

Effects of skin tone on photoacoustic imaging and oximetry

Thomas R. Else^{©, a, b}, Lina Hacker^{©, a, b, †}, Janek Gröhl^{©, a, b}, Ellie V. Bunce^{©, a, b}, Ran Tao^{©, a, b}
and Sarah E. Bohndiek^{©, a, b, *}

^aUniversity of Cambridge, CRUK Cambridge Institute, Cambridge, United Kingdom

^bUniversity of Cambridge, Department of Physics, Cambridge, United Kingdom

ABSTRACT. **Significance:** Photoacoustic imaging (PAI) provides contrast based on the concentration of optical absorbers in tissue, enabling the assessment of functional physiological parameters such as blood oxygen saturation (sO_2). Recent evidence suggests that variation in melanin levels in the epidermis leads to measurement biases in optical technologies, which could potentially limit the application of these biomarkers in diverse populations.

Aim: To examine the effects of skin melanin pigmentation on PAI and oximetry.

Approach: We evaluated the effects of skin tone in PAI using a computational skin model, two-layer melanin-containing tissue-mimicking phantoms, and mice of a consistent genetic background with varying pigmentations. The computational skin model was validated by simulating the diffuse reflectance spectrum using the adding-doubling method, allowing us to assign our simulation parameters to approximate Fitzpatrick skin types. Monte Carlo simulations and acoustic simulations were run to obtain idealized photoacoustic images of our skin model. Photoacoustic images of the phantoms and mice were acquired using a commercial instrument. Reconstructed images were processed with linear spectral unmixing to estimate blood oxygenation. Linear unmixing results were compared with a learned unmixing approach based on gradient-boosted regression.

Results: Our computational skin model was consistent with representative literature for *in vivo* skin reflectance measurements. We observed consistent spectral coloring effects across all model systems, with an overestimation of sO_2 and more image artifacts observed with increasing melanin concentration. The learned unmixing approach reduced the measurement bias, but predictions made at lower blood sO_2 still suffered from a skin tone-dependent effect.

Conclusion: PAI demonstrates measurement bias, including an overestimation of blood sO_2 , in higher Fitzpatrick skin types. Future research should aim to characterize this effect in humans to ensure equitable application of the technology.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International 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.29.S1.S11506](https://doi.org/10.1117/1.JBO.29.S1.S11506)]

Keywords: photoacoustic imaging; spectral unmixing; racial bias; Fitzpatrick skin types; melanin

Paper 230232SSR received Aug. 17, 2023; revised Nov. 2, 2023; accepted Nov. 28, 2023; published Dec. 20, 2023.

*Address all correspondence to Sarah E. Bohndiek, seb53@cam.ac.uk

†Present address: University of Oxford, Department of Oncology

1 Introduction

Photoacoustic imaging (PAI) is an emerging molecular imaging modality that provides contrast based on the absorption of light by different molecules (chromophores) in tissue. These molecules may be endogenous, such as hemoglobin, melanin, lipids and water, or exogenous, covering a range of contrast agents.^{1,2} PAI has the potential to recover the concentration of each of these chromophores based on their unique optical absorption spectra, by acquiring data at multiple wavelengths and subjecting the reconstructed images to multispectral processing methods.³ For example, the recovery of relative concentrations of oxy- and deoxy-hemoglobin is commonly used to derive oxygenation maps from PAI with linear spectral unmixing.⁴ Obtaining accurate maps of chromophore concentrations, however, is limited by spatial variations in light fluence, which are significant at depth and are highly dependent on the surrounding distribution of optical absorbers and scatterers.³ This is often referred to as ‘spectral coloring’. Several factors can affect tissue optical properties, including age,⁵ low blood perfusion,⁶ or skin tone,⁷ which consequently affect the ability to extract quantitative information from PAI.

Measurement biases introduced by the differential absorption of melanin in darker skin tones have been identified in several technologies where light passes through the skin before making a quantitative measurement, including pulse oximeters,^{8–18} bilirubinometers,^{19,20} wearable technologies,^{21–24} cerebral oximeters,²⁵ and optical reflectance measurements.²⁶ These biases can have significant implications for patient management, an issue that came to the fore in the management of COVID-19 patients, where the over-estimation of blood oxygenation by pulse oximetry in black patients may have led to under-diagnosis of hypoxemia compared to white patients.¹⁷ PAI also relies on the transmission of light through the skin and is thus also expected to be subject to similar limitations.

Skin tone is typically defined according to the Fitzpatrick scale, which provides a ranking based on the response of the skin to ultraviolet light determined by a questionnaire, from type I (always burns, never tans) to type VI (never burns, pigmented skin). A few studies have begun to emerge that examine skin tone measurement bias in PAI.^{27–30} Higher Fitzpatrick skin type grading was shown to reduce PAI signal-to-noise ratio deep in the tissue and introduced biases in the estimated blood oxygenation.^{27–30} An empirical calibration scheme has been proposed to correct these biases.²⁹ These early studies highlight the key challenges associated with varying skin tones in PAI, how we measure skin tone, and how we develop fluence compensation schemes. The first of these key challenges could be tackled using more quantitative measures of skin tone, for example, measurements based on colorimeters. The second may be improved by data-driven approaches, which improve on standard linear unmixing schemes for photoacoustic oximetry but have not been tested in cases of varying skin tone.^{4,31–35}

Here, we undertook a systematic evaluation of the effects of skin melanin concentration on PAI, using a combination of simulation, phantom, and animal models to examine the highlighted challenges. We illustrate that linear spectral unmixing is highly dependent on skin tone and introduces biases in photoacoustic oximetry. These biases mirror observations made in pulse oximetry.¹⁷ We demonstrate that learned spectral decoloring can partially compensate for this bias, recovering less biased oxygenation values. We underscore these findings in animal studies, showing a systematic overestimation of blood oxygenation with increasing skin pigmentation. Our results provide a theoretical basis upon which we can interpret future studies in patients with different skin tones.

2 Methods

2.1 Computational Models

To assess the effects of skin tone on PAI in a controlled setting, we developed a computational forearm model imitating the structure and physical properties of the skin. The model was constructed in Python using absorption and reduced scattering spectra obtained from the literature.³⁶ Three variations of the model were used: a layer-only model for calibration with adding-doubling reflectance simulations, a varying blood oxygenation model for optical modeling, and a realistic model for full photoacoustic simulation. These distinctions were made to allow thorough characterization of the models under computational constraints. To bridge the gap between simulation

and experiment, we also simulated optical absorption in the agarose phantom model (described in Sec. 2.5).

The first model consisted of three parallel layers. The first layer represented the epidermis, modelled as 60 μm thick, with optical absorption from a varying concentration of melanosomes. The melanosome volume fraction was chosen by first simulating diffuse reflectance (Sec. 2.2) over a range previously identified in the literature (1.3% to 43%³⁷) and adjusting the simulation range to align as closely as possible with individual typology angle (ITA) values associated with different Fitzpatrick types.³⁸ The resulting parameters were then qualitatively verified using a publicly available integrating sphere forearm skin reflectance dataset (Fig. S1 of the [Supplementary Material](#)).³⁹ We settled on six logarithmically-spaced values of melanosome concentration between 2% v/v and 40% v/v, representing the six Fitzpatrick skin types. The second layer, representing the dermis, was 1 mm thick, containing blood (0.2% v/v, 70% oxygenation), water (65% v/v) and lipid (20% v/v).³⁶ The third layer, representing generic tissue, filled the rest of the simulation volume. It consisted of blood (2.5% v/v, 70% oxygenation), water (65% v/v), and lipid (20% v/v). The absorption spectra for these three layers are shown in Fig. 1(a). Scattering properties were defined by a combination of Rayleigh and Mie scattering, assuming a constant scattering anisotropy of 0.9. These values are representative of typical values in each region.³⁶ This simplified layer model was used for validation purposes, using adding-doubling simulations of diffuse reflectance. The structure formed the basis of all simulations, with blood vessels added for Monte-Carlo simulations.

To model the effects of changing blood oxygenation in a blood vessel on the photoacoustic initial pressure, Monte-Carlo simulations were run with the varying blood oxygenation model, giving the optical absorption distribution. This model was based on the layer structure described above with the addition of a cylinder of radius 1.25 mm, at a depth of 2.5 mm below the epidermis, representing a blood vessel of varying oxygenation (between 0% and 100%).

To evaluate the effects of a more realistic tissue model, a third model, the realistic forearm model, was subjected to both optical and acoustic simulations, giving full photoacoustic data that was processed using normal reconstruction and analysis procedures. Two cylinders were added to the layer model, one representing an artery (100%-oxygenated blood, radius 1.25 mm at a depth of 2.5 mm below the epidermis) and one representing a vein (70% oxygenated blood, radius 0.625 mm at a depth of 2.5 mm).

