Model-based spectroscopic analysis of the oral cavity: impact of anatomy

Sasha McGee
Jelena Mirkovic
Massachusetts Institute of Technology
G.R. Harrison Spectroscopy Laboratory
Building 6-205, 77 Massachusetts Avenue
Cambridge, Massachusetts 02139

Vartan Mardirossian
Alphi Elackattu
Boston Medical Center
4th Floor, FGH Building
820 Harrison Avenue
Boston, Massachusetts 02118

Chung-Chieh Yu
Massachusetts Institute of Technology
G.R. Harrison Spectroscopy Laboratory
Building 6-205, 77 Massachusetts Avenue
Cambridge, Massachusetts 02139

Sadru Kabani
George Gallagher
Boston University Goldman School of Dental Medicine
Department of Oral and Maxillofacial Pathology
100 East Newton Street G-04
Boston, Massachusetts 02118

Robert Pistey
Boston Medical Center
Department of Anatomic Pathology
670 Albany Street, 3rd floor (BIO-3)
Boston, Massachusetts 02118

Luis Galindo
Massachusetts Institute of Technology
G.R. Harrison Spectroscopy Laboratory
Building 6-205, 77 Massachusetts Avenue
Cambridge, Massachusetts 02139

Kamran Badizadegan
Harvard Medical School and Massachusetts General
    Hospital
Department of Pathology
55 Fruit Street, WRN219
Boston, Massachusetts 02114

Zimmern Wang
Boston Medical Center
4th Floor, FGH Building
820 Harrison Avenue
Boston, Massachusetts 02118

Ramachandra Dasari
Michael S. Feld
Massachusetts Institute of Technology
G.R. Harrison Spectroscopy Laboratory
Building 6-205, 77 Massachusetts Avenue
Cambridge, Massachusetts 02139

Address all correspondence to Sasha McGee, G.R. Harrison Spectroscopy Laboratory, Massachusetts Institute of Technology, 77 Massachusetts Ave. Bldg. 6-218M-Cambridge, Massachusetts 02139; Tel: 617-258-7667; Fax: 617-253-4513; E-mail: samcgee@mit.edu

Gregory Grillone
Boston Medical Center
4th Floor, FGH Building
820 Harrison Avenue
Boston, Massachusetts 02118

Abstract. In order to evaluate the impact of anatomy on the spectral properties of oral tissue, we used reflectance and fluorescence spectroscopy to characterize nine different anatomic sites. All spectra were collected *in vivo* from healthy oral mucosa. We analyzed 710 spectra collected from the oral cavity of 79 healthy volunteers. From the spectra, we extracted spectral parameters related to the morphological and biochemical properties of the tissue. The parameter distributions for the nine sites were compared, and we also related the parameters to the physical properties of the tissue site. *k*-Means cluster analysis was performed to identify sites or groups of sites that showed similar or distinct spectral properties. For the majority of the spectral parameters, certain sites or groups of sites exhibited distinct parameter distributions. Sites that are normally keratinized, most notably the hard palate and gingiva, were distinct from nonkeratinized sites for a number of parameters and frequently clustered together. The considerable degree of spectral contrast (differences in the spectral properties) between anatomic sites was also demonstrated by successfully discriminating between several pairs of sites using only two spectral parameters. We tested whether the 95% confidence interval for the distribution for each parameter, extracted from a subset of the tissue data could correctly characterize a second set of validation data. Excellent classification accuracy was demonstrated. Our results reveal that intrinsic differences in the anatomy of the oral cavity produce significant spectral contrasts between various sites, as reflected in the extracted spectral parameters. This work provides an important foundation for guiding the development of spectroscopic-based diagnostic algorithms for oral cancer. © *2008 Society of Photo-Optical Instrumentation Engineers.* [DOI: 10.1117/1.2992139]

1 Introduction

Reflectance and fluorescence spectroscopy have been widely investigated as noninvasive tools for detecting oral cancer.[1–5] Although numerous studies have reported differences in spectral parameters when comparing diseased and normal tissues, in many cases, multiple anatomic sites (i.e., the gingiva, buccal mucosa, etc.) are combined without taking into account the normal spectral contrasts (differences in the spectral

properties).[5–8] Therefore, spectral contrasts due to anatomy may be incorrectly attributed to malignancy. A critical step in applying these techniques to the detection of oral malignancy is to first develop an understanding of the spectral contrasts among anatomic sites in the oral cavity.

