. However, at these wavelengths, blood absorption is relatively low, so, in practice, this may result in the change in photoacoustic amplitude being below the noise floor, again resulting in inaccurate sO_2 estimation. There is, therefore, a "spectral coloring dilemma" in photoacoustic imaging when a fluence correction is not applied, a fundamental trade-off between accuracy and SNR, which limits the practical applicability of a linear spectroscopic inversion even for the simple tube geometry studied here.

Having said that, the tube study (see the results in Fig. 7) also showed that there may be pairs of wavelengths for which both the spectral coloring is low and there is sufficient SNR to obtain a fairly accurate sO_2 estimate. Unsurprisingly, the accuracy was improved when a fluence correction, although a partial, only depth-dependent, correction was used. (Of course, a complete and accurate fluence correction, were it available, could be used to remove the problem of spectral coloring altogether.) Nevertheless, the difficulty in making this observation practically useful is that the wavelengths that satisfy these requirements are by no means obvious. Here, they were found by testing hundreds of wavelength pairs using a model that, by definition in this phantom study, was accurate. In an *in vivo* context, an accurate description of the tissue anatomy and physiological status is rarely available *a priori*, and a simplified model will need to be used to compute the suitable wavelengths, bringing with it inevitable uncertainty.

It should also be noted that this study made a number of optimistic assumptions: First, and perhaps most notably, the images, $p_0(\lambda)$, contained no reconstruction artifacts. In practice, a photoacoustic image is not an exact depiction of the initial pressure distribution in the tissue due to instrumentation related factors, such as detector bandwidth limitations and limited detection aperture: the latter being a particularly significant source of error with the linear or planar detection arrays commonly used in clinical and preclinical photoacoustic imaging. Second, the spectral coloring was modest due to the very simple absorber (one tube), its superficial location (1 mm depth), and the limited sO_2 range investigated. Third, the SNR was higher than typically found in *in vivo* images. This was, therefore, a study of a best-case scenario. Despite this, there were relatively few cases for which reasonable sO_2 estimates were obtained. This suggests caution should be exercised when estimating blood oxygen saturation with a linear inversion *in vivo*, where the situation is likely to be considerably more complex and noisy.

In summary, the aim of this work was to investigate whether there are circumstances in which a linear spectroscopic inversion, which implicitly does not account for spectral coloring, can provide accurate *in vivo* sO_2 estimates. On a narrow reading, the conclusion is that there are, or could be, by choosing the right pair of wavelengths. The challenge will be finding these wavelengths with confidence. This will likely require carefully controlled experimental phantom studies or numerical simulations that fully account for the imaging instrumentation and image reconstruction process and the geometry and physiological status of anatomical region of interest. Given the multitude of confounding factors, it is our conclusion that a simple linear inversion should be applied *in vivo* with very considerable caution and only following a careful and comprehensive validation.

By highlighting the challenges presented when trying to estimate blood oxygenation using photoacoustic imaging, the findings of this study will help inform the practical application of spectroscopic photoacoustic techniques to the *in vivo* measurement of blood sO_2. This is becoming ever more important as photoacoustic imaging is translated to practical biomedical applications in the clinical and life sciences and commercial systems are increasingly available.

6 Appendix A: Modeling Light Transport in Tissue: Monte Carlo Simulations

There are a number of ways in which light distributions in scattering media can be modeled. The radiative transfer equation (RTE) is accurate, but the requirement for spatial and angular discretization quickly leads to excessive computational demands, especially for 3D domains. The diffusion approximation to the RTE, while computationally efficient, does not accurately model the fluence close to the surface, a region of interest for photoacoustics, or in nonturbid media. Here, a MC model of light transport was used. MC models estimate solutions to the RTE by summing a large number of "photon" random walks. They are highly parallelizable as the photons do not interact, and therefore, the computational effort scales well with domain size. The code used here was MCX,[50] which has been validated extensively and is GPU-accelerated, offering relatively short computation times even for a large number of photons required for convergence. A method to determine the number of photons required for convergence is presented below.

