Photoacoustic discrimination of vascular and pigmented lesions using classical and Bayesian methods

Jennifer A. Swearingen
University of Missouri
Department of Dermatology
One Hospital Drive, MA111
Columbia, Missouri 65211

Scott H. Holan
University of Missouri
Department of Statistics
134F Middlebush Hall
Columbia, Missouri 65211

Mary M. Feldman
ProPath Laboratories
8267 Elmbrook Drive, Suite 100
Dallas, Texas 75247

John A. Viator
University of Missouri
Department of Dermatology
and
Department of Biological Engineering
240C Christopher S. Bond Life Sciences Center
1201 East Rollins Road
Columbia, Missouri, 65211

Abstract. Discrimination of pigmented and vascular lesions in skin can be difficult due to factors such as size, subungual location, and the nature of lesions containing both melanin and vascularity. Misdiagnosis may lead to precancerous or cancerous lesions not receiving proper medical care. To aid in the rapid and accurate diagnosis of such pathologies, we develop a photoacoustic system to determine the nature of skin lesions *in vivo*. By irradiating skin with two laser wavelengths, 422 and 530 nm, we induce photoacoustic responses, and the relative response at these two wavelengths indicates whether the lesion is pigmented or vascular. This response is due to the distinct absorption spectrum of melanin and hemoglobin. In particular, pigmented lesions have ratios of photoacoustic amplitudes of approximately 1.4 to 1 at the two wavelengths, while vascular lesions have ratios of about 4.0 to 1. Furthermore, we consider two statistical methods for conducting classification of lesions: standard multivariate analysis classification techniques and a Bayesian-model-based approach. We study 15 human subjects with eight vascular and seven pigmented lesions. Using the classical method, we achieve a perfect classification rate, while the Bayesian approach has an error rate of 20%.
[DOI: 10.1117/1.3316297]

Keywords: acoustic transducer; classification; optoacoustic; polyvinylidene fluoride; ultrasonic.

Paper 09093RR received Mar. 18, 2009; revised manuscript received Dec. 15, 2009; accepted for publication Dec. 15, 2009; published online Mar. 1, 2010.

1 Introduction

Discrimination of vascular and pigmented lesions may pose a challenging task for smaller lesions for which gross examination is inconclusive. Additionally, other factors may contribute to the difficulty of lesion classification, such as subungual formations where optical attenuation and scattering are caused by the compacted keratin structure of the nail.[1] The clinical danger of misclassification arises when a melanoma or a precancerous pigmented lesion is identified as a vascular lesion. To improve diagnosis, we developed a photoacoustic probe used previously in skin measurements[2–4] and used classical and Bayesian statistical classification methods for differentiating between pigmented and vascular lesions.

Several approaches exist for diagnosis of pigmented and vascular lesions, including gross examination, dermoscopy, digital imaging, and ultimately biopsy. Each has its advantages and shortcomings. The photoacoustic procedure described in this paper provides an alternative that is rapid, noninvasive, and relies on the distinct spectral structures in absorption for melanin and hemoglobin. These differences provide robust information in the photoacoustic response that can be used for statistical classification of the lesions. Although the photoacoustic procedure requires specialized laser instrumentation, the classification methods can be adapted to simpler systems, including digital imaging and white light spectroscopy. However, modifications of the photoacoustic method may allow for depth profiling and imaging, in addition to lesion classification.

Atypical vascular structures within pigmented lesions may confound the ability to differentiate skin pathologies.[5,6] In these and other cases, dermoscopy has been used to diagnose skin lesions as it provides the ability to examine small structures not normally visible to the human eye. Dermoscopy has been used extensively in diagnosing pigmented lesions as well as vascular lesions.[7] This device is particularly useful for diagnosis of a Spitz nevus that may be red in color and may resemble a vascular lesion.[8] In amelanotic lesions, however, the efficacy of dermoscopy is reduced.[9,10] Furthermore, the sensitivity of dermoscopy varies based on the clinician's experience.[11]

Spectrophotometric intracutaneous analysis, otherwise known as SIAscopy, uses spectral information to image pigmented structures in skin.[12] While the imaging instrumentation is simple, the extraction algorithms use inverse computations that are intensive and may yield unstable solutions, resulting in bad imaging information. Other systems using digital imaging utilize polarized light to extract information

Address all correspondence to John A. Viator, University of Missouri, 240C Bond Life Sciences Center, 1201 East Rollins Road, Columbia, Missouri 65211-7310. Tel: 573-884-2862; Fax: 573-884-9676; E-mail: viatorj@missouri.edu

1083-3668/2010/15(1)/016019/9/$25.00 © 2010 SPIE

about skin pigmentation[13] and have been successful in differentiating pigmented and vascular lesions.[14] The approach described in this paper focuses on simple photoacoustic predictors for model-based classification of skin lesions that provide measures of uncertainty in group membership assignment.

Biomedical photoacoustics uses pulsed laser irradiation to transduce optical energy into acoustic waves. This transduction serves two purposes. First, it enables selective targeting of subsurface light absorbers due to the great differences in optical absorption of tissue chromophores. Second, it enables robust propagation of information embedded in the resulting acoustic wave, as such waves travel with much less attenuation in tissue compared to photons propagating in turbid media. In our application, photoacoustics arise when pulsed laser irradiation induces rapid thermoelastic expansion in skin chromophores. Unlike photoacoustic spectroscopy,[15] which uses modulated continuous wave irradiation, photoacoustic generation by thermoelastic expansion uses nanosecond pulsed laser energy. This energy is absorbed by a small volume such that subsequent heating induces rapid expansion. This rapid expansion results in a mechanical disturbance in tissue that manifests itself as a transient pulse of acoustic energy.

We developed a photoacoustic probe that incorporates an optical fiber and a piezoelectric sensor in a small handpiece to induce and detect acoustic waves generated at two key laser wavelengths in human skin. The handpiece is small and is easy to place on most areas of skin, except for places such as the ears and nose, where the curvature is too great for proper probe placement. The two key wavelengths are 422 and 530 nm, purple and green light, respectively, and are two isosbestic points in the hemoglobin absorption spectrum where the absorption of oxy- and deoxygenated hemoglobin is equal. Thus, analysis of the photoacoustic waves did not depend on hemoglobin oxygen saturation. Also, these two wavelengths assured high absorption in pigmented and vascular structures in skin.