Only a few studies have investigated the relationship between spectral contrasts and anatomic differences. In a study by Kolli et al., fluorescence spectra were collected from the buccal mucosa and dorsal surface of the tongue of 21 healthy volunteers, in addition to patients.[9] Analysis of data collected from healthy volunteers revealed that the ratio of fluorescence emission at 450 nm when excited at 335 and 375 nm, respectively, was statistically different between the buccal mucosa and dorsal surface of the tongue. Differences in the fluorescence peak intensities at 337, 365, and 410 nm excitation were also noted by Gillenwater et al. in a study of nine sites in the oral cavity of eight healthy volunteers.[10] Another small study of 10 healthy volunteers by Dhingra et al. noted differences in the intensity of the fluorescence emission among various sites.[11] Although these three studies provide support for the fact that spectral contrasts due to anatomy exist, they lack a detailed description of the differences and only examine a few sites.

In a larger study by de Veld et al., 9295 fluorescence spectra were collected from 97 healthy volunteers to determine whether differences in the mucosa at various sites in the oral cavity resulted in changes in the fluorescence intensity and lineshape.[12] In this study, both classification and clustering methods were used to detect spectral contrasts among the 13 anatomic sites evaluated. The authors found that the dorsal surface of the tongue and vermilion border of the lip each displayed unique spectral properties, and the other anatomic sites were similar enough to be combined into a single group; however, significant differences in total fluorescence intensity between almost all the different sites were noted at multiple wavelengths. In this large study, the authors tested the class overlap between each of the 11 anatomic sites (the dorsal surface of the tongue and vermilion border of the lip were excluded) against the remaining 10 sites. Because the larger group consisted of multiple sets of sites with small differences in their means and standard deviations, the resultant distribution has a large spread and is more likely to overlap with the single site against which it is compared. This would diminish the ability to observe differences. There may be added value in separating certain sites, because small changes with cancer would be more easily distinguished for narrow distributions (i.e., increased power of a diagnostic test).

None of the four studies cited above used a physical model of tissue fluorescence, in which the spectral features are modeled by parameters relating to specific tissue components. Therefore, it is difficult to directly interpret or provide an explanation for the observed differences in fluorescence emission. Additionally, to the best of our knowledge, no studies have been performed that examine differences in the reflectance properties of oral tissue due to anatomy.

In the present study, we use a physical model to extract parameters from both diffuse reflectance and intrinsic fluorescence spectra that are related to the morphological and biochemical properties of the tissue.[13] We compare the extracted parameter distributions for various anatomic sites, and also relate the physical spectral parameters to the anatomic features of the tissue sites. We perform k-means cluster analysis in order to identify and classify groups of sites that share similar or distinct spectral properties.

2 Materials and Methods

2.1 In Vivo Data Collection

Healthy volunteers (HVs) were recruited at Boston Medical Center (BMC) and at the Massachusetts Institute of Technology (MIT). The protocol for *in vivo* data collection was approved by the institutional review board at BMC and the Committee on the Use of Humans as Experimental Subjects at MIT. Written informed consent was obtained from all subjects to indicate their willingness to participate in the study. Relevant background information was obtained from each subject, such as their smoking history, alcohol consumption, and history of lesions in the upper aerodigestive tract. Study subjects without a prior history of a lesion of the oral cavity, larynx, or esophagus (benign or malignant) were considered healthy volunteers, regardless of their smoking status. Smokers were included as HVs because in our patient study we have found that those presenting with oral lesions represent a mixture of smokers and nonsmokers. Less than 1% of the subjects had a history of alcohol abuse.