6.1 *Level of Uncertainty in* sO_2 *Estimates Due to MC Noise*

Although MC models are considered the "gold standard" (Ref. 51), it is important to recognize that due to the stochastic nature of simulations, the estimate of the fluence in each voxel will exhibit some variance. As the number of photons increases, the variance will reduce, and the fluence estimate will converge to the solution given by the RTE. For these simulations, it is necessary to use sufficient photons that the variance (the "MC noise") falls below an acceptable level. To assess what level this needs to be, we examined the effect of the MC noise on two-wavelength sO_2 estimates calculated from initial pressure distributions generated using the MC models. The standard deviation in oxygenation at a particular voxel, σ_{sO_2}, can be estimated from the standard deviation in the initial acoustic pressure simulated for that voxel, $\sigma_{p_0(\lambda_i)}$, by assuming that the initial pressure at the two wavelengths, $p_0(\lambda_1)$ and $p_0(\lambda_2)$, is uncorrelated[52] and using the standard formula:

$$\sigma_{sO_2} = sO_2 \sqrt{\left(\frac{\partial sO_2}{\partial p_0(\lambda_1)}\right)^2 (\sigma_{p_0(\lambda_1)})^2 + \left(\frac{\partial sO_2}{\partial p_0(\lambda_2)}\right)^2 (\sigma_{p_0(\lambda_2)})^2},$$

$$(13)$$

where the partial derivatives of sO_2 with respect to initial pressure can be computed from Eqs. (4) and (5). The standard deviation $\sigma_{p_0(\lambda_i)}$ was estimated from multiple simulations in the same domain with the same number of photons, by reseeding the random number generator to ensure different photon trajectories from simulation to simulation. Running the simulation multiple times at a pair of wavelengths allowed σ_{sO_2} to be calculated. The criterion used to ensure a sufficient number of photons was that

the spatially varying standard deviation σ_{sO_2} was everywhere much less than the precision required in the oxygenation estimates. Performing the analysis above for the tube phantom using 10 model runs at 500 and 800 nm (chosen due to the significant difference in the absorption of blood at these wavelengths) yielded relative uncertainties of <0.16% in the initial pressure in the tube. Using Eq. (13) to propagate these errors to sO_2 estimates made using $p_0(\lambda)$ averaged over the 10 model runs produced a maximum σ_{sO_2} value in the tube of 0.34%, indicating that uncertainties in sO_2 estimates due to MC noise are negligible. Note that this was corroborated by directly computing the standard deviation in the sO_2 estimates themselves.

Disclosures

The authors have no relevant financial interests in this article and no potential conflicts of interest to disclose.