We also ran optical simulations of the cylindrical phantoms (Sec. 2.5), with approximately identical optical absorption and scattering properties as the agarose phantoms. We defined the model in Python, with a grid size of 100 μm in an array of size (220 \times 220 \times 220). The tube in the center of the phantom was modeled as a cylindrical absorber of radius 0.75 mm containing blood. Concentric absorbers were added representing the inner part of the phantom (23 mm diameter, water absorption and intralipid scattering) and the outer layer of the phantom (1 mm thick, varying melanosome volume fraction, intralipid scatterer). The absorption coefficients of the melanin layer were defined by finding the melanosome volume fraction that gave the same optical absorption coefficient at 700 nm as measured in the double integrating sphere system (Table S1 in the [Supplementary Material](#), Sec. 2.6).

2.2 Adding-Doubling Simulations

The layer-only skin model was validated using the adding-doubling model of diffuse reflectance⁴⁰ to allow for fast evaluation of the reflectance spectra at several wavelengths (450 nm to 900 nm in 20 nm steps) in the visible and near-infrared ranges. We additionally calculated the reflectance at 685 nm to enable direct comparison with literature values.⁴¹ The reflectance spectra were converted into an RGB image and an ITA metric using a D65 standard light source and a 10-deg standard observer in the Python color-science library.⁴² ITA has been proposed as a quantitative measure of skin tone and can be obtained from low-cost colorimeters. ITA was calculated from skin reflectance spectra using the following equation:

$$\text{ITA} = \arctan\left(\frac{L^* - 50}{b^*}\right) \times \frac{180}{\pi}, \quad (1)$$

where L^* and b^* are taken from the CIELAB color space, representing luminance [ranging from black (0) to white (100)] and the range from yellow to blue, respectively.

2.3 Monte-Carlo Simulations

We simulated the PAI optical forward model using the varying blood oxygenation model described above using a custom version of MCX (<https://github.com/IMSY-DKFZ/mcx>), giving initial pressure distributions that correspond to an idealized photoacoustic image. The model was defined on a grid with a pixel size of $60\ \mu\text{m}$, which allowed the simulation to run in a sufficiently short time while ensuring that the simulations were sufficiently detailed. The light transport through the model was simulated with 10^7 photons per simulation, at 21 equally spaced wavelengths between 700 and 900 nm inclusive.⁴³

We used an illumination geometry, which approximates a clinical PAI system (MSOT Acuity, iThera Medical GmbH, Munich, Germany), consisting of a single optical fiber with a beam divergence of 8.66 deg, 43.2 mm above the surface of the probe, directed at an angle of 22.4 deg from the imaging plane, intersecting the imaging plane 2.8 mm outside of the probe.⁴⁴ Optical simulations of the agarose phantoms were run using the same customized version of MCX with an illumination geometry approximating a pre-clinical photoacoustic system (MSOT inVision 256-TF, iThera Medical GmbH, Munich Germany), defined on a grid of size $100\ \mu\text{m}$. The illumination geometry consisted of five radially distributed fiber bundle pairs.⁴⁴ The version of MCX used in these experiments is available on GitHub: <https://github.com/BohndiekLab/melanin-phantom-simulation-paper>. 10^7 photons were simulated per fiber bundle, for a total of 10^8 photons. All simulations were run on a computer with 2 NVIDIA Quadro 8000 GPUs (48 GB GPU RAM each), Intel Xeon Gold 6230 CPU at 2.10 GHz (40 cores), and 256 GB RAM.

2.4 Acoustic Simulations

To determine the impact of acoustic modeling, we ran the forward acoustic model in j-Wave⁴⁵ to calculate the pressure time series detected by the ultrasound transducer array for the initial pressure distributions. Simulations were run in three dimensions, with a grid size of $240\ \mu\text{m}$, and a speed of sound of $1500\ \text{m s}^{-1}$. Three-dimensional initial pressure distributions were taken from Monte-Carlo simulations and down-sampled by a factor of four, and their propagation was simulated. The resulting pressure distributions were evaluated using band-limited interpolation sensors in the position of the sensors in the photoacoustic system. The pressure time series were recorded and processed in the same way as the experimental data.

2.5 Phantom Design

Phantoms were used to validate the simulations in a real photoacoustic system. A custom-made 3D-printed polylactic acid mold was constructed to fabricate a two-layer cylindrical phantom with an inner scattering compartment and an outer compartment containing melanin to model the epidermis. The mold consisted of two cylindrical compartments, one of radius 9 mm and one of radius 10 mm. A polyvinyl chloride tube (inner diameter 1.5 mm, outer diameter 2.1 mm; VWR 228-3857) was placed in the center of the phantom mold to enable blood flow through the phantom.

Agarose-based phantoms were fabricated using standard methods.⁴⁶ The inner phantom mixture consisted of a mixture of agarose for mechanical robustness and intralipid for scattering. 1.5% (w/v) agarose (Sigma-Aldrich 05039) was added to Milli-Q water and heated until it dissolved. Once the mixture had cooled to $\sim 40^\circ\text{C}$, 2.08% (v/v) of pre-warmed intralipid (20% emulsion, Sigma-Aldrich I141) was added and mixed, before being poured into the inner compartment of the phantom mold and allowed to set, constituting the inner 9 mm radius of the phantom. Synthetic melanin (Sigma, M8631, $2.5\ \text{mg mL}^{-1}$ stock in dimethyl sulfoxide) was added to a separate batch of the base mixture at the desired concentrations (Table S1 in the [Supplementary Material](#)). Once the inner compartment was set, the outer layer mixture was added and again allowed to set at room temperature (Fig. S2 in the [Supplementary Material](#)).

2.6 Phantom Characterization

A custom double integrating sphere system was used to characterize the optical absorption and reduced scattering coefficients of the phantom material. Samples were placed between two integrating spheres (Avantes, AvaSphere-50, 50 mm internal diameter), both of which were

connected to a spectrometer (Avantes, Starline Avaspec-2048) via an optical fiber. The reflectance integrating sphere was connected to a broadband light source (Avantes, Avalight-HAL-s-mini) via an optical fiber. Circular samples (approximate diameter: 5 cm) of the agarose mixture with and without melanin were prepared with a thickness of ~ 3 mm. Sample thickness was measured using vernier callipers. Each agarose sample was placed between the two spheres and the transmittance and reflectance were measured. Optical scattering and absorption coefficients were calculated using the inverse adding doubling method,⁴⁷ assuming a scattering anisotropy factor of 0.89 and refractive index of 1.34.⁴⁸ These measurements were made on three replicates of each sample, over a wavelength range of 450 to 900 nm at 1 nm steps.

2.7 Blood-Flow Circuit

A blood-flow circuit was used to control the oxygenation of blood flowing through the tube centered in the two-layer cylindrical phantom, enabling photoacoustic measurements of blood with known ground-truth blood oxygenation as described previously.^{49,50} Briefly, we measured the oxygen partial pressure (pO_2) and the photoacoustic spectrum of the blood simultaneously. We used fresh whole blood from Wistar rats (6 to 9 months, Charles River Laboratories). All blood samples were anti-coagulated in ethylenediaminetetraacetic acid (EDTA), stored at 4°C, and processed within 48 h. The blood was pumped around the circuit using a peristaltic pump (Fisher Scientific CTP100), and the oxygenation was controlled by initially oxygenating the blood using hydrogen peroxide (30% (w/w) in deionized water, Sigma-Aldrich 7722-84-1) and then slowly injecting 0.03% w/v sodium dithionite (ACROS Organics 7775-14-6, in PBS) using a syringe pump ($10 \mu\text{L min}^{-1}$, Harvard, MKCB2159V). To measure the blood pO_2 , two oxygen fluorescence quenching needle probes (Oxford Optronix, NX-BF/O/E) were placed before and after the phantom; the average of these measurements was taken as the ground-truth pO_2 in the phantom. pO_2 was converted to blood oxygenation (sO_2) using the Severinghaus equation.⁵¹ The values were extracted using custom code on an Arduino UNO and fed to a computer over USB where they were read and stored using custom code in MATLAB. Experiments were conducted at 37°C.

2.8 Animal Procedures

All animal procedures were conducted under personal and project licences issued under the United Kingdom Animals (Scientific Procedures) Act, 1986 and reviewed by the Animal Welfare and Ethical Review Board at the CRUK Cambridge Institute. The protocols were approved locally under compliance forms CFSB2317 and CFSB2348.