Reflectance and fluorescence spectra were collected from nine different anatomic sites within the oral cavity using the Fast Excitation Emission Matrix (FastEEM) instrument, which has been previously described.[14] The calibration procedures are also described in Ref 14. Briefly, a xenon arc flash lamp (2.9-μs pulse) and a 308-nm XeCl excimer laser (8-ns pulse) serve as the excitation light sources for the reflectance and fluorescence measurements, respectively. Fluorescence excitation wavelengths between 308 and 460 nm were generated by using the excimer laser to sequentially pump a series of dyes. A 1.3-mm diameter optical fiber probe, consisting of a central delivery fiber (200-μm core) and six collection fibers (200-μm core), was used to deliver the excitation light and collect the fluorescence emission and diffuse reflectance. A 1-mm quartz shield at the tip of the probe ensured a fixed geometry when the probe was placed in contact with the tissue during data collection. Emission light was collected over the range of 300–800 nm. The optical fiber probe was disinfected with CIDEX OPA (Advanced Sterilization Products, Irvine, CA) before each session, according to the manufacturer's specifications. In a single data set, five measurements each of the reflectance and fluorescence spectra were collected in ~1.5 s, and the average of each of the five spectra was recorded. The standard deviation for the five reflectance and fluorescence spectra, respectively, was also recorded. Approximately three to five data sets were acquired for each tissue sample examined. Data sets exhibiting large standard deviations were excluded from the analysis. Additionally, data sets that did not overlap in lineshape or intensity were excluded, and the remaining data sets were averaged. If all the measurements were inconsistent, then the data for the tissue sample were excluded from the analysis.

2.2 Model-Based Analysis

The measured diffuse reflectance was fit using the model of Zonios et al.[15] The inputs to the model are the reduced scat-

tering coefficient, $\mu_s'(\lambda)$, and absorption coefficient, $\mu_a(\lambda)$. Reflectance spectra were fit over the range of 380–700 nm using a constrained nonlinear least squares fitting algorithm. A double power law equation was used to represent the wavelength dependence of μ_s',

$$\mu_s'(\lambda) = A\left(\frac{\lambda}{\lambda_0}\right)^{-B} + C\left(\frac{\lambda}{\lambda_0}\right)^{-4} \quad [\text{mm}^{-1}]. \qquad (1)$$

Here, the wavelength, λ, is expressed in units of microns and λ_0 is equal to 1 μm. A power law description of μ_s' has been widely used to model data collected from both cells and tissue.[16–19] We add the second power law term, $C\lambda^{-4}$, because it was found to be particularly important for fitting the spectra at wavelengths <400 nm, where there is significant scattering from smaller particles (Rayleigh scattering). From modeling the scattering, we extract three parameters: A, B, and C. Parameter A is a scaling parameter related to the overall magnitude of scattering. In a study of skin, the collagen fibers of the dermis were shown to be the major source of light scattering, contributing mostly to the scattering from larger particles (Mie scattering) and, to a lesser extent, Rayleigh scattering.[20] Parameter B reflects the size of the scattering particles.[16] C represents the magnitude of scattering by small scatterers.

The absorption coefficient was modeled as the sum of the contributions from two absorbers, hemoglobin and β-carotene,

$$\mu_a(\lambda) = \mu_a^{\text{Hb}}(\lambda) + \mu_a^{\beta C}(\lambda) \quad [\text{mm}^{-1}]. \qquad (2)$$

Hemoglobin absorption was modeled using a correction for vessel packaging previously described in the literature.[21–23] In our analysis, we adopt the formulation of the model developed by Svaasand et al.,[23] as presented by van Veen et al.[21] Vessel packaging accounts for alterations in the shape and intensity of the hemoglobin absorption peaks as a result of the strong absorption of light at 420 nm by hemoglobin in blood vessels. Using this model, the absorption coefficient of hemoglobin is represented by the product of the absorption coefficient of hemoglobin in whole blood, μ_a^{blood}, a wavelength-dependent correction factor that alters the shape of the absorption spectrum, K, and the volume fraction of blood sampled, ν. The absorption coefficient for hemoglobin was described as follows:

$$\mu_a^{\text{Hb}}(\lambda) = K(\lambda)\nu\mu_a^{\text{blood}}(\lambda) \quad [\text{mm}^{-1}]. \qquad (3)$$