In the photoacoustic procedure, laser energy was delivered via the optical fiber probe to the skin surface, where it was absorbed by melanin and blood vessels. The short pulse duration of the laser ensured transduction of optical energy into acoustic waves, analysis of which gives the exact optical energy absorption. The amount of energy absorbed by the lesion chromophores depended on the laser fluence at the lesion and the optical absorption coefficient at the given wavelength. The amplitude of the resulting photoacoustic wave was proportional to these quantities, thus a simple analysis of the photoacoustic amplitudes at 422 and 530 nm was possible for lesion classification.

In this study, we tested 15 human subjects by irradiated lesions in skin. Of the 15 subjects, 8 had vascular lesions and 7 had pigmented lesions. Owing to the distinct spectral absorption of melanin and hemoglobin, the ratio of the photoacoustic amplitudes at these wavelengths are distinct for the two chromophores, with melanin giving a ratio of about 1.4 to 1 and hemoglobin giving a ratio closer to 4.0 to 1. Based on these distinctly different ratios, we used two classification methods to distinguish between melanoma and vascular lesions based on different optical properties.

The first classification rule we propose begins by assuming that there exist two distinct populations of lesions as defined by their photoacoustic response at 422 and 530 nm. Then classification into one of these populations proceeds using standard multivariate classification techniques[16,17] that rely only on the first two sample moments (i.e., mean and variance) of the photoacoustic ratios of the lesions and the assumption that the photoacoustic ratios follow a Gaussian distribution. Henceforth, we refer to this as the standard approach. Note that although our rule under the standard approach is developed using only photoacoustic ratio as a predictor, other predictors can be incorporated into the rule with straightforward modification to the existing rule.[16]

To develop this classification rule we created a training set of photoacoustic measurements on known samples. We began by assuming that the probability density functions (pdfs) associated with the two populations, vascular and pigmented lesions, were Gaussian. However, we acknowledge that this assumption may not be satisfied. Nonetheless, the rule we propose is shown empirically to be robust to departures from this assumption.

The second classification rule we proposed was based on a Bayesian probit regression model that used the photoacoustic ratio at 422 and 530 nm and the age of the subject as predictors. In particular, we considered two different models. The first model used the photoacoustic ratio as the sole predictor, whereas the second model used both the photoacoustic ratio and age as predictors.

In all cases, the classification rules were developed using the fifteen human subject profiles. For the in sample (IS) results the classification rules were developed using information from all 15 subjects. Note that this technique can cause a bias since subject information is used both to develop the rule as well as to classify. To alleviate this source of bias we also developed our classification rules using a cross validation (CV) (or leave-one-out) approach. This method uses 14 subjects to create a rule used to classify the lesion of the 15th subject (i.e., the subject left out). The results presented here are for the analysis conducted using IS.

Using the standard method, we achieved an error rate of 0% in classifying pigmented and vascular lesions. In contrast, using the Bayesian method, we achieved an error rate of 47% for the model using only the photoacoustic ratio and an error rate of 20% for the model using the photoacoustic ratio and age. Though the standard approach had a smaller error rate, the Bayesian models provided a measure of uncertainty in the classification. For instance, a lesion classified as vascular using the Bayesian method would also include information regarding the distribution of the estimated probability of group membership (i.e., this approach provides a natural measure of uncertainty on the probability of group membership for a given subject) so that a high amount of uncertainty in assigning group membership could prompt the clinician to perform further diagnosis.

2 Materials and Methods

2.1 *Photoacoustic Apparatus*

The photoacoustic device we used was similar to that used in our previous work in quantifying epidermal melanin with minor modifications.[4] The laser system was coupled to an optical parametric oscillator (OPO) to enable tuning the system to more than one output wavelength. The apparatus consisted of

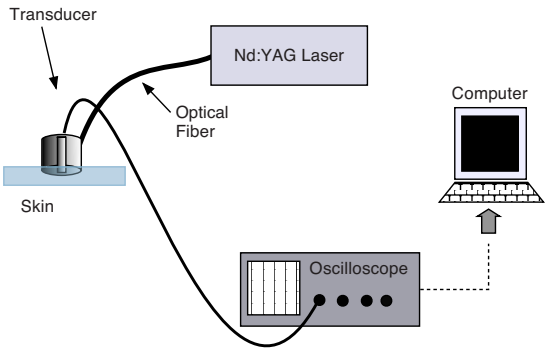

Fig. 1 Apparatus for photoacoustic depth determination of melanin content. The frequency-doubled 532-nm Nd:YAG laser had a pulse duration of 5 ns. Laser light was delivered to the probe via 1000-μm-diam optical fiber. A PVDF element within the probe detected photoacoustic waves.

a frequency-tripled Nd:YAG laser operating at 355 nm pumping an OPO (Opotek, Carlsbad, California) for a tunable output from 410 to 2400 nm (Fig. 1). The pulse duration was 5 ns. The OPO was coupled to a 1000-μm optical fiber (Thorlabs, Newton, New Jersey) that terminated in a cylindrical acrylic handpiece. This handpiece was put in contact with the skin of a human subject to measure photoacoustic waves generated in skin lesions. The handpiece was 16 mm in diameter and approximately 22 mm long. Within the handpiece was a miniature acoustic sensor made from polyvinylidene difluoride (PVDF). The design of the probe is more fully described in Viator et al.[2] The optical fiber was directed to a 1.5-mm spot at the bottom of the probe, which was placed onto the skin surface for measurements. Laser energy ranged from 3 to 6 mJ/pulse in all experiments, though pulse energy was determined to 5% accuracy for each measurement. Radiant exposures were approximately 0.15 to 0.35 J/cm^2. Laser spot size was estimated to be 1 to 2 mm with a Gaussian profile output from the optical fiber. The voltage waveform from the PVDF sensor was transmitted through a semirigid coaxial cable to a 350-MHz instrumentation amplifier (SR445, Stanford Research Systems, Sunnyvale, California) with a gain of 125. The amplified signal was sent to an oscilloscope (TDS 2024, Tektronix, Wilsonville, Oregon) with a bandwidth of 200 MHz and was triggered by a photodiode (DET-210, Thorlabs, Inc., Newton, New Jersey) that monitored laser output. The probe is shown in Fig. 2.