To understand the effects of skin pigmentation in a model *in vivo* setting, we compared non-pigmented albino mice ($n = 10$, C57BL/6 albino, Janvier Labs, France, 10 to 13 weeks old) to pigmented “B6” mice ($n = 10$, C57BL/6J, Charles River, 10 to 13 weeks old). The B6 mice were further subdivided qualitatively based on the visual appearance of pigmentation patches in color photographs (Fig. S3 in the [Supplementary Material](#)) into a non-pigmented group ($n = 8$ scans) and a pigmented group ($n = 11$ scans). Albino mice were imaged once and B6 mice were imaged twice (2 weeks apart) to allow the development of pigmentation after hair removal. Repeat measurements of the same mouse were treated as independent measurements for statistical analysis given the time lapse between measurements.

2.9 Photoacoustic Imaging

Phantoms and mice were imaged using a pre-clinical multispectral photoacoustic tomography system (MSOT inVision-256TF, iThera Medical GmbH).⁴⁶ Briefly, excitation pulses were provided by a tunable optical parametric oscillator, pumped by a nanosecond Nd:YAG laser (10 Hz repetition rate up to 7 ns pulse duration). A custom optical fiber assembly creates a diffuse ring of uniform illumination over the imaging plane within the sample. The sample is coupled to the transducers using a water bath. For ultrasound detection, an array of 256 toroidally focused ultrasound transducers covering a 270 deg circular arc and radius of 4 cm was used with a center frequency of 5 MHz and 60% bandwidth.

For phantom imaging, a single cross-sectional slice of the phantom was imaged continuously during the blood deoxygenation procedure until the blood was completely deoxygenated. Phantoms were placed in a custom animal holder (iThera Medical GmbH, Munich, Germany),

wrapped in a thin polyethylene membrane, with water used to couple the phantom to the membrane. Imaging was performed at wavelengths from 700 to 900 nm inclusive with 10 nm steps and frame averaging was enabled (10 averages) to improve the signal-to-noise ratio.

Mice were prepared for imaging according to our standard operating procedure.⁴⁶ Briefly, the mice were shaved and depilatory cream was used to remove excess hair. The mice were anaesthetized using <3% isoflurane in a 50% pure oxygen/medical air mixture and placed in a custom animal holder (iThera Medical GmbH, Munich, Germany), wrapped in a thin polyethylene membrane, with ultrasound gel (Aquasonic Clear, Parker Labs) used to couple the skin to the membrane. The holder was then placed within the PAI system and immersed in deionized water maintained at 36°C for Hb and HbO₂ imaging acquisition. The animal respiratory rate was maintained in the range 70 to 80 b.p.m. with 1.5% isoflurane concentration for the entire scan. A single cross-sectional slice through the kidneys and spleen was imaged. Frame averaging was enabled (four averages) to minimize the effects of breathing motion artifacts. 10 repeat scans were taken at each of the following wavelengths 700, 730, 750, 760, 770, 800, 820, 840, 850, and 880 nm.

2.10 Image Reconstruction and Processing

Simulated, phantom, and *in vivo* photoacoustic data were processed using the Python photoacoustic tomography analysis toolkit (PATATO, <https://github.com/bohndieklab/patato>).

Acoustic simulation images were reconstructed by filtered backprojection with a grid spacing of 60 μm and a field of view of 18 mm \times 40 mm. Phantom and *in vivo* images were reconstructed by filtered backprojection with a grid spacing of 75 μm and a field of view of 25 mm. A band-pass filter was applied to all of the time series data, with a bandpass range of 5 to 7 MHz.

Linear spectral unmixing was applied to all simulated, phantom, and *in vivo* datasets using the matrix pseudo-inverse function in NumPy⁵² and literature values for the optical absorption spectra of oxy- and deoxy-hemoglobin.³⁶ Photoacoustic estimates of the blood oxygenation (denoted with a superscript EST $s\text{O}_2^{\text{EST}}$ to differentiate it from the true blood oxygenation, $s\text{O}_2$) were made pixel-wise from the unmixing components as the ratio of the oxyhemoglobin to the total hemoglobin. To evaluate the potential for a machine-learning approach to improve photoacoustic oximetry in the presence of melanin, a learned spectral decoloring³¹ algorithm was trained on a numerical tissue model containing varying levels of melanin in the epidermis layer (0.1% to 5% melanosome v/v). We simulated 500 unique volumes 19.2 mm³ in volume with a resolution of 0.3 mm that contained up to nine blood vessels with a random radius and blood oxygenation. The tissue background contained a blood concentration of 1% (v/v) and an oxygenation of 70%. These design parameters were developed independently of this study and simulations were obtained using the SIMPA framework.⁴⁴ From the simulated initial pressure distributions, we used a histogram-based gradient boosting regression tree as the inversion algorithm with a maximum depth of 16 and otherwise default hyperparameters as implemented in the scikit-learn (version 1.2.1) Python package.

Unless otherwise specified, manual polygon regions of interest (ROIs) were drawn and average spectra or unmixed components were calculated by masking the defined regions. In phantoms, the ROIs were taken around the blood tube and the melanin compartment. The melanin compartment was defined by drawing an ROI around the whole phantom and eroding the polygon by 1 mm. The difference mask between the drawn and eroded polygons was applied to the photoacoustic images, giving the multispectral data within the melanin region. Mean photoacoustic and unmixed component values were extracted. In mice, ROIs were drawn around the whole mouse body. The body ROI was eroded by 1 mm to subdivide it into a “skin” region (outer 1 mm) and a “body” region (inner region). Negative pixel values were included in the median spectra calculation to illustrate image reconstruction artifacts. For unmixed $s\text{O}_2^{\text{EST}}$ calculation, pixels were excluded if either of the unmixed oxy- or deoxy-hemoglobin coefficients were negative; the overall $s\text{O}_2^{\text{EST}}$ was calculated as the ratio of the median oxyhemoglobin component to the median total hemoglobin. The median was chosen because of the high negative pixel percentage in the body of the pigmented mice. Pixels near the open region of the detector were manually excluded from the analysis as the image reconstruction is less well posed in that region, leading to artifacts.

Line plots of the photoacoustic signal through the mice were generated by constructing a line through the kidneys and vena cava of the mice. To ensure constant placement of this line, it was constructed to be 6.5 mm from the surface of the mouse closest to the spline, and perpendicular to an axis through the spine and aorta. This construction was observed to pass through the kidneys and vena cava in albino and non-pigmented mice. The same construction was applied to the heavily pigmented mice, where the kidneys and spleen were not visible, to give a direct comparison. The values of the reconstructed image, total hemoglobin map, and blood oxygenation were calculated along this line using the `interpnd` function of SciPy.⁵²

2.11 Statistics

Statistical analyses were conducted in Python using the `statsmodels` library. The relationship between continuous variables was quantified by linear regression. A log transformation was applied where specified to ensure that the residual normality assumption is upheld in each linear model. The residual normality assumption was tested using the Shapiro-Wilk test ($p > 0.1$ for all models). For the *in vivo* data, the relationship between each response variable and the mouse pigmentation was assessed using a linear model, with both mouse strain (Albino or B6) and pigmentation (not pigmented or pigmented) as independent variables. The model related the response variable (sO_2 or photoacoustic signal) to a sum of the effect due to mouse strain and pigmentation with no interaction term.

3 Results

3.1 Computational Model of Skin Shows Good Agreement with Literature Based on Adding-Doubling Reflectance Measurements

We first validated our computational model of human skin, with the assigned optical absorption spectra in the epidermis [Fig. 1(a)], by calculating diffuse reflectance spectra using the adding-doubling method. We chose six values of melanosome concentration to represent the typical physiological range observed in the population. The calculated spectra show decreasing reflectance with increasing melanosome concentration and broadly increasing reflectance with wavelength, as one would expect given the profile of optical absorption of melanin [Figs. 1(a) and 1(b)]. We then calculated the ITA for each reflectance spectrum, allowing direct comparison

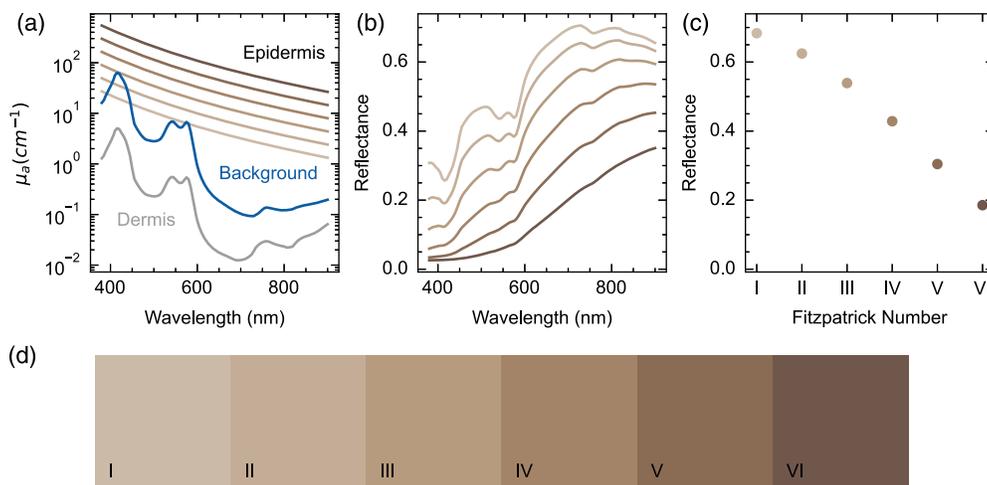


Fig. 1 Diffuse reflectance simulations validate Monte-Carlo simulation parameters and enable the assignment of simulations to equivalent Fitzpatrick skin types. (a) Optical absorption spectra in the tissue model were assigned according to the desired melanosome concentration (2% to 40% v/v, gradations of brown) in the skin layer, dermis below (gray) and background absorber (blue). (b) Diffuse reflectance spectra were calculated using the adding-doubling method. (c) Reflectance values simulated at 685 nm cover the physiological range of skin reflectance. (d) Red, green, and blue images were simulated from the reflectance spectra and the six models were assigned to the six Fitzpatrick skin types (shown by inset label I to VI).