The correction factor K is given by

$$K(\lambda) = \left\{\frac{1 - \exp[-2\mu_a^{\text{blood}}(\lambda)R]}{2R\mu_a^{\text{blood}}(\lambda)}\right\}, \qquad (4)$$

where R is the effective vessel radius (in millimeters) and $\mu_a^{\text{blood}}(\lambda)$ is the absorption coefficient of blood (in inverse millimeters). The volume fraction, υ, was calculated using the following ratio:

$$\nu = \text{cHb}/150, \qquad (5)$$

where cHb is the extracted concentration of hemoglobin (in milligrams per milliliter). The concentration of hemoglobin in

whole blood was assigned a value of 150 mg/mL. The equation for μ_a^{blood} was given by

$$\mu_a^{\text{blood}}(\lambda) = 0.1 \ln(10) \times 150[\alpha\varepsilon_{\text{HbO2}}(\lambda) + (1-\alpha)\varepsilon_{\text{Hb}}(\lambda)] \quad [\text{mm}^{-1}], \qquad (6)$$

where α is the oxygen saturation, ε is the extinction coefficient (in $1/[\text{cm}(\text{mg}/\text{mL})]$), HbO_2 denotes oxyhemoglobin, and Hb denotes deoxyhemoglobin.[24] We further impose a lower limit on the extracted effective vessel radius of 2.5 μm, because the minimum diameter of a capillary is on the order of 5–7 μm.[25] The absorption coefficient for β-carotene was modeled as follows:

$$\mu_a^{\beta C}(\lambda) = 0.1 \ln(10)c\beta C\varepsilon_{\beta\text{Car}}(\lambda) \quad [\text{mm}^{-1}], \qquad (7)$$

where $c\beta C$ represents the concentration of β-carotene (in milligrams per milliliter) and $\varepsilon_{\beta\text{Car}}$ is the extinction coefficient β-carotene (in $1/[\text{cm}(\text{mg}/\text{mL})]$).[26] The absorption coefficient was calculated as the sum of the contributions of each absorber,

$$\mu_a(\lambda) = K(\lambda)\nu\mu_a^{\text{blood}}(\lambda) + 0.1 \ln(10)c\beta C\varepsilon_{\beta\text{Car}}(\lambda) \quad [\text{mm}^{-1}]. \qquad (8)$$

From the reflectance spectra we derive seven parameters: A, B, C, cHb, the oxygen saturation (α), effective vessel radius, and cβC.

Using a model that has been previously described, the distortions in the measured fluorescence spectra due to scattering and absorption were removed, in order to obtain what we refer to as the instrinsic fluorescence.[13,27] Unlike the measured fluorescence, the intrinsic fluorescence can be modeled as a linear combination of the component fluorophores present within the tissue. The spectra of the component fluorophores (basis spectra) for each excitation wavelength were extracted from the intrinsic fluorescence using multivariate curve resolution (MCR) and, in some cases, a spectrum collected from the pure chemical.

Parameters were extracted from the intrinsic fluorescence spectra in the subsequent data analysis in one of two ways, which we refer to as non-normalized fluorescence data or area-normalized fluorescence data. In the case of the non-normalized fluorescence data, the basis spectrum of each fluorophore at a given excitation wavelength was normalized by dividing the intensity at each wavelength point in the spectrum by the intensity at the peak. Following this step, a linear combination of the basis spectra was used to fit the intrinsic fluorescence spectrum for the sample. For area-normalized fluorescence data, the intensity at each wavelength in the basis spectrum was divided by the total area under the curve. This same process was applied to the intrinsic fluorescence spectrum for the sample. The resultant intrinsic fluorescence spectrum for the sample was then fit with a linear combination of the basis spectra. In a final step, the extracted contribution of each fluorophore was divided by the total sum of the contributions for all the fluorophores at the specific excitation wavelength. The extracted contribution for each fluorophore in the area-normalized fluorescence data represents the fraction of the total area under the emission curve contributed by the component fluorophore.

The fluorescence excited at 308 nm was modeled by a linear combination of tryptophan, collagen, and the reduced form of nicotinamide adenine dinucleotide (NADH) over the range of 370–600 nm. The fluorescence excited at 340 nm was modeled by a linear combination of NADH and collagen over the range of 380–600 nm.