2.2 *Human Subjects*

We tested 15 human subjects with vascular or pigmented lesions. Eight subjects had vascular lesions and seven had pigmented lesions. This study was approved by the Institutional Review Board at the University of Missouri (protocol number 1089418) and informed and signed consent was obtained from all volunteers according to the Declaration of Helsinki. To ensure proper placement of the photoacoustic probe, we placed a plastic annulus centered on the lesion. The inner diameter of the annulus was equal to the outer diameter of the circular probe. Thus, since the acoustic sensor was centered within the probe, the sensor was centered above the lesion. With a laser spot size of less than 2 mm, the laser spot was generally within each lesion that we measured.

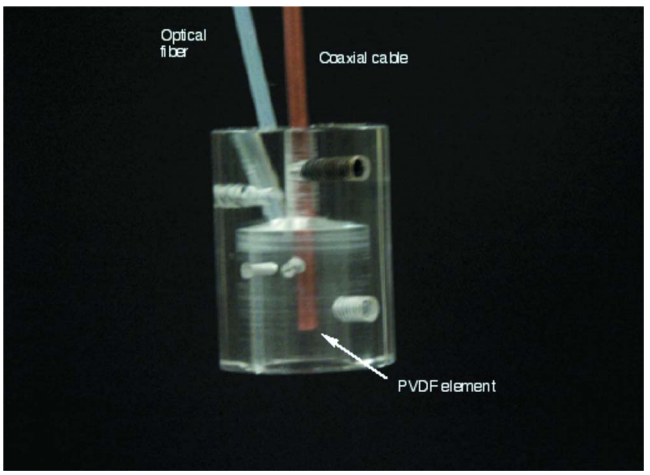

Fig. 2 Photoacoustic probe. The optical fiber delivered 532-nm laser light with a pulse duration of 5 ns. Acoustic detection was obtained by a PVDF element within the probe cavity. The PVDF was recessed approximately 3 mm from the bottom of the probe, thus initial photoacoustic waves were delayed approximately 2 μs from the laser pulse.

2.3 *Photoacoustic Analysis*

Analysis of photoacoustic data was simple and straightforward. We determined the amplitude of the photoacoustic wave by a peak-to-peak measurement. This method was used as the complicated geometry of the vasculature and pigmented structures almost ensured high diffractive content in the acoustic waves, making a textbook linear analysis impossible. However, the peak-to-peak amplitude provided a quick, unambiguous measure of the photoacoustic response in the lesion. The amplitudes at each wavelength were normalized by pulse energy, then the normalized amplitude at 422 nm was divided by the normalized amplitude at 530 nm, giving the photoacoustic ratio. These ratios were used in the subsequent statistical analysis.

2.4 *Statistical Classification Methods*

2.4.1 *Classical approach*

Following the notation and exposition of Johnson and Wichern[16] and Talbert et al.,[17] we begin by stipulating the photoacoustic ratios come from two distinct populations: Π_1, the pigmented lesion, and Π_2, the vascular lesion. Let $f_1(x)$ and $f_2(x)$ denote the pdfs associated with the random variable x of the two populations, where x denotes the ratio of photoacoustic response. Further, let $p(2|1)$ be the conditional probability of classifying a ratio as belonging to Π_2 when it belongs to Π_1 and $p(1|2)$ be defined similarly. In this simplified analysis, we assume equal costs of misclassification (see Ref. 16 for further discussion). Thus, for any classification rule, the average or expected cost of misclassification (ECM) is given by $\mathrm{ECM}=p(2|1)p_1+p(1|2)p_2$, where $p_i(i=1,2)$ is the prior probability of Π_i and $p_1+p_2=1$. A reasonable classification rule would seek to minimize ECM. Let Ω denote the sample space of all possible observations, R_1 the set of all x observations from Π_1, and $R_2=\Omega-R_1$. The rule that minimizes ECM is then given by

$$R_1: \frac{f_1(x)}{f_2(x)} \geqslant \left(\frac{p_2}{p_1}\right), \quad R_2: \frac{f_1(x)}{f_2(x)} < \left(\frac{p_2}{p_1}\right).$$

In the case that we that we assume $p_1=p_2=1/2$, this rule simplifies to

$$R_1: \frac{f_1(x)}{f_2(x)} \geqslant 1 \quad R_2: \frac{f_1(x)}{f_2(x)} < 1.$$

To implement these rules we evaluate the density function ratio at a new value x_0. Furthermore, since we do not know the population first and second moments of the pdfs, we substitute the sample quantities to establish a classification rule. Specifically, for normally distributed variables and assuming $\sigma_1^2 \neq \sigma_2^2$, yields the rule

$$\ln\left(\frac{s_2}{s_1}\right) + \frac{1}{2}\left[\frac{(x_0-\overline{x_2})^2}{s_2^2} - \frac{(x_0-\overline{x_1})^2}{s_1^2}\right] \geqslant \ln\left(\frac{p_2}{p_1}\right) \Rightarrow x_0 \in \Pi_1, \tag{1}$$

$$\ln\left(\frac{s_2}{s_1}\right) + \frac{1}{2}\left[\frac{(x_0-\overline{x_2})^2}{s_2^2} - \frac{(x_0-\overline{x_1})^2}{s_1^2}\right] < \ln\left(\frac{p_2}{p_1}\right) \Rightarrow x_0 \in \Pi_2, \tag{2}$$

where s_i^2, $i=1,2$, denotes the sample variances. This rule allocates the photoacoustic ratio to the vascular or pigmented lesion populations. Assuming $\sigma_1^2=\sigma_2^2$, Eqs. (1) and (2) can be simplified to

$$\frac{(x_0-\overline{x_2})^2 - (x_0-\overline{x_1})^2}{2s^2} \geqslant \ln\left(\frac{p_2}{p_1}\right) \Rightarrow x_0 \in \Pi_1,$$

$$\frac{(x_0-\overline{x_2})^2 - (x_0-\overline{x_1})^2}{2s^2} < \ln\left(\frac{p_2}{p_1}\right) \Rightarrow x_0 \in \Pi_2,$$

where s^2 is the pooled variance defined by

$$\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}.$$

Here, n_i is the total number of measurements for $i=1,2$. Again, assuming equal cost of misclassification, equal variance and equal prior probabilities produces a rule where

$$\frac{(x_0-\overline{x_2})^2 - (x_0-\overline{x_1})^2}{2s^2} \geqslant 0$$

indicates the measurement was performed on a lesion from Π_1, the pigmented lesion. The point x_0 is the ratio of the 422- and 530-nm photoacoustic amplitudes, and should not be confused with the ratio of the pdfs evaluated at x_0. Note that under equal cost of misclassification and equal prior probabilities, this amounts to a threshold rule that is the midpoint between the sample means.