Table 1 Parameters of the skin simulations (melanosome concentration, absorption coefficient) and quantities derived from adding-doubling simulations (reflectance, ITA, and HEX color value). The equivalent Fitzpatrick skin type was assigned based on literature values.⁵³

Melanosome concentration (v/v, %)	Absorption coefficient at 685 nm (cm ⁻¹)	Reflectance at 685 nm	ITA (deg)	Assigned Fitzpatrick skin type	HEX color
2.0	3.45	0.72	67	I	#CBB9A8
3.6	6.28	0.65	56	II	#C3AC95
6.6	11.43	0.57	41	III	#B69A7F
12.1	20.81	0.46	20	IV	#A28367
22.0	37.89	0.33	-6.9	V	#8A6B54
40.0	68.98	0.20	-47	VI	#6E574C

with *in vivo* measurements from skin colorimeters. These ITA values (67 deg to -47 deg) are consistent with previous observations made *in vivo* using the Fitzpatrick skin type (Table 1).⁵³ Reflectance at 685 nm was evaluated for each tissue model, giving a range of reflection coefficients from 0.20 to 0.72 over our simulated range of melanosome concentrations [Fig. 1(c), Table 1] again consistent with previous literature.⁴¹ To enable qualitative evaluation of our skin model, the reflection coefficients were converted into red, green, and blue (RGB) images [Fig. 1(d), Table 1]. Our results align closely with literature values made of skin reflectance using a single integrating sphere (Fig. S1 in the [Supplementary Material](#)).³⁹ For the remainder of the paper, results from different Fitzpatrick skin types are color coded according to the calculated RGB values for consistency [Fig. 1(d)].

3.2 Photoacoustic Simulations with the Skin Model Show Spectral Coloring Due to Skin Pigmentation

Multispectral Monte-Carlo simulations using the computational skin model with an embedded blood vessel showed how the optical absorption values are affected by melanosome concentration and hence Fitzpatrick skin type [Fig. 2(a)]. In simulations of fully oxygenated blood, we observed an increase in the optical absorption within the skin at all wavelengths with increasing Fitzpatrick type [Fig. 2(b)]. The optical absorption in the epidermis increased more than 10-fold between Fitzpatrick I and Fitzpatrick VI at 700 nm [379 a.u. in FP I, 4128 a.u. in FP VI, Fig. 2(c)]. Spectral coloring is observed within the blood vessel region, which manifests as a decrease in the mean-normalized absorption at lower wavelengths and a corresponding increase at higher wavelengths [Fig. 2(d)]. Absorption measurements within the blood vessel region, however, decrease by a factor of 1.7 from Fitzpatrick I to Fitzpatrick VI [44.6 a.u. in FP I, 26.7 a.u. in FP VI, Fig. 2(e)]. Combined optical and acoustic simulations of a realistic tissue model revealed additional image reconstruction artifacts associated with backprojection, including negative pixels and streaking artifacts (Fig. S4 in the [Supplementary Material](#)). Quantitative results from the blood and skin regions in the acoustic simulations were consistent with those made with solely optical simulations, so the remainder of the simulation study was performed using the Monte-Carlo optical simulation alone for computational practicality.

Spectral coloring observed in the blood vessel biased the linear spectral unmixing blood oxygenation estimate [Fig. 3(a)]. Encouragingly, the bias can be somewhat overcome when using learned spectral unmixing trained on independently generated numerical tissue models containing varying melanin concentrations [Fig. 3(a)]. Considering the spectral changes observed [Figs. 2(d) and 2(e)], it appears that learned unmixing can better cope with the spectral corruption that presents as a reduction in normalized optical absorption at low wavelengths and a relative increase at higher wavelengths for higher Fitzpatrick skin types. With fully oxygenated blood, linear unmixing estimates of photoacoustic sO_2^{EST} reveal a non-linear relationship with Fitzpatrick skin type [Fig. 3(b)] and a linear relationship with melanosome concentration [$p < 0.001$, gradient = 0.218, 95% CI (0.200, 0.235)]. Blood oxygenation estimated with learned

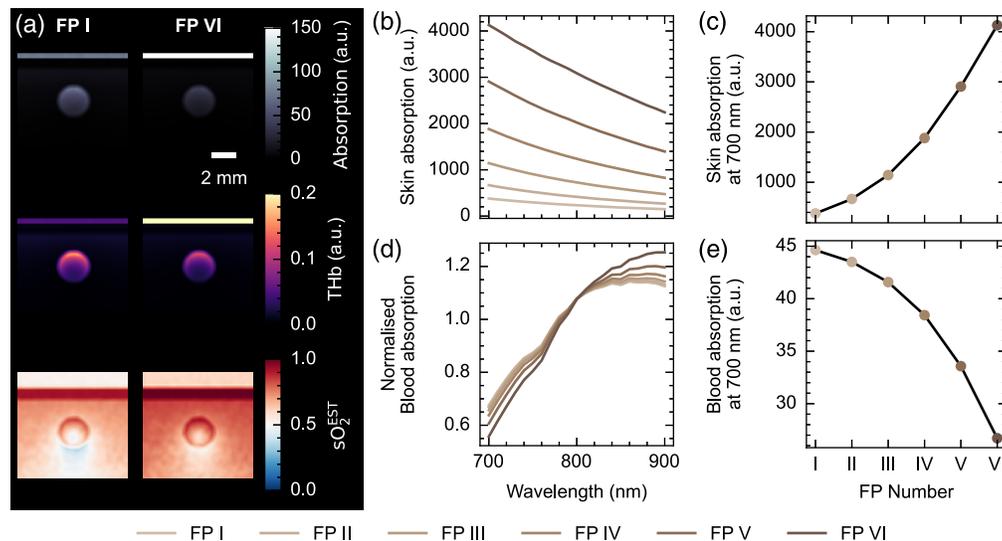


Fig. 2 Monte-Carlo simulations of light transport through the skin model show decreased signal and increased spectral coloring in an underlying blood vessel with increasing Fitzpatrick skin type. (a) Photoacoustic initial pressure (upper), unmixed total hemoglobin (THb, middle), and unmixed blood oxygenation (sO_2^{EST} , lower) in representative examples for fully oxygenated blood in the vessel. (b) Epidermis PA initial pressure as a function of wavelength for all Fitzpatrick types and (c) at 700 nm as a function of Fitzpatrick type. (d) Normalized blood PA initial pressure as a function of wavelength for all Fitzpatrick types and (e) at 700 nm as a function of Fitzpatrick (FP) type. Fitzpatrick types are denoted with graded brown coloration according to the RGB definitions in Fig. 2(d). Scale bar = 2 mm applies to all images in panel (a).

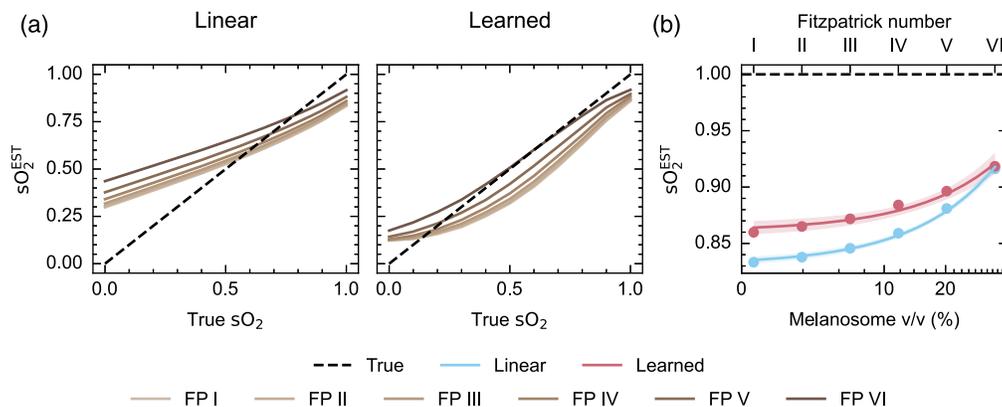


Fig. 3 Estimation of blood oxygenation in simulated data using linear and learned spectral unmixing. (a) sO_2^{EST} estimates increase with increasing Fitzpatrick type when using linear spectral unmixing, whereas estimates are closer to the ground truth when using learned unmixing, particularly at high and low blood oxygenation levels. (b) sO_2^{EST} estimates for fully oxygenated blood using both approaches are shown as a function of Fitzpatrick type and melanosome volume fraction. The shaded area shows 95% confidence intervals.

unmixing also increases linearly with melanosome concentration but with a smaller gradient [$p < 0.001$, gradient = 0.150, 95% CI (0.117, 0.183)] [Fig. 3(b)].