2.3 Outlier Removal and k-Means Cluster Analysis

Outliers for each of the parameters were identified for each site separately based on the interquartile range (IQR) exclusion criteria.[28] Applying these criteria, parameter values were excluded if they exceeded the upper quartile by 1.5 times the IQR, or fell below the lower quartile by 1.5 times the IQR. Because each parameter was evaluated separately, the number of exclusions varied for each parameter. Following this procedure, we performed k-means cluster analysis to determine if there were specific sites with unique spectral properties, without assuming a priori that the nine anatomic sites were well-defined groups (unsupervised classification). The analysis was performed using MATLAB (Mathworks, Natick, MA). Outliers in the data were removed in order to prevent skewing of the cluster centroid and to maximize the accuracy of the results. Analysis was performed for each parameter separately for $k=2$, 3, and 4 clusters using a random initial clustering assignment and the city-block distance measure. Each separation entailed 20 replicate runs of the clustering algorithm. For each iteration, the distance from all points to its assigned centroid was calculated and the final result of the function represented the iteration that produced the minimum sum of all the distances. For each k-means separation (i.e. $k=2$, 3, or 4), the percentage of each site within each of the clusters was identified after subtracting the probability of a site being assigned to a cluster purely due to chance (i.e., 25% for $k=4$). We define Δ as the difference between the maximum percentage of a given site assigned to a cluster and the percentage that would be expected to be assigned due to chance. All sites for which the value for Δ for all three separations or at least two separations exceeded 30.0% were identified for each parameter. These selection criteria for Δ were used to identify sites displaying significant clustering. For each parameter, sites meeting these criteria were either clustered together or separately based on the whether they shared the same or unique cluster assignments.

In order to evaluate the degree of spectral contrast, a logistic regression model was developed to differentiate specific sites. A receiver operator characteristic (ROC) curve was then generated after performing leave-one-out cross-validation. The accuracy of the separation was evaluated by determining the sensitivity, specificity, and area under the ROC curve (AUC). The sensitivities and specificities were selected based on the Youden index, which represents the point in the ROC curve with the maximum vertical distance from the 45 degree line.[29]

3 Results

3.1 General Description of the Data

The complete set of HV data consisted of 781 spectra. Because of instrumentation error, evidence of movement (large variations among the three to five data sets collected from

Table 1 Summary of the number of spectra collected from each of the nine sites in the oral cavity for the calibration, validation, and total data sets. BM: buccal mucosa, DT: dorsal surface of the tongue, FM: floor of the mouth, GI: gingiva, HP: hard palate, LT: lateral surface of the tongue, RT: retromolar trigone, SP: soft palate, and VT: ventral surface of the tongue.

Site	Calibration data	Validation data	Total data
	Number of spectra		
BM	88	12	100
DT	110	14	124
FM	98	14	112
GI	38	20	58
HP	32	22	54
LT	101	11	112
RT	24	9	33
SP	30	13	43
VT	63	11	74
Total	584	126	710

each tissue sample), or too few spectra collected from a given anatomic site, 9.1% of the spectra were excluded. The final set of healthy volunteer data included 710 spectra from 79 subjects. The average age (±standard deviation) of the subjects was 41.5 ± 16 years. Data were collected from the following nine sites in the oral cavity: buccal mucosa (BM), dorsal surface of the tongue (DT), floor of the mouth (FM), gingiva (GI), hard palate (HP), lateral surface of the tongue (LT), retromolar trigone (RT), soft palate (SP), and ventral surface of the tongue (VT). Table 1 summarizes the number of spectra collected from each site within the oral cavity for the entire data set. Fewer spectra were collected from sites such as the HP, RT, and SP because the architecture of the oral cavity in some subjects made these areas less accessible to the probe.