Acknowledgments

References

1. K. Kwong and J. Belliveau, "Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation," *Proc. Natl. Acad. Sci. U. S. A.* **89**, 5675–5679 (1992).
2. A. Kleinschmidt et al., "Simultaneous recording of cerebral blood oxygenation changes during human brain activation by magnetic resonance imaging and near-infrared spectroscopy," *J. Cereb. Blood Flow Metab.* **16**, 817–826 (1996).
3. L.-D. Liao et al., "Imaging brain hemodynamic changes during rat forepaw electrical stimulation using functional photoacoustic microscopy," *Neuroimage* **52**, 562–570 (2010).
4. M. Gerling et al., "Real-time assessment of tissue hypoxia in vivo with combined photoacoustics and high-frequency ultrasound," *Theranostics* **4**, 604–613 (2014).
5. P. Vaupel et al., "Oxygenation of human tumors: evaluation of tissue oxygen distribution in breast cancers by computerized O2 tension measurements," *Cancer Res.* **51**(12), 3316–3322 (1991).
6. J. W. Severinghaus and Y. Honda, "History of blood gas analysis," *J. Clin. Monit.* **3**(2), 135–138 (1987).
7. A. Villringer et al., "Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults," *Neurosci. Lett.* **154**, 101–104 (1993).
8. A. P. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," *Phys. Med. Biol.* **50**, R1–R43 (2005).
9. H. Friedman, W. Koroshetz, and N. Qureshi, "Tissue plasminogen activator for acute ischemic stroke," *N. Engl. J. Med.* **333**(24), 1581–1587 (1996).
10. A. P. Gibson et al., "Three-dimensional whole-head optical tomography of passive motor evoked responses in the neonate," *Neuroimage* **30**, 521–528 (2006).
11. P. Beard, "Biomedical photoacoustic imaging," *Interface Focus* **1**, 602–631 (2011).
12. L. Li, L. Lin, and L. V. Wang, "Multiscale photoacoustic tomography," *Opt. Photonics News* **29**(4), 32–39 (2018).
13. W. Choi, E. P. Seungwan, and J. Chulhong, "Clinical photoacoustic imaging platforms," *Biomed. Eng. Lett.* **8**(2), 139–155 (2018).
14. E. Z. Zhang et al., "*In vivo* high-resolution 3D photoacoustic imaging of superficial vascular anatomy," *Phys. Med. Biol.* **54**, 1035–1046 (2009).
15. S. Wray et al., "Characterisation of the near-infrared absorption spectra of cytochrome-AA3 and haemoglobin for the non-invasive monitoring of cerebral oxygenation," *Biochim. Biophys. Acta* **933**, 184–192 (1988).
16. W. C. Vogt et al., "Photoacoustic oximetry imaging performance evaluation using dynamic blood flow phantoms with tunable oxygen saturation," *Biomed. Opt. Express* **10**, 449–464 (2019).
17. J. Gröhl et al., "Estimation of blood oxygenation with learned spectral decoloring for quantitative photoacoustic imaging (LSD-qPAI)" arXiv:1902.05839 (2019).
18. M. A. Naser et al., "Improved photoacoustic-based oxygen saturation estimation with SNR-regularized local fluence correction," *IEEE Trans. Med. Imaging* **38**, 561–571 (2019).
19. B. Cox et al., "Quantitative spectroscopic photoacoustic imaging: a review," *J. Biomed. Opt.* **17**, 061202 (2012).
20. J. Laufer et al., "Quantitative spatially resolved measurement of tissue chromophore concentrations using photoacoustic spectroscopy: application to the measurement of blood oxygenation and haemoglobin concentration," *Phys. Med. Biol.* **52**, 141–168 (2008).
21. S. Tzoumas et al., "Eigen spectra optoacoustic tomography achieves quantitative blood oxygenation imaging deep in tissues," *Nat. Commun.* **7**, 12121 (2016).
22. C. Cai et al., "End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging," *Opt. Lett.* **43**, 2752–2755 (2018).
23. A. Hussain et al., "Quantitative blood oxygen saturation imaging using combined photoacoustics and acousto-optics," *Opt. Lett.* **41**, 1720–1723 (2016).
24. S. Lee et al., "Non-invasive monitoring of the therapeutic response in sorafenib-treated hepatocellular carcinoma based on photoacoustic imaging," *Eur. Radiol.* **28**, 372–381 (2018).
25. J. Reber et al., "Non-invasive measurement of brown fat metabolism based on optoacoustic imaging of hemoglobin gradients," *Cell Metab.* **27**, 689–701.e4 (2018).
26. I. Quiros-Gonzalez et al., "Optoacoustics delineates murine breast cancer models displaying angiogenesis and vascular mimicry," *Br. J. Cancer* **118**, 1098–1106 (2018).
27. K. Okumura et al., "Evaluation of renal oxygen saturation using photoacoustic imaging for the early prediction of chronic renal function in a model of ischemia-induced acute kidney injury," *PLoS One* **13**, e0206461 (2018).
28. J. Lv et al., "Hemispherical photoacoustic imaging of myocardial infarction: in vivo detection and monitoring," *Eur. Radiol.* **28**, 2176–2183 (2018).
29. D. J. Lawrence et al., "Spectral photoacoustic imaging to estimate in vivo placental oxygenation during preeclampsia," *Sci. Rep.* **9**, 558 (2019).
30. M. Martinho Costa et al., "Quantitative photoacoustic imaging study of tumours in vivo: baseline variations in quantitative measurements," *Photoacoustics* **13**, 53–65 (2019).
31. C. Wood et al., "Photoacoustic-based oxygen saturation assessment of murine femoral bone marrow in a preclinical model of leukemia," *Photoacoustics* **14**, 31–36 (2019).
32. A. Needles et al., "Development and initial application of a fully integrated photoacoustic micro-ultrasound system," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **60**(5), 888–897 (2013).
33. V. Ermolayev et al., "Simultaneous visualization of tumour oxygenation, neovascularization and contrast agent perfusion by real-time three-dimensional optoacoustic tomography," *Eur. Radiol.* **26**, 1843–1851 (2016).
34. C. Arthuis et al., "Real-time monitoring of placental oxygenation during maternal hypoxia and hyperoxygenation using photoacoustic imaging," *PLoS One* **12**, e0169850 (2017).
35. M. Tomaszewski et al., "Oxygen enhanced optoacoustic tomography (OE-OT) reveals vascular dynamics in murine models of prostate cancer," *Theranostics* **7**, 2900–2913 (2017).
36. M. Tomaszewski et al., "Oxygen-enhanced and dynamic contrast-enhanced optoacoustic tomography provide surrogate biomarkers of tumor vascular function, hypoxia, and necrosis," *Cancer Res.* **78**, 5980–5991 (2018).
37. J. Lavaud et al., "Photoacoustic imaging as an innovative technique for the exploration of blue rubber bleb naevus," *Br. J. Dermatol.* **181**(3), 596–597 (2019).