3.3 Blood Flow Phantoms Confirm Melanin-Associated Bias in an Experimental Photoacoustic System

Cylindrical phantoms were enclosed in a melanin-containing outer layer with concentrations equivalent to Fitzpatrick types below I, II, and IV [Fig. 4(a)]. A blood flow circuit was connected

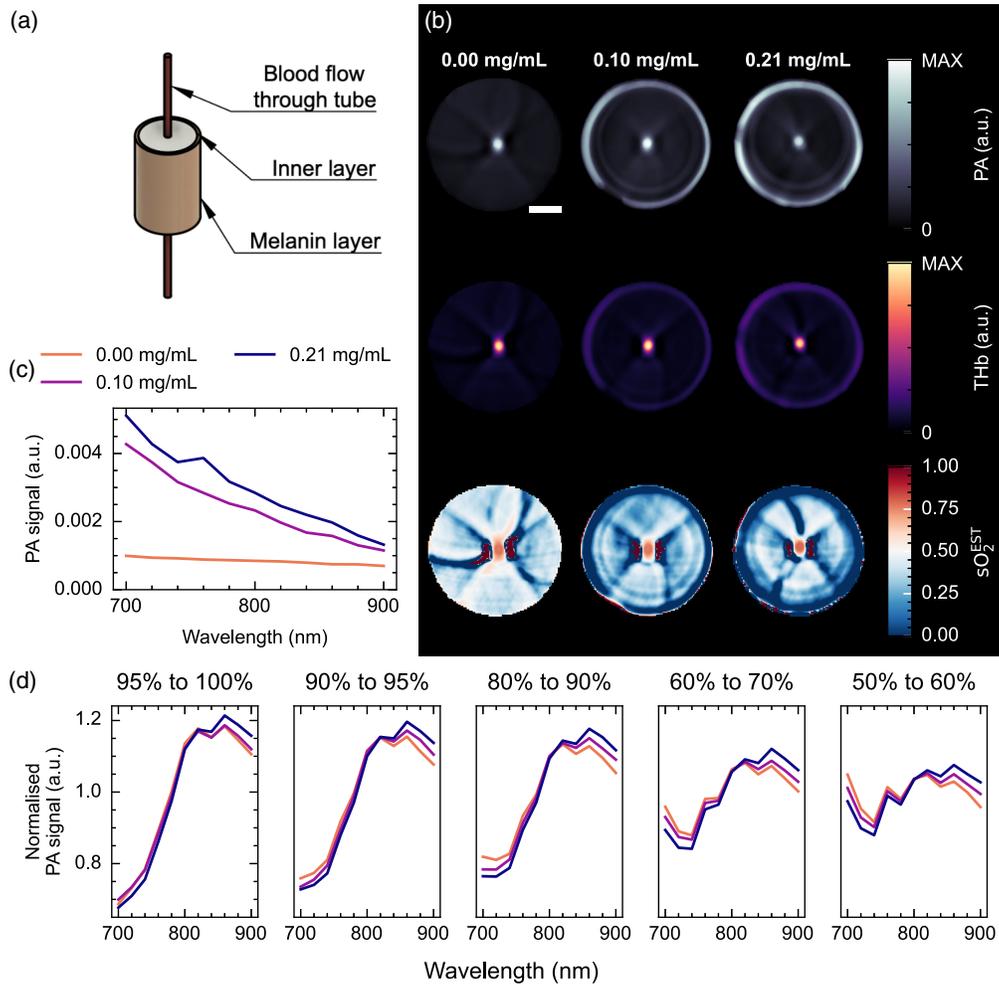


Fig. 4 Skin-mimicking blood-flow phantoms show spectral coloring. (a) Schematic diagram of the agarose phantom setup, showing a 1 mm-thick outer layer containing synthetic melanin and a blood inclusion with varying oxygenation. (b) Reconstructed photoacoustic (PA) images at 800 nm (upper), linear unmixed total hemoglobin (THb, middle), and linear unmixed blood oxygenation (sO_2^{EST} , lower) maps. Image reconstruction artifacts can be observed near the blood inclusion and near the outer melanin layer (particularly clear in sO_2^{EST}). (c) The photoacoustic signal in the outer cylinder of the phantom shows an increase with increasing melanin concentration. (d) Normalized photoacoustic spectra from the blood region show the effect of spectral coloring over a range of blood oxygenation levels. Scale bar = 5 mm.

to a tube through the center of the phantom, allowing continuous monitoring of the blood oxygenation. The optical properties of the phantoms were confirmed using double integrating sphere measurements (Fig. S5 and Table S1 in the [Supplementary Material](#)). By imaging the phantoms over time, as the blood was adjusted from fully oxygenated to fully deoxygenated, the photoacoustic measurements were related to the known blood oxygenation (Fig. S6 in the [Supplementary Material](#)). Images of these phantoms show strong absorption in the melanin outer layer and blood flow tube, with image reconstruction artifacts visible around the blood inclusion and near the melanin layer, particularly visible with the higher concentrations of melanin [Fig. 4(b)]. The photoacoustic signal in the outer layer increases with melanin concentration [Figs. 4(b) and 4(c)]. Wavelength-dependent spectral coloring is observed in the blood spectrum with increasing melanin concentration, with normalized photoacoustic signals relatively lower at shorter wavelengths and higher at longer wavelengths, consistent with the optical absorption properties of melanin [Fig. 4(d)] and simulation findings.

Furthermore, examining the unmixed sO_2^{EST} as a function of the reference blood sO_2 measurements reveals the effect of spectral coloring on PAI signal quantification [Figs. 5(a) and 5(b)],

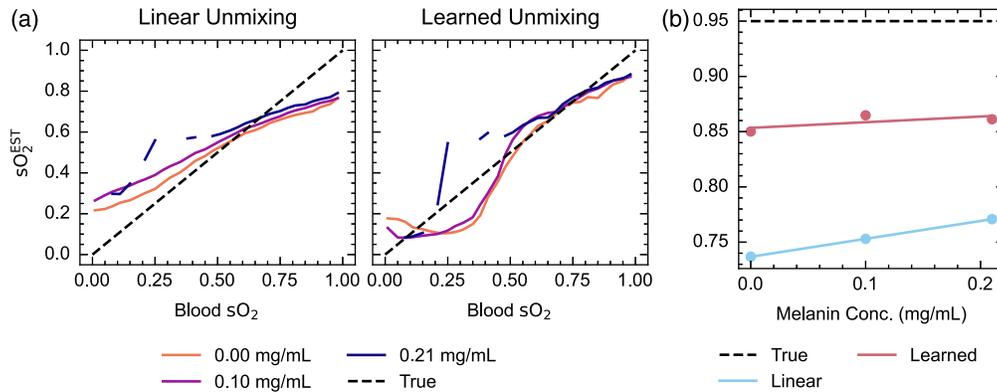


Fig. 5 Comparison of linear unmixing and learned unmixing of blood oxygenation in skin-mimicking phantoms. Linear unmixing and learned unmixing blood oxygenation estimates (sO_2^{EST}) are calculated and compared for varying melanin concentrations (a) Note that discontinuities in this plot arose from the rapid decrease in blood oxygenation from around 0.5 in the 0.21 mg mL^{-1} . This is also shown as a function of melanin concentration at a high blood oxygenation, 95% (b). The line of best fit from a linear model is shown. Linear unmixing sO_2^{EST} increases significantly ($p < 0.001$) with melanin concentration, whereas learned unmixing increases but not significantly ($p = 0.50$).

with higher melanin concentrations giving higher photoacoustic estimates of blood oxygenation when using linear spectral unmixing. Learned unmixing again reduces the disparity between low and high melanin concentrations, particularly at higher blood oxygenation levels [Fig. 5(b)]. Linear unmixing estimates of sO_2^{EST} increase with melanin concentration [$p < 0.001$; gradient = 0.162 mL mg^{-1} , 95% CI (0.161 mL mg^{-1} , 0.163 mL mg^{-1})], whereas learned estimates do not [$p = 0.50$; gradient = 0.051 mL mg^{-1} , 95% CI ($-0.603 \text{ mL mg}^{-1}$, 0.706 mL mg^{-1})]. Analyzing simulations of digital twins of these phantoms in the tomographic PAI system showed consistent trends (Figs. S7 and S8 in the [Supplementary Material](#)).