The performance of our classification scheme can in principal be assessed by calculating the apparent error rate (APER) defined as the fraction of observations in the training sample that are misclassified by the sample. The APER can be easily described by the confusion matrix

	Π_1	Π_2
Π_1	n_{1_c}	$n_{1_m}=n_1-n_{1_c}$,
Π_2	$n_{2_m}=n_2-n_{2_c}$	n_{2_c}

where the columns denote predicted membership and the rows are actual membership. Thus, $\text{APER}=(n_{1_m}+n_{2_m})/(n_1+n_2)$.

As previously discussed, evaluating our classification scheme on the same sample used to develop our rule would result in a bias; that is, it would reduce the APER and make our rule appear to perform better than it does in actuality. To alleviate this bias we evaluate our rule using cross validation (leave-one-out). Specifically, we leave out one observation from the training set, develop our classification rule, and then evaluate the performance on the left out observation. We repeat this process for all the items in the sample and report this figure as our estimated APER.

2.4.2 *Bayesian-model-based classification*

An alternative method for classifying vascular and pigmented lesions is to model the probability of having a vascular lesion as a function of several covariates [i.e., photoacoustic ratio (PR), age, and lesion color]. Following the notation and exposition of Albert,[18] let $y_1, y_2, \ldots, y_n$ denote binary observations where $y_i=1$ if the patient presents a vascular lesion ($y_i=0$ if the patient presents a pigmented lesion). Associated with the i'th response are a set of k covariates $x_{i1}, x_{i2}, \ldots, x_{ik}$. For the model described here $k=2$, corresponding to a model with PR and age as predictors. Although the model we describe has PR and age as predictors, other models can easily be estimated by making appropriate modifications. The model we consider is a probit regression model where the probability that $y_i=1$, p_i, is expressed as

$$p_i = P(y_i=1) = \Phi(\beta_{\text{PR}}\text{PR}_i + \beta_{\text{age}}\text{age}_i), \tag{3}$$

where β_{PR} and β_{age} denote the regression coefficients associated with photoacoustic ratio and age respectively, and $\Phi(\cdot)$ denotes the standard normal cumulative distribution function (cdf).

Due to the small number of subjects, we considered Bayesian estimation so that measures of uncertainty would be more accurately reflected. In particular, we regard this problem as a missing data problem and use the automatic Gibbs sampling method described in Albert and Chib[19] for simulating from the posterior distribution. Specifically, suppose there exists a continuous random variable Z_i associated with lesion such that if Z_i is positive, then the patient has a vascular lesion; otherwise the patient has a pigmented lesion. Further, the continuous lesion measurement (Z_i) is related to the PR and age covariates through the normal linear regression model

$$Z_i = \beta_{\text{PR}}\text{PR}_i + \beta_{\text{age}}\text{age}_i + \epsilon_i,$$

where ϵ_i $(i=1,\ldots,n=15)$ are independent and identically distributed normal random variables with mean 0 and variance 1. Therefore, it follows that

$$P(y_i=1) = P(Z_i>0) = \Phi(\beta_{\text{PR}}\text{PR}_i + \beta_{\text{age}}\text{age}_i).$$

In our analysis, we chose a uniform prior distribution for β (i.e., our prior distribution gives equal weight to all values of

β_{PR} and β_{age} and is thus noninformative). Therefore, the posterior density is given by

$$p(\beta|y) \propto \prod_{i=1}^{n} p_i^{y_i}(1-p_i)^{1-y_i}.$$

As a result, this problem can be considered as a missing data problem where we have a normal linear regression model on the latent data $Z_1,\ldots,Z_n$ and the observed responses are such that they are missing or incomplete in that we only observe $Z_i>0$ $(y_i=1)$ or $Z_i\leqslant 0$ $(y_i=0)$.

Estimation of our model proceeds using an automatic Gibbs sampling algorithm. In particular we employ the algorithm know as "data augmentation" (Albert and Chib[19]), by adding the (unknown) latent data $Z_1,\ldots,Z_n$ to the parameter vector β and sampling from the joint posterior distribution of Z and β. Let $[u|v]$ denote the shorthand notation for the conditional distribution of u given v. Then it can be shown that the conditional posterior distribution of β is

$$[\beta|Z,y] \sim \mathcal{N}\{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Z,(\mathbf{X}'\mathbf{X})^{-1}\},$$

where $\mathbf{X}$ is the design matrix with columns corresponding to the predictors PR and age. Further, given a value of $\beta,Z_1,\ldots,Z_n$ are independent, with

$$[Z_i|\beta,y] \sim \begin{cases} \mathcal{N}(\mathbf{x}_i\beta,1)I(Z_i>0) & \text{if } y_i=1 \\ \mathcal{N}(\mathbf{x}_i\beta,1)I(Z_i<0) & \text{if } y_i=0, \end{cases}$$

where $\mathbf{x}_i$ denotes the vector $(\mathrm{PR}_i,\mathrm{age}_i)'$, and $I(\cdot)$ denotes the indicator function [i.e., $I(Z_i<0)=1$ if $Z_i<0$ and equals 0 otherwise]. Therefore given $\beta=(\beta_{PR},\beta_{age})'$ we can simulate the latent data Z from the truncated normal with the truncation depending on the value of the binary response. See Ref. 18, pp. 216–219, for a comprehensive discussion.

This model can be estimated using the R contributed package "Learn Bayes" (Ref. 20) or via stand alone code. The Markov chain Monte Carlo simulations we conducted for all models considered consisted of 1 chain with 10,000 iterations. The first 5000 iterations were discarded for burn in and convergence was assessed through trace plots of the posterior distributions.

3 Results

3.1 *Gross Examination and Histology*

All lesions were digitally photographed and examined for color, size, and location on the body. They were also classified as pigmented or vascular after gross examination by a dermatologist. Typical pigmented and vascular lesions are shown in Fig. 3. In addition, all lesions were also biopsied and examined histologically to verify whether they were pigmented or vascular, as shown in Fig. 3. The determination by gross examination and histology is provided in Table 1.