38. C. Lutzweiler et al., "Real-time optoacoustic tomography of indocyanine green perfusion and oxygenation parameters in human finger vasculature," *Opt. Lett.* **39**, 4061–4064 (2014).
39. G. Diot, A. Dima, and V. Ntziachristos, "Multispectral opto-acoustic tomography of exercised muscle oxygenation," *Opt. Lett.* **40**, 1496–1499 (2015).
40. E. Fakhrejahani et al., "Clinical report on the first prototype of a photoacoustic tomography system with dual illumination for breast cancer imaging," *PLoS One* **10**, e0139113 (2015).
41. Y. Zhu et al., "Light emitting diodes based photoacoustic imaging and potential clinical applications," *Sci. Rep.* **8**, 9885 (2018).
42. H. Yoon, G. P. Luke, and S. Y. Emelianov, "Impact of depth-dependent optical attenuation on wavelength selection for spectroscopic photoacoustic imaging," *Photoacoustics* **12**, 46–54 (2018).
43. S. L. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**, 5007–5008 (2013).
44. R. Hochuli, P. C. Beard, and B. Cox, "Effect of wavelength selection on the accuracy of blood oxygen saturation estimates obtained from photoacoustic images," *Proc. SPIE* **9323**, 93231V (2015).
45. R. J. Paproski et al., "Multiwavelength photoacoustic imaging of inducible tyrosinase reporter gene expression in xenograft tumors," *Sci. Rep.* **4**, 5329 (2014).
46. E. S. Lee and R. E. Forthofer, "Strategies for variance estimation," Chapter 4 in *Analyzing Complex Survey Data*, 2nd ed., pp. 22–39, SAGE Publications, Thousand Oaks, California (2006).
47. J. Yao and L. V. Wang, "Photoacoustic microscopy," *Laser Photonics Rev.* **7**, 758–778 (2013).
48. J. Laufer et al., "*In vivo* preclinical photoacoustic imaging of tumor vasculature development and therapy," *J. Biomed. Opt.* **17**, 056016 (2012).
49. A. P. Jathoul et al., "Deep *in vivo* photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter," *Nat. Photonics* **9**, 239–246 (2015).
50. Q. Fang and D. A. Boas, "Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units," *Opt. Express* **17**, 20178–20190 (2009).
51. Q. Fang, "Mesh-based Monte Carlo method using fast ray-tracing in Plücker coordinates," *Biomed. Opt. Express* **1**, 165–75 (2010).
52. P. R. Bevington and D. K. Robinson, "Instrumental and statistical uncertainties," Chapter 3.1 in *Data Reduction and Error Analysis for the Physical Sciences*, 3rd ed., pp. 36–39, McGraw-Hill, New York (1969).

Biographies of the authors are not available.