3.4 Mouse Models Underscore the Severity of the Impact of Skin Pigmentation on Photoacoustics *in vivo*

Two closely related mouse models showing differential pigmentation were used to observe the effects of skin pigmentation on PAI *in vivo* (Figs. S3 and S9 in the [Supplementary Material](#)). Qualitative inspection of the image data clearly shows the increase in photoacoustic signal from the skin region in the presence of pigmentation and an associated decrease in photoacoustic signals measured from the body region in those pigmented mice [Figs. 6(a) and 6(b)]. The spectra extracted from the ROIs show spectral coloring is clearly present in the pigmented mice. Spectra also demonstrate an increased presence of artifacts, such as negative pixels [Fig. 6(b)]. Line profiles across the mice [Figs. 6(c)–6(e)] further emphasize these findings, highlighting the high level of optical absorption in the skin of the pigmented mice, artifactual increase in the total hemoglobin signals from the same region, and an increase in noise in the body of the mouse in the sO_2^{EST} image.

Considering data from a single wavelength [700 nm, Fig. S9(A) in the [Supplementary Material](#)], the photoacoustic signal in the skin region of pigmented mice is significantly higher than in non-pigmented and albino mice ($p = 0.010$). Interestingly, there is also a significantly higher signal in the skin of non-pigmented B6 mice than in albino mice ($p < 0.001$). There was no significant difference between the photoacoustic signal, taken at 700 nm, in the body of the albino mice compared to the non-pigmented B6 mice ($p = 0.95$), however, there is a significantly lower signal in the body of the pigmented B6 mice ($p < 0.001$) [Fig. S9(B) in the [Supplementary Material](#)]. The photoacoustic signal within the skin shows a negative correlation with the photoacoustic signal in the body of the mouse ($p < 0.001$) [Fig. 7(a)]. Unmixed estimates of blood oxygen saturation are significantly higher [Fig. 7(b)] in pigmented mice compared to non-pigmented ($p = 0.025$) and albino mice ($p = 0.019$). There is a strong positive correlation between the skin photoacoustic signal and the sO_2^{EST} in the body of the mouse

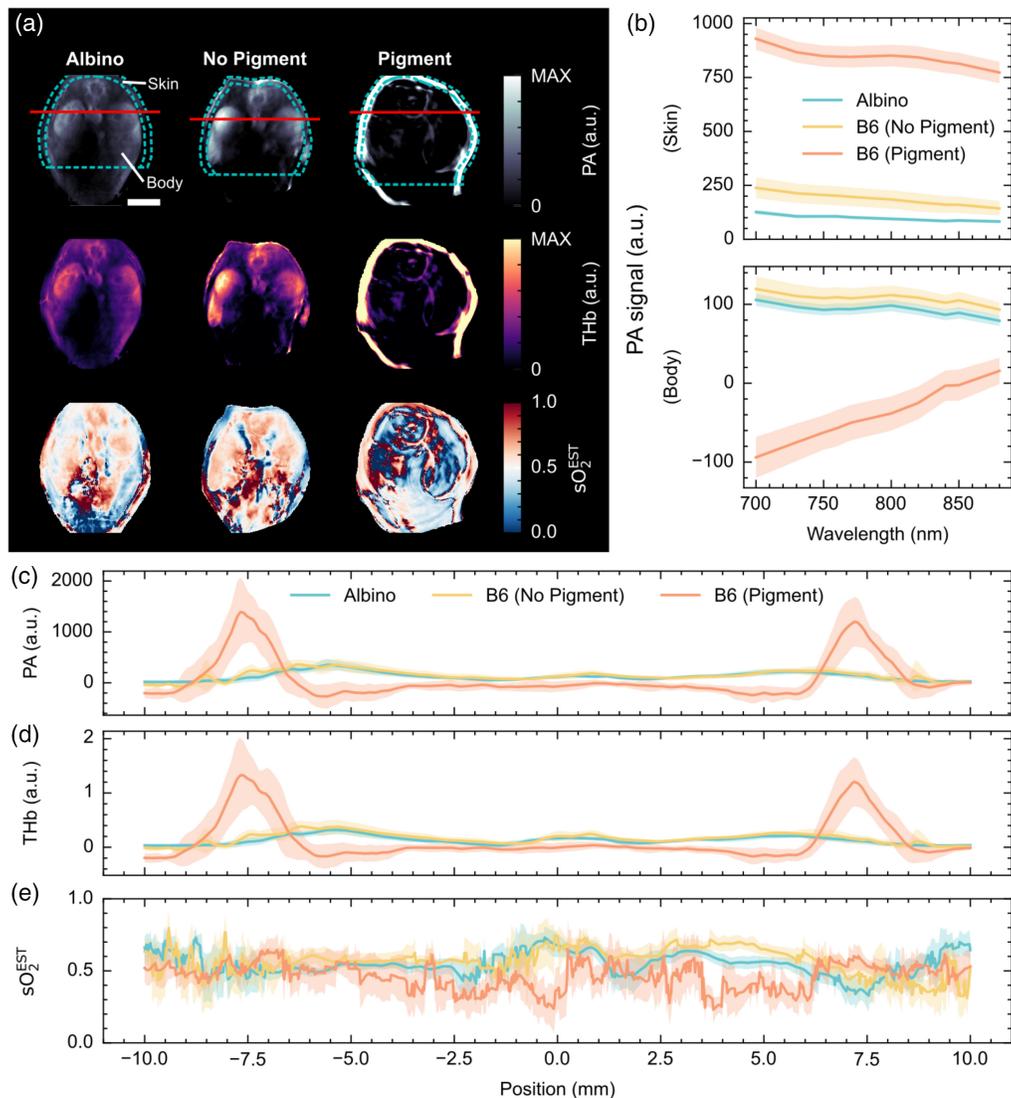


Fig. 6 Photoacoustic images of mouse spleen and kidneys in albino and pigmented mice show substantial variations in signal intensity within the skin region and body regions. (a) Reconstructed photoacoustic images at 700 nm (upper), linear unmixed total hemoglobin (THb, middle), and linear unmixed blood oxygenation (sO_2^{EST} , lower) maps. Regions of interest (body = central region of the blue dashed line, skin = outer region of the blue dashed line) and line profiles used for quantitative analysis are shown on the single wavelength image. (b) Median photoacoustic spectra from the skin and body of albino mice (blue, $n = 10$), non-pigmented B6 mice (yellow, $n = 8$) and pigmented B6 mice (orange, $n = 11$). Line profiles through the kidneys, taken 6.5 mm from the surface of the mouse nearest the spine, of the photoacoustic signal at 700 nm (c), THb (d), and sO_2^{EST} (e). The shaded area shows the standard error. Scale bar = 5 mm.

[Fig. 7(c), $p < 0.001$], demonstrating that a higher melanin pigmentation in the skin leads to an overestimation of the underlying sO_2 . Comparing these results to Fig. 6(e), on the other hand, where negative pixels were included, appears to show a decrease in sO_2^{EST} in pigmented mice compared to albino and non-pigmented mice across the single line plot. These findings highlight the potential influence of different approaches to image reconstruction and quantification on the observed bias.

Learned spectral unmixing was applied to the mouse data, with fair performance in albino and non-pigmented B6 mice. In pigmented mice, no clear improvement was observed over linear unmixing, with reconstruction artifacts dominating the oxygenation maps (Fig. S10 in the [Supplementary Material](#)).