3.2 Reflectance and Fluorescence Modeling

Figure 1 shows representative examples of the excellent quality of the fits to the reflectance and intrinsic fluorescence spectra, respectively. Examples of how the quality of the modeled fit can deteriorate when either β-carotene or vessel packaging are not accounted for are shown in Figs. 1(a) and 1(b), respectively. β-carotene has a prominent absorption peak at ~450 nm and a secondary peak at 478 nm. The latter absorption peak is especially prominent in the spectrum in Fig. 1(a). The vessel packaging model predicts that the differences in the depth of the 420-nm hemoglobin peak as compared to the 540- and 580-nm peaks will diminish as the size of the effective vessel radius increases. The uncorrected hemoglobin absorption spectrum underfits the 540- and 580-nm peaks. In Figs. 1(c) and 1(d), the fluorescence emission at 308- and 340-nm excitation, respectively, is shown. The fluorescence
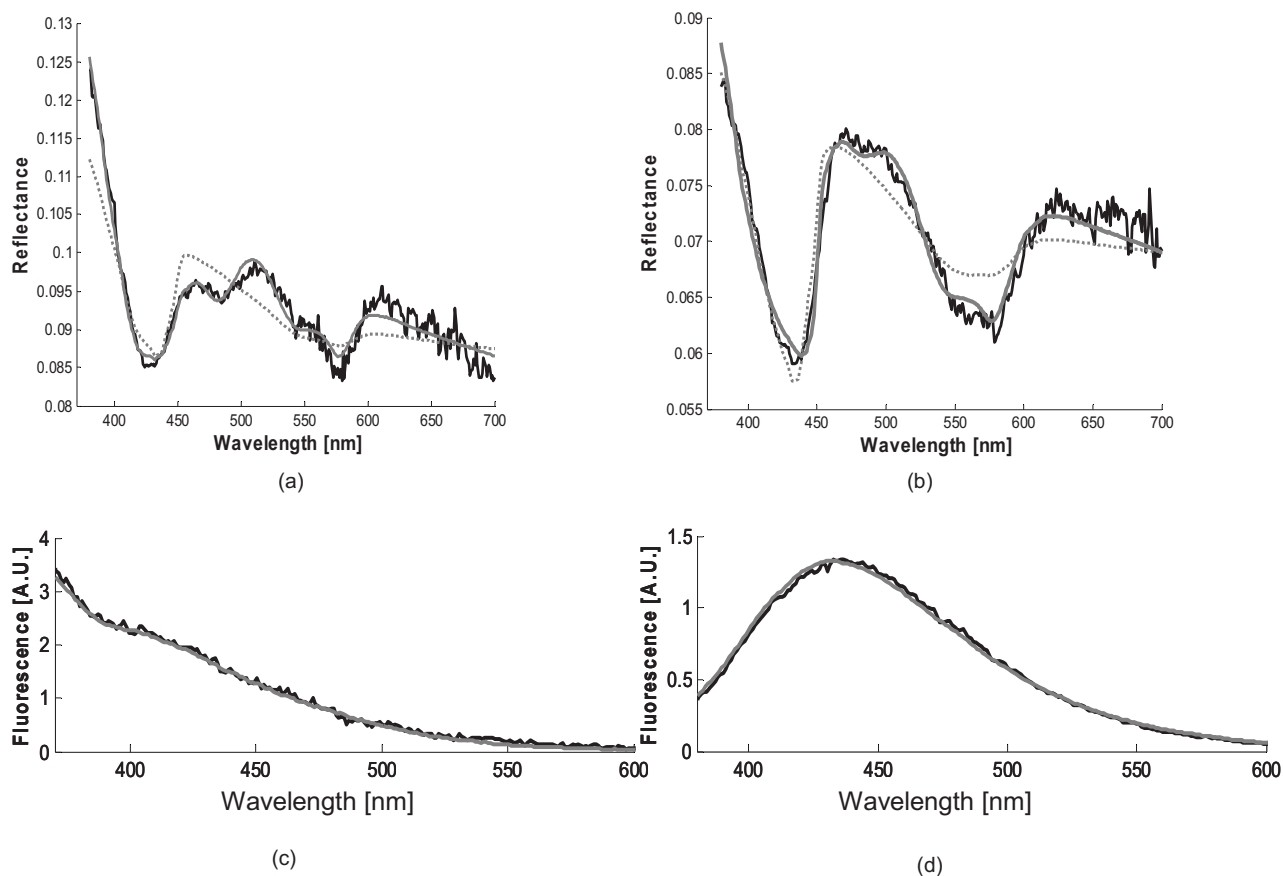
Fig. 1 (a) Plot of a reflectance spectrum measured from the buccal mucosa (thick black line) and the modeled fit excluding β-carotene (dashed gray line). The solid gray line demonstrates the fit when β-carotene is included as an absorber. (b) Plot of a reflectance spectrum measured from the lateral surface of the tongue (thick black line). The gray lines show the modeled fit excluding the effect of vessel packaging (dashed gray line) and including vessel packaging (solid gray line). (c) Plot of the intrinsic fluorescence at 308-nm excitation (black line) and the modeled fit (gray line) for a DT site. (d) Plot of the intrinsic fluorescence at 340-nm excitation (black line) and the modeled fit (gray line) for a HP site.