3.2 *Photoacoustic Signals*

Typical photoacoustic waveforms for 422 and 530 nm for pigmented and vascular lesions are shown in Fig. 4. The graphs representing photoacoustic generation in pigmented lesions are denoted as Figs. 4(a) and 4(b). In Fig. 4(a), the photoacoustic waveform was generated by irradiating the lesion with 422-nm laser light, while in Fig. 4(b), the wavelength 530 nm was used. The corresponding wavelengths for the vascular lesion were used in the graphs in Figs. 4(c) and 4(d). All photoacoustic waveforms occur at about 3.5 μs, as the separation of the acoustic sensor from the skin surface was approximately 5 mm. In the pigmented lesion, the amplitude of the waves at both wavelengths were similar, with a ratio of 422 to 530 nm about 1.4 to 1. In the vascular lesion, the ratio of peak-to-peak amplitudes is about 4.0 to 1.

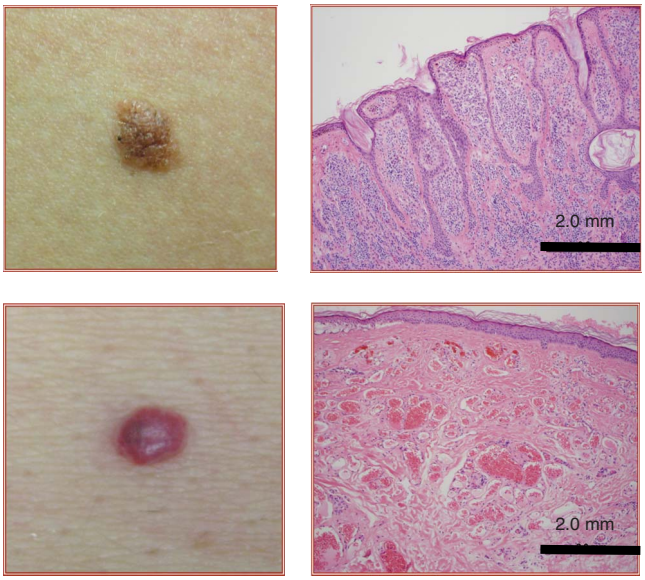

Fig. 3 Photographs and histology of representative (upper) pigmented and (lower) vascular lesions.

3.3 *Classical Classification of Pigmented and Vascular Lesions*

The results of statistical classification of pigmented and vascular lesions are summarized in Table 2. Using the ratios of photoacoustic amplitudes at 422 and 530 nm and the CV scheme, the APER (see Sec. 2.4 for a formal definition) for the standard approach was 6.7%. Using IS classification, the APER is 0.0%.

3.4 *Bayesian Classification of Pigmented and Vascular Lesions*

The Bayesian classification approach used probit regression models, incorporating PR and age as predictors, to differentiate pigmented and vascular lesions. Using PR and age as predictors, the CV method gave an APER of 26.7%, while the IS gave an APER of 20.0%. As a predictor, PR was significant at the 5% level (i.e., the 95% credible interval did not contain zero), whereas age was not. Note that even though age was not a significant predictor in the model (i.e., the 95% credible interval contained zero), the predictor is useful in terms of classification. Finally, using only PR as a predictor, the APER using CV and IS were both 46.7%.

As previously eluded to, the Bayesian approach not only provided classification, but gave information about the certainty of classification, as shown in Fig. 5. This figure shows the estimated probability of classification for each of the 15

Table 1 Photoacoustic determination of vascular and pigmented lesions.

Subject	Size (mm)	Age	Lesion Type	PA Determination	Ratio
1	4	54	vascular	vascular	7
2	3-by-2	35	vascular	vascular	3.1
3	2-by-2	27	pigmented	pigmented	1.63
4	4-by-3	57	vascular	pigmented	0.97
5	3-by-4	56	vascular	vascular	2.31
6	3-by-2	21	pigmented	pigmented	0.5
7	8-by-5	28	pigmented	pigmented	0.8
8	11-by-7	27	pigmented	pigmented	0.7
9	3	37	vascular	vascular	6.7
10	4	43	pigmented	pigmented	1
11	9	69	pigmented	pigmented	1
12	2	57	vascular	vascular	5.3
13	3	57	pigmented	vascular	4.2
14	3	42	vascular	vascular	3.9
15	3	43	vascular	vascular	2.6

Sizes denoted by a single dimension indicate diameter. Note that size was not used as a predictor in our classification.

human subjects. Each histogram represents the distribution of the estimated probability, obtained through Markov chain Monte Carlo, of group membership for each of the 15 subjects. Further, each histogram is denoted in the title by the true nature of the lesion as determined by histology. Our classification rule assigns a subject to the vascular lesion group if the estimated probability (the mean of the posterior distribution of the estimated probabilities obtained from Markov chain Monte Carlo) is greater than one half. Thus, values less than one half indicate pigmented, while values greater than one half indicate vascular. For instance, subjects 1 and 2 are both classified as vascular. However, subject 1 can be definitively classified as vascular (i.e., the 95% credible interval for the estimated probability is strictly greater than 0.5), whereas the classification of subject 2 is not definitive (i.e., the 95% credible interval for the estimated probability contains values less than 0.5)

Finally, an important aspect of conducting classification is the possibility of incorporating the cost of misclassification. The "traditional" Bayesian classification, which is similar to our classical approach (Ref. 21, pp. 304–309) can easily and explicitly account for the cost of misclassification. In our context (using a probit regression model), incorporating the cost of misclassification into our model could be done by changing our decision rule from one that classifies pigmented if the estimated probability is less than 0.5 and vascular if the estimated probability is greater than 0.5 to a decision rule that classifies pigmented if the estimated probability is less than 0.25 and vascular if the estimated probability is greater than 0.25 (or some other user specified threshold).