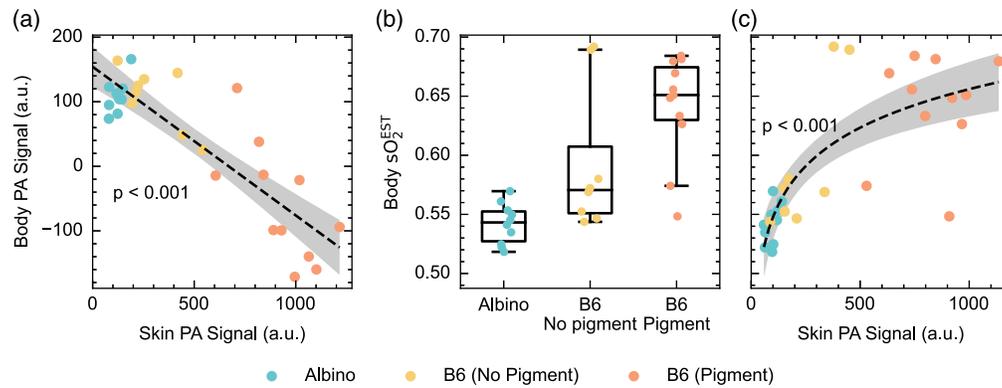


Fig. 7 Increased skin pigmentation reduces the photoacoustic signal in the body, and linear unmixing estimates of blood oxygenation are increased in pigmented mice. (a) The mean skin photoacoustic signal negatively correlates with the mean body photoacoustic signal ($p < 0.001$). (b) sO_2^{EST} , averaged across the body of B6 mice is significantly higher than that of albino mice ($p = 0.025$), and pigmentation is a significant factor in the linear model ($p = 0.019$). (c) sO_2^{EST} averaged across the body region, correlates with skin photoacoustic signal ($p < 0.001$). Averages were taken across the regions defined in Fig. 6.

4 Discussion

Measurement bias in optical oximetry due to the differential light absorption of melanin in darker skin is widely acknowledged, yet the impact on PAI data has barely been tested. Tackling this challenge is key to ensuring equity in the use of PAI and in biomedical optics more generally. To understand the underlying physics behind this phenomenon, we undertook a systematic study using a computational skin model, a simple tissue-mimicking phantom, and animals with varying skin pigmentation on the same genetic background. We then applied two different spectral unmixing approaches, linear unmixing and learned unmixing, revealing how spectral coloring caused by melanin can influence quantitative PAI.

Our findings confirm the presence of a measurement bias in PAI, with increased spectral coloring in the presence of melanin in all models. The spectral coloring trend is consistent with the optical properties of melanin; the absorption coefficient of melanin decreases with wavelength, so light propagation through tissue is more greatly attenuated at shorter wavelengths, a phenomenon that becomes more apparent as melanin concentration increases. Spectral coloring leads to a substantial overestimation of unmixed sO_2^{EST} with increasing Fitzpatrick skin type. The difference in sO_2^{EST} quantification between Fitzpatrick I and Fitzpatrick VI may be as large as 9 percentage points based on our simulations, which is similar to typical differences observed between cancer models.⁵⁴ It is, therefore, unlikely that sO_2 derived from linear spectral unmixing in its present form could be applied equitably when translated into the clinic for cancer detection or staging. The measurement of relative changes may be more robust to measurement biases, however, translation to the in vivo setting remains challenging, given the need for paired measurements and ground-truth data.

Fortunately, we observed that fundamental limitations in linear spectral unmixing, which have been characterized previously,^{4,49,50} may be partially mitigated using a learned spectral unmixing approach. Learned spectral unmixing demonstrated improved quantification relative to the ground truth in simulations and phantoms, and reduced the disparity between Fitzpatrick skin types, particularly in highly oxygenated blood. Learned unmixing, however, could not overcome the fundamental limitations of our backprojection-based reconstruction pipeline, which cannot accurately reconstruct the true initial-pressure distribution, particularly in regions adjacent to strongly absorbing structures. Strong optical absorbers lead to significant artifacts, such as streaking and negative pixels, as can be seen clearly in the center of the images of heavily pigmented mice.⁵⁵ Blood oxygen quantification was poor in negative pixel regions both with linear and learned unmixing.

Our study showed consistent agreement between Monte-Carlo simulations, phantoms, and in vivo studies, both in quantitative measures and image reconstruction artifacts. We expect our

observations to transfer well into human data, however, there are further features of *in vivo* imaging that were not considered here and could further limit the use of PAI equitably in the clinic. In this study, we did not consider the effects of high acoustic impedance structures, such as bones, which could reflect the downward propagating photoacoustic wave from the epidermis, giving image reconstruction artifacts. Furthermore, the appearance of image reconstruction artifacts would be dependent on the ultrasound detection geometry, which differs between the pre-clinical photoacoustic system used here, and typical handheld photoacoustic systems used in a clinical setting. The effects of noise were not considered in our simulations, which would dominate in regions of low light fluence, such as deep in the tissue, particularly in the higher Fitzpatrick skin types. Additional noise could adversely affect the spectral unmixing results, particularly as the energy of different wavelengths will vary, and so affect the noise profiles. The presence of noise, acoustic reflection artifacts, and differences in ultrasound detection geometry, may explain the disagreement between this study and recently published *in vivo* data, which, contrary to this study, showed a decrease in linear spectral unmixing estimates of sO_2 with increasing Fitzpatrick type.²⁹ Future studies should explore ways to improve image reconstruction in the presence of strong optical absorbers and acoustic heterogeneity, for example with model-based reconstruction methods. A substantial expansion of the available data from human subjects of varying skin pigmentation is required to thoroughly characterize this phenomenon across a range of PAI systems and methods.

5 Conclusion

PAI demonstrates a clear measurement bias with the presence of melanin in the skin. Spectral unmixing results were highly dependent on melanin concentration, which could introduce unintended biases into future human photoacoustic studies. Reproducible and equitable application of quantitative PAI will likely require a combination of physics-based, data-driven, and empirical methods to account for varying light fluence and tissue properties fairly across the population.

Disclosures

The authors have no conflicts of interest to declare.

Code and Data Availability

All data and code used in the preparation of this paper and not otherwise referenced in-line will be made available on the University of Cambridge Apollo repository (<https://doi.org/10.17863/CAM.100220>) or GitHub (<https://github.com/BohndiekLab/melanin-phantom-simulation-paper>), respectively.

Acknowledgments

This work was funded by: Cancer Research UK (Grant No. SB, TRE; C9545/A29580); the MedAccel program of the National Physical Laboratory financed by the Department for Business, Energy and Industrial Strategy's Industrial Strategy Challenge Fund (LH); Cancer Research UK RadNet Cambridge [Grant No. C17918/A28870] (EB); the Herchel Smith Studentship, Cambridge Trust (RT); and the Walter Benjamin Stipendium of the Deutsche Forschungsgemeinschaft (JG). We thank the Cancer Research UK Cambridge Institute Imaging Core and Biological Resources Unit for their support in conducting this research.

References

1. P. Beard, "Biomedical photoacoustic imaging," *Interface Focus* **1**, 602–631 (2011).
2. J. Weber, P. C. Beard, and S. E. Bohndiek, "Contrast agents for molecular photoacoustic imaging," *Nat. Methods* **13**, 639–650 (2016).
3. F. M. Brochu et al., "Towards quantitative evaluation of tissue absorption coefficients using light fluence correction in optoacoustic tomography," *IEEE Trans. Med. Imaging* **36**, 322–331 (2017).
4. R. Hochuli et al., "Estimating blood oxygenation from photoacoustic images: can a simple linear spectroscopic inversion ever work?," *J. Biomed. Opt.* **24**, 121914 (2019).