excited at 308 nm was modeled by a linear combination of tryptophan, collagen, and NADH [Fig. 1(c)]. The fluorescence excited at 340 nm was modeled by a linear combination of three fluorophores: NADH, and two additional fluorophores with peaks at approximately 401 and 427 nm, respectively [Fig. 1(d)]. A preliminary examination of the data revealed that by using only NADH and a single additional fluorophore (collagen), we could not reliably model the intrinsic fluorescence signal at 340-nm excitation for the entire data set. Figure 2(a) shows a plot of NADH and the two additional spectral components used to fit the spectra at 340-nm excitation. We refer to the two extracted collagen components as Coll401 (340 nm) and Coll427 (340 nm), where the number represents the wavelength at which the peak emission occurs. To determine if the two fluorophores we extracted could be due to different types of collagen, we collected the fluorescence emission for dry collagen IV (Sigma-Aldrich C9879) and collagen I (Sigma-Aldrich C5533), the major forms present in the basement membrane and stroma, respectively.[30,31] Figure 2(b) shows the fluorescence emission excited at 340 nm. The peaks for collagen I and collagen IV occur at 409 and 422 nm, respectively, which is similar to what we obtained from the tissue data.

The following six fluorescence ratios were examined based on the analysis of the non-normalized fluorescence data:

NADH/collagen (308 nm), NADH/tryptophan (308 nm), NADH/Coll401 (340 nm), NADH/Coll427 (340 nm), NADH/total collagen (340 nm), and Coll401/Coll427 (340 nm). The number in parentheses indicates the excitation wavelength. Total collagen refers to the sum of Coll401 and Coll427. Analysis of the area-normalized fluorescence data yielded six additional parameters: tryptophan (308 nm), collagen (308 nm), NADH (308 nm), NADH (340 nm), Coll401 (340 nm), and Coll427 (340 nm). The 308- and 340-nm excitations were specifically chosen for our analysis because data were not consistently available for all samples at the higher wavelengths due to a low signal and operator error.

3.3 Model-Based Analysis Parameter Distributions

Table 2 summarizes the mean, standard deviation (σ), and the relative standard deviation (RSD) ($[\sigma/\text{mean}] \times 100$) for each parameter for each individual site, as well as when all nine sites were combined into a single group. In each case, the statistical measures were calculated after applying the IQR exclusion criteria. The relative standard deviation varied widely across sites for a given parameter, even among sites with similar means. For 16 of the 19 parameters, the RSD for the combined group of sites exhibited a value that ranked among the highest three values.
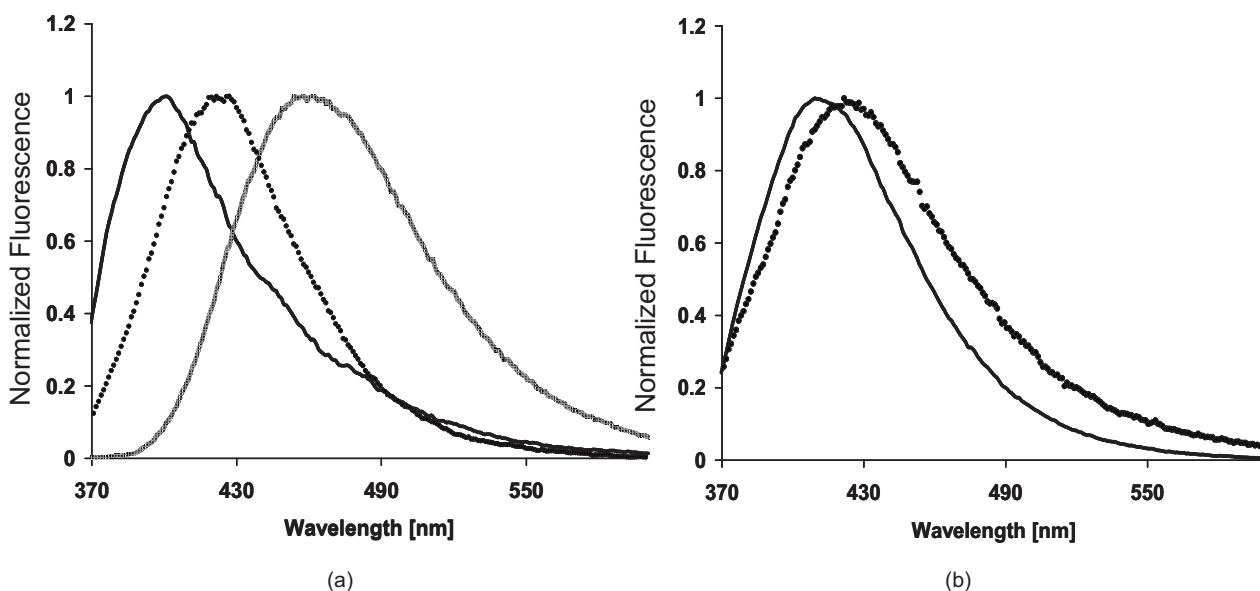
Fig. 2 (a) Plot of the two fluorescence spectra extracted at 340-nm excitation using multivariate curve resolution, Coll401 (solid black line) and Coll427 (dotted black line), and NADH (shaded gray line). The fluorescence peaks occur at approximately 401 and 427 nm, respectively. These three components were used to model the intrinsic fluorescence spectra at 340-nm excitation. (b) Plot of the fluorescence emission from collagen I (solid line) and collagen IV (dotted line) chemicals at 340-nm excitation. The fluorescence peaks occur at approximately 409 and 422 nm, respectively.