4 Discussion

Photoacoustic generation and detection was used to depth profile and image pigmented and vascular lesions in human skin.[2,22] We extended these techniques to a method for discriminating these lesions for rapid, noninvasive diagnosis. The advantages of the photoacoustic system are that the optical contrast, assuming the appropriate laser wavelengths are used, is one to two orders of magnitude different between melanin and hemoglobin, the primary chromophores in these lesions and that the generated acoustic wave is robust and relatively immune to the optical scattering nature of skin. The confounding factor is that most lesions are not purely pigmented or vascular.[5,10] Indeed, virtually all vascular lesions will contain a pigmented component from the superficial melanin layer above the network of blood vessels that comprise the lesion itself. Likewise, many pigmented lesions will contain vascular structures, and even those that exist in the avascular epidermis will be in close proximity to the blood supply in the papillary dermis. However, the classification of these lesions relies on the fact that either melanin or hemoglobin dominates the optical nature of the skin. Further, using the robust photoacoustic waves generated in skin, we used two statistical methods to classify lesions in the presence of this dual nature. In particular, we used a standard approach

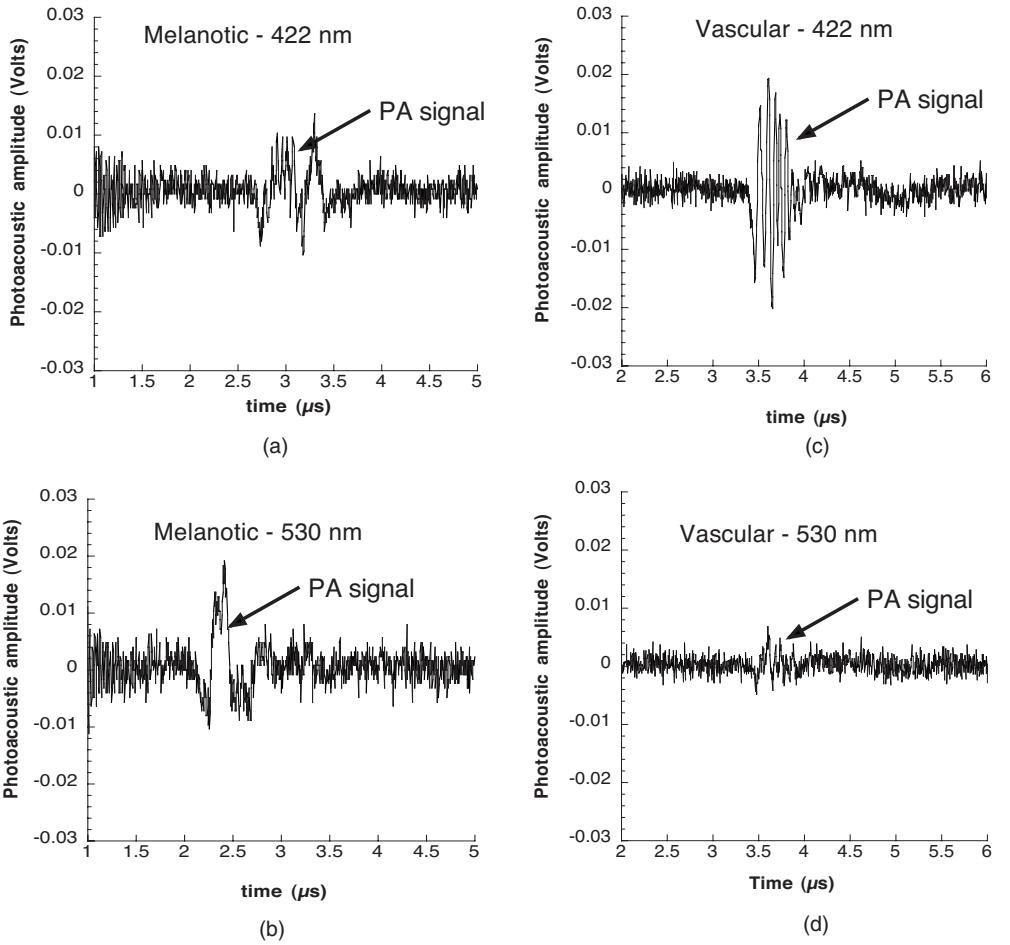

Fig. 4 Representative photoacoustic signals from (left) pigmented and (right) vascular lesions at 422 and 530 nm. Note that these signals were averaged over 128 pulses.

based on multivariate classification techniques as well as a Bayesian probit regression model. Both methods enable classification of lesions using PR predictors. However, it is fundamental to the Bayesian approach to obtain the posterior distribution of the estimated probability of group membership, which can be used as a measure of uncertainty of group membership assignment. This rapid and accurate characterization of skin pigmentation will provide clinicians with valuable information in guiding appropriate triage and therapy of skin lesions *in vivo*.

Table 2 Classification of pigmented and vascular lesions using the standard method and Bayesian probit models.

	APER (%)	
Method	IS	CV
Standard approach	0	6.7
Bayesian model—PR only	46.7	46.7
Bayesian model—PR and age	20	26.7

Dermoscopy has been used for skin diagnosis and its use is becoming increasingly popular. This photoacoustic technique can be used as an adjunct to dermoscopy and gross visualization, as there is still much error in lesion diagnosis. For instance, in a study by Agenziano et al.,[23] 40 clinicians used dermoscopy to diagnose various melanocytic lesions. Using various criteria, the 95% confidence interval of interobserver agreement clustered around 0.5. For example, using the *ABCD* method for melanoma diagnosis, the confidence interval was 0.47 to 0.48. Thus, dermoscopy, while a valuable tool, should not be used solely as a means of diagnosing vascular or pigmented lesions.[6] Furthermore, there are identifying traits of lesions that may be similar between melanocytic and nonmelanocytic lesions, such as streaks and blue pigmentation.[10] Considering the limitations of dermoscopy, photoacoustics may provide additional information for differential diagnosis. The most obvious advantage of photoacoustics being its potential for depth resolution and imaging. While the current work shows the ability of photoacoustics to classify lesion type, additional work may show that lesions can be further classified by depth. In particular, deeper pigmented lesions that appear blue due to Rayleigh scattering may be shown to be distinct from other lesion types, such as a Spitz nevus. One final advantage is that the photoacoustic