5. J. A. Iglesias-Guitian et al., "A biophysically-based model of the optical properties of skin aging," *Comput. Graphics Forum* **34**, 45–55 (2015).
6. K. Poorzargar et al., "Accuracy of pulse oximeters in measuring oxygen saturation in patients with poor peripheral perfusion: a systematic review," *J. Clin. Monit. Comput.* **36**, 961–973 (2022).
7. V. G. Kanellis, "A review of melanin sensor devices," *Biophys. Rev.* **11**, 843–849 (2019).
8. A. T. Adams et al., "Equity-driven sensing system for measuring skin tone-calibrated peripheral blood oxygen saturation (OptoBeat): development, design, and evaluation study," *JMIR Biomed. Eng.* **7**, e34934 (2022).
9. J. N. Adler et al., "Effect of skin pigmentation on pulse oximetry accuracy in the emergency department," *Acad. Emerg. Med.* **5**, 965–970 (1998).
10. P. E. Bickler, J. R. Feiner, and J. W. Severinghaus, "Effects of skin pigmentation on pulse oximeter accuracy at low saturation," *Anesthesiology* **102**, 715–719 (2005).
11. C. J. Crooks et al., "Pulse oximeter measurements vary across ethnic groups: an observational study in patients with COVID-19," *Eur. Respir. J.* **59**, 2103246 (2022).
12. S. J. Ebmeier et al., "A two centre observational study of simultaneous pulse oximetry and arterial oxygen saturation recordings in intensive care unit patients," *Anaesth. Intensive Care* **46**, 297–303 (2018).
13. J. R. Feiner, J. W. Severinghaus, and P. E. Bickler, "Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender," *Anesth. Analg.* **105**, S18–S23 (2007).
14. N. R. Henry et al., "Disparities in hypoxemia detection by pulse oximetry across self-identified racial groups and associations with clinical outcomes," *Crit. Care Med.* **50**, 204–211 (2022).
15. A. L. Holder and A. K. I. Wong, "The big consequences of small discrepancies: why racial differences in pulse oximetry errors matter," *Crit. Care Med.* **50**, 335–337 (2022).
16. A. Jubran and M. J. Tobin, "Reliability of pulse oximetry in titrating supplemental oxygen therapy in ventilator-dependent patients," *Chest* **97**, 1420–1425 (1990).
17. M. W. Sjoding et al., "Racial bias in pulse oximetry measurement," *N. Engl. J. Med.* **383**, 2477–2478 (2020).
18. Z. Vesoulis et al., "Racial discrepancy in pulse oximeter accuracy in preterm infants," *J. Perinatol.* **42**, 79–85 (2022).
19. L. Holland and K. Blick, "Implementing and validating transcutaneous bilirubinometry for neonates," *Am. J. Clin. Pathol.* **132**, 555–561 (2009).
20. P. Szabo et al., "Detection of hyperbilirubinaemia in jaundiced full-term neonates by eye or by bilirubinometer?," *Eur. J. Pediatrics* **163**, 722–727 (2004).
21. B. Bent et al., "Investigating sources of inaccuracy in wearable optical heart rate sensors," *NPJ Digital Med.* **3**, 1–9 (2020).
22. P. J. Colvonen et al., "Limiting racial disparities and bias for wearable devices in health science research," *Sleep* **43**, zsaal159 (2020).
23. P. J. Colvonen, "Response to: investigating sources of inaccuracy in wearable optical heart rate sensors," *NPJ Digital Med.* **4**, 1–2 (2021).
24. A. Shcherbina et al., "Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort," *J. Pers. Med.* **7**(2), 3 (2017).
25. A. Afshari et al., "Evaluation of the robustness of cerebral oximetry to variations in skin pigmentation using a tissue-simulating phantom," *Biomed. Opt. Express* **13**, 2909–2928 (2022).
26. T. Phan et al., "Quantifying the confounding effect of pigmentation on measured skin tissue optical properties: a comparison of colorimetry with spatial frequency domain imaging," *J. Biomed. Opt.* **27**, 036002 (2022).
27. X. Li et al., "Optoacoustic mesoscopy analysis and quantitative estimation of specific imaging metrics in Fitzpatrick skin phototypes II to V," *J. Biophotonics* **12**, e201800442 (2019).
28. G. S. P. Fernandes et al., "Impact of skin pigmentation on photoacoustic imaging using linear array transducer: a pilot in vivo study," in *IEEE Int. Ultrasonics Symp. (IUS)*, 2022-October, IEEE, pp. 1–4 (2022).
29. Y. Mantri and J. V. Jokerst, "Impact of skin tone on photoacoustic oximetry and tools to minimize bias," *Biomed. Opt. Express* **13**(2), 875 (2022).
30. W. Yim, et al., "3D-bioprinted phantom with human skin phototypes for biomedical optics," *Adv. Mater.* **35**, 2206385 (2023).
31. J. Gröhl et al., "Learned spectral decoloring enables photoacoustic oximetry," *Sci. Rep.* **11**, 6565 (2021).
32. S. Tzoumas et al., "Eigenspectra optoacoustic tomography achieves quantitative blood oxygenation imaging deep in tissues," *Nat. Commun.* **7**, 12121 (2016).
33. I. Olefir et al., "Deep learning-based spectral unmixing for optoacoustic imaging of tissue oxygen saturation," *IEEE Trans. Med. Imaging* **39**, 3643–3654 (2020).
34. S. Tzoumas and V. Ntziachristos, "Spectral unmixing techniques for optoacoustic imaging of tissue pathophysiology," *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **375**, 20170262 (2017).
35. C. Cai et al., "End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging," *Opt. Lett.* **43**, 2752 (2018).

36. S. L. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**, R37–R61 (2013).
37. S. L. Jacques, "Origins of tissue optical properties in the UVA, visible, and NIR regions," in *OSA TOPS on Adv. in Opt. Imaging and Photon Migration*, pp. 364–371 (1996).
38. A. Chardon, I. Cretois, and C. Hourseau, "Skin colour typology and suntanning pathways," *Int. J. Cosmet. Sci.* **13**, 191–208 (1991).
39. C. C. Cooksey, D. W. Allen, and B. K. Tsai, "Reference data set of human skin reflectance," *J. Res. Natl. Inst. Stand. Technol.* **122**, 26 (2017).
40. 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**, 559 (1993).
41. N. G. Jablonski and G. Chaplin, "The evolution of human skin coloration," *J. Hum. Evol.* **39**, 57–106 (2000).
42. T. Mansencal et al., "Colour 0.4.2," (2022).
43. 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 (2009).
44. J. Gröhl et al., "SIMPACT: an open-source toolkit for simulation and image processing for photonics and acoustics," *J. Biomed. Opt.* **27**, 083010 (2022).
45. A. Stanzola et al., "j-wave: an open-source differentiable wave simulator," *SoftwareX* **22**, 101338 (2023).
46. J. Joseph et al., "Evaluation of precision in photoacoustic tomography for preclinical imaging in living subjects," *J. Nucl. Med.* **58**, 807–814 (2017).
47. S. A. Prahl, "Everything I think you should know about inverse adding–doubling," Oregon Medical Laser Center, St. Vincent Hospital 1344, pp. 1–74 (2011).
48. B. Pierscionek, "Refractive index of the human lens surface measured with an optic fibre sensor," *Ophthalm. Res.* **26**(1), 32–35 (1994).
49. M. Gehring, S. E. Bohndiek, and J. Brunner, "Development of a blood oxygenation phantom for photoacoustic tomography combined with online pO₂ detection and flow spectrometry," *J. Biomed. Opt.* **24**, 121908 (2019).
50. L. Hacker et al., "Photoacoustics resolves species-specific differences in hemoglobin concentration and oxygenation," *J. Biomed. Opt.* **25**, 095002 (2020).
51. J. W. Severinghaus, "Simple, accurate equations for human blood O₂ dissociation computations," *J. Appl. Physiol.: Respir. Environ. Exerc. Physiol.* **46**(3), 599–602 (1979).
52. P. Virtanen et al., "SciPy 1.0: fundamental algorithms for scientific computing in Python," *Nat. Methods* **17**, 261–272 (2020).
53. M. Osto et al., "Individual typology angle and Fitzpatrick skin phototypes are not equivalent in photodermatology," *Photochem. Photobiol.* **98**, 127–129 (2022).
54. I. Quiros-Gonzalez et al., "Photoacoustics delineates murine breast cancer models displaying angiogenesis and vascular mimicry," *Br. J. Cancer* **118**, 1098–1106 (2018).
55. X. L. Deán-Ben and D. Razansky, "Photoacoustic image formation approaches—a clinical perspective," *Phys. Med. Biol.* **64**, 18TR01 (2019).

Thomas R. Else completed his MSci degree in experimental and theoretical physics in 2019 and is currently pursuing a PhD in medical sciences, both at the University of Cambridge. His research focuses on clinical translation of photoacoustic imaging, with a focus on equitable application of the technology through open access software and evaluation of skin tone biases.

Lina Hacker completed her PhD in medical sciences at the University of Cambridge focusing on technical and biological validation of photoacoustic imaging biomarkers. She is now a Junior Research Fellow at the University of Oxford.

Janek Gröhl received his master's degree in medical computer science from the University of Heidelberg and Heilbronn University of Applied Sciences in 2016. He received his PhD from the medical faculty of the University of Heidelberg in April 2021 and did his PhD research under Prof. Lena Maier-Hein at the German Cancer Research Center. Since December 2020, he is working as a postdoctoral fellow with Prof. Sarah Bohndiek funded by the Walter Benjamin Programme of the German Research Foundation focusing on data-driven quantitative photoacoustic imaging.

Ellie V. Bunce completed her MPharmacol degree (Hons) at the University of Bath in 2021. She is currently pursuing a PhD in medical sciences at the University of Cambridge, investigating the potential of vascular targeted therapies to enhance cancer radiotherapy response using novel imaging approaches.

Ran Tao completed her MRes in electronics and photonics at the University of Cambridge in 2021. She is currently pursuing a PhD in physics at the University of Cambridge, focusing on quantitative optical property estimation for early cancer detection.

Sarah E. Bohndiek completed her PhD in radiation physics at the University College London in 2008 and then worked in both the United Kingdom (at Cambridge) and the United States (at Stanford) as a postdoctoral fellow in molecular imaging. Since 2013, she has been a Group Leader at the University of Cambridge, where she is jointly appointed in the Department of Physics and the Cancer Research UK Cambridge Institute. She was appointed as full professor of biomedical physics in 2020. Sarah was recently awarded the CRUK Future Leaders in Cancer Research Prize and SPIE Early Career Achievement Award in recognition of her innovation in biomedical optics.