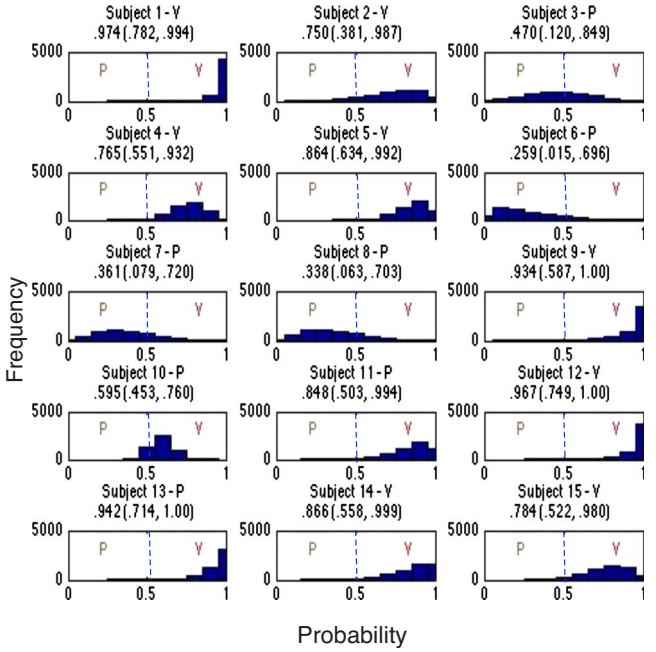

Fig. 5 Montage of histograms for the 15 subjects showing the posterior distribution of the estimated probability of belonging to the vascular lesion group. Subjects with posterior mean greater than 0.5 are classified to the vascular lesion group, whereas the subjects with posterior mean less than 0.5 are classified to the pigmented lesion group. The title of each subject shows a V or P indicating vascular or pigmented according to histological analysis. Additionally, for completeness, the posterior mean and 95% credible interval (in parentheses) are presented.

technique is objective and can be automated; dermoscopy is subjective and diagnosis depends on the skill and experience of the observer.

Using spectral information for hemoglobin and melanin, the ratios of absorption are about 11 to 1 and 2.2 to 1, respectively, using hemoglobin data compiled by Prahl[24] in 1999 and melanin information taken from Jacques et al.[25] The melanin absorption was approximated by the relationship, $\mu_a = 1.70 \times 10^{12} \lambda^{-3.48}$, where μ_a is the absorption coefficient of melanin in inverse centimeters and λ is the wavelength in nanometers. Using this spectral information, there should be two distributions of photoacoustic ratios with means of 11 for hemoglobin and 2.2 for melanin. However, the data shown in this study give different, though still distinct, means for the two pigments. For the vascular lesions, the mean ratio and standard deviation was 4.0 (2.2) and for pigmented lesions is was 1.4 (1.3).

Additional sources of error are evident in Fig. 4 and were not addressed in our analysis. While the photoacoustic signal for vascular lesions at 422 has a signal-to-noise ratio (SNR) of 7 to 1, all other photoacoustic signals, though evident from random noise, are less robust and have SNRs closer to 2 to 1. The signals seen in Fig. 4 are a complex amalgam of noise and acoustic energy integrated from complicated arrangements of chromophores in human skin, in addition to mismatches in acoustic impedance resulting in numerous reflections at skin layer interfaces.

The departure in theory for the expected ratios for the vascular lesions may be explained by the presence of superficial melanin in the epidermis. This blending of photoacoustic signals from hemoglobin and melanin would have depressed the expected ratio. However, the ratio for pigmented lesions was also depressed from a theoretical value of 2.2 to 1 to 1.4 to 1. This difference in ratios is most probably explained by the fact that the melanin absorption model is not proven to be precise, with differences in absorption in melanin by a factor of 2 or more. Nevertheless, the distributions formed by the two populations of ratios from pigmented and vascular lesions were shown to be distinct enough to separate the two groups using the statistical methods described in this paper.

Future developments may include improvements of probe design to increase SNR. We are incorporating a focused PVDF transducer with a acoustic receiver (BR-640A, Ritec, Inc., Warwick, Rhode Island) to enable us to decrease laser fluence to within ANSI standards. The original handpiece used in this study was made from PVDF film without focusing, coupled directly to a linear amplifier, then sent to an oscilloscope. This receiver enables inline filtering within the appropriate range of frequencies expected from photoacoustic waves in tissue, typically 1 to 50 MHz. The receiver has significantly higher gain, up to 64 dB. Using a stabilized probe mounted on an articulating arm will enable higher signal averaging. Using these improvements, signal acquisition can be improved along with signal processing.

As previously discussed, if one is careful about the ultrasound considerations, it is possible to decrease the fluence and increase the SNR. Thus, improving the ultrasound detector, including improved focusing and sensitivity, could have a significant impact on the effectiveness of our classification. In particular, increasing the SNR would reduce the uncertainty in the estimated distribution of the PR for both populations (i.e., the vascular and pigmented populations). As a result, the 95% credible intervals (and confidence intervals) for the estimated PRs will be narrower. Thus, it will be more likely for the estimated distributions of the two populations to be disjoint. Ultimately, this would culminate in superior discriminating power in both the classical and Bayesian approaches.

The standard classification approach used in our analysis was based on the assumption that the two populations follow a Gaussian distribution. Although this assumption may be violated our empirical results suggest that the classification method is fairly robust to these departures. Furthermore, note that the results of both types of analysis are based on a small number of subjects. Therefore, the results presented here should be interpreted in the context of a pilot study, and future work will necessarily include a large-scale clinical study to develop classification rules and model parameter estimates suitable for use in practice.

Acknowledgments

This work was supported by the Departments of Dermatology and Biological Engineering at the University of Missouri and the National Institutes of Health (NIH) Clinical Loan Repayment Program and NIH R21 CA 139186–01. We acknowledge the resources of the Christopher S. Bond Life Sciences Center where the photoacoustic experiments were performed. We thank Jon Dyer, MD, of the Department of Dermatology

for valuable suggestions. We also acknowledge Louanne Chance for her help in organizing the patient studies.

References

1. R. P. Braun, R. Baran, F. A. Le Gal, S. Dalle, S. Ronger, R. Pandolfi, O. Gaide, L. E. French, P. Laugier, J. H. Saurat, A. A. Marghoob, and L. Thomas, "Diagnosis and management of nail pigmentations," *J. Am. Acad. Dermatol.* **56**, 835–847 (May 2007).
2. J. A. Viator, G. Au, G. Paltauf, S. L. Jacques, S. A. Prahl, H. Ren, Z. Chen, and J. S. Nelson, "Clinical testing of a photoacoustic probe for port wine stain depth determination," *Lasers Surg. Med.* **30**, 141–148 (2002).
3. J. A. Viator, B. Choi, M. Ambrose, J. Spanier, and J. S. Nelson, "In vivo port-wine stain depth determination with a photoacoustic probe," *Appl. Opt.* **42**, 3215–3224 (June 2003).
4. J. A. Viator, J. Komadina, L. O. Svaasand, G. Aguilar, B. Choi, and J. Stuart Nelson, "A comparative study of photoacoustic and reflectance methods for determination of epidermal melanin content," *J. Invest. Dermatol.* **122**, 1432–1439 (June 2004).
5. G. Argenziano, I. Zalaudek, R. Corona, F. Sera, L. Cicale, G. Petrillo, E. Ruocco, R. Hofmann-Wellenhof, and H. P. Soyer, "Vascular structures in skin tumors: a dermoscopy study," *Arch. Dermatol.* **140**, 1485–1489 (Dec. 2004).
6. G. Argenziano, I. Zalaudek, G. Ferrara, R. Johr, D. Langford, S. Puig, H. P. Soyer, and J. Malvehy, "Dermoscopy features of melanoma incognito: indications for biopsy," *J. Am. Acad. Dermatol.* **56**, 508–513 (Mar. 2007).
7. I. Zalaudek, G. Argenziano, A. Di Stefani, G. Ferrara, A. A. Marghoob, R. Hofmann-Wellenhof, H. P. Soyer, R. Braun, and H. Kerl, "Dermoscopy in general dermatology," *Dermatology (Basel, Switz.)* **212**, 7–18 (2006).
8. E. E. Ducharme and N. B. Silverberg, "Selected applications of technology in the pediatric dermatology office," *Semin Cutan Med. Surg.* **27**, 94–100 (Mar. 2008).
9. V. de Giorgi, S. Sestini, D. Massi, V. Maio, and B. Giannotti, "Dermoscopy for 'true' amelanotic melanoma: a clinical dermoscopic-pathologic case study," *J. Am. Acad. Dermatol.* **54**, 341–344 (Feb. 2006).
10. A. Scope, C. Benvenuto-Andrade, A. L. Agero, and A. A. Marghoob, "Nonmelanocytic lesions defying the two-step dermoscopy algorithm," *Dermatol. Surg.* **32**, 1398–1406 (Nov. 2006).
11. D. Piccolo, A. Ferrari, K. Peris, R. Diadone, B. Ruggeri, and S. Chimenti, "Dermoscopic diagnosis by a trained clinician vs. a clinician with minimal dermoscopy training vs. computer-aided diagnosis of 341 pigmented skin lesions: a comparative study," *Br. J. Dermatol.* **147**, 481–486 (Sep. 2002).
12. K. Terstappen, O. Larkö, and A. M. Wennberg, "Pigmented basal cell carcinoma-comparing the diagnostic methods of SIAscopy and dermoscopy," *Acta Derm Venereol* **87**, 238–242 (2007).
13. D. S. Gareau, J. Lagowski, V. M. Rossi, J. A. Viator, G. Merlino, M. Kulesz-Martin, and S. L. Jacques, "Imaging melanoma in a murine model using reflectance-mode confocal scanning laser microscopy and polarized light imaging," *J. Investig. Dermatol. Symp. Proc.* **10**, 164–169 (Nov. 2005).
14. H. Kang, B. Jung, and J. S. Nelson, "Polarization color imaging system for on-line quantitative evaluation of facial skin lesions," *Dermatol. Surg.* **33**, 1350–1356 (Nov. 2007).
15. A. Rosencwaig, *Photoacoustics and Photoacoustic Spectroscopy*, Wiley, New York (1980).
16. R. Johnson and D. Wichern, *Applied Multivariate Statistical Analysis*, Prentice-Hall, Upper Saddle River, NJ (1998).
17. R. J. Talbert, S. H. Holan, and J. A. Viator, "Photoacoustic discrimination of viable and thermally coagulated blood using a two-wavelength method for burn injury monitoring," *Phys. Med. Biol.* **52**, 1815–1829 (Apr. 2007).
18. J. Albert, *Bayesian Computation with R*, Springer, Seattle, WA (2007).
19. J. Albert and S. Chib, "Bayesian analysis of binary and polychotomous response data," *J. Am. Stat. Assoc.* **88**, 669–679 (1993).
20. J. Albert, *Learn Bayes Package*, R Foundation for Statistical Computing, Vienna, Austria (2008).
21. K. Mardia, J. Kent, J. Bibby, *Multivariate Analysis*, Academic, London (1979).
22. R. G. Kolkman, M. J. Mulder, C. P. Glade, W. Steenbergen, and T. G. van Leeuwen, "Photoacoustic imaging of port-wine stains," *Lasers Surg. Med.* **40**, 178–182 (Mar. 2008).
23. G. Argenziano, H. P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, M. Binder, L. Cerroni, G. De Rosa, G. Ferrara, R. Hofmann-Wellenhof, M. Landthaler, S. W. Menzies, H. Pehamberger, D. Piccolo, H. S. Rabinovitz, R. Schiffner, S. Staibano, W. Stolz, I. Bartenjev, A. Blum, R. Braun, H. Cabo, P. Carli, V. De Giorgi, M. G. Fleming, J. M. Grichnik, C. M. Grin, A. C. Halpern, R. Johr, B. Katz, R. O. Kenet, H. Kittler, J. Kreusch, J. Malvehy, G. Mazzocchetti, M. Oliviero, F. Ozdemir, K. Peris, R. Perotti, A. Perusquia, M. A. Pizzichetta, S. Puig, B. Rao, P. Rubegni, T. Saida, M. Scalvenzi, S. Seidenari, I. Stanganelli, M. Tanaka, K. Westerhoff, I. H. Wolf, O. Braun-Falco, H. Kerl, T. Nishikawa, K. Wolff, and A. W. Kopf, "Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet," *J. Am. Acad. Dermatol.* **48**, 679–693 (May 2003).
24. S. A. Prahl "Optical absorption of hemoglobin" (1999), http://omlc.ogi.edu/spectra/hemoglobin/index.html.
25. S. L. Jacques and D. J. McAuliffe, "The melanosome: threshold temperature for explosive vaporization and internal absorption coefficient during pulsed laser irradiation," *Photochem. Photobiol.* **53**, 769–775 (June 1991).