For parameter A, most of the sites showed very similar mean values, except the BM and RT, which exhibited a mean that was significantly higher than all other sites based on a multiple comparison test ($p < 0.05$) but not from each other. The HP and DT demonstrated slightly lower mean values that were significantly different from all other sites, except the GI.

The GI and HP both exhibited values of B of 0.6, which was significantly higher than for other sites. Three groups were easily distinguished for parameter C. The FM, LT, and VT formed one group, demonstrating the highest values of C. The BM, DT, RT, and SP formed another group with intermediate values of parameter C. At the other extreme, the keratinized GI and HP were easily distinguished by a small C parameter.

In examining the absorption parameters, we found that the GI, HP, and SP demonstrated low values of cHb compared to other sites. The mean oxygen saturations were comparable for all nine sites, and the combined set of sites had a mean value of 60%. The mean effective vessel radius for the nine sites ranged from approximately 3 to 6 μm.

For the NADH/tryptophan and NADH/collagen fluorescence ratios at 308-nm excitation, the FM was statistically different from all other sites. The extracted contribution of Coll427 (340 nm) was slightly higher than that of Coll401 (340 nm) for every site except the LT and RT. The ratio of NADH to Coll401, Coll427 and total collagen, respectively was highest for the GI and HP, largely due to the low collagen emission for these sites. The GI and HP demonstrated the lowest mean Coll401/Coll427 ratio of all the sites. Examining the individual contributions of Coll401 and Coll427 extracted from the non-normalized fluorescence, both values were found to be significantly lower than all other sites.

As shown in Table 2, the RSD values for the area-normalized fluorescence parameters were considerably smaller than those of the non-normalized fluorescence data, particularly for individual sites. Similar to the results for the non-normalized fluorescence data, the mean NADH (308 nm) and tryptophan (308 nm) parameters for the FM were statistically different from all other sites. However, this was not the case for the mean collagen (308 nm) value. The GI and HP displayed the two lowest mean collagen (308 nm) values. Similar to the results for the Coll401/Coll427 ratio, the mean Coll427 (340 nm) value exceeded the mean Coll401 (340 nm) value for each site except for the LT and RT. Although the GI and HP displayed low mean Coll401 (340 nm) values compared to other sites, as was observed for the non-normalized fluorescence, they demonstrated mean Coll427 (340 nm) values, which were comparable to other sites.

3.4 k-Means Cluster Analysis

k-Means clustering was used as an objective and quantitative means of identifying and determining the magnitude of differences and similarities among sites, based on the spectral data summarized in Table 2. The results are shown in Table 3. For each parameter, we list the site(s) meeting the criteria for Δ described earlier and the range of values for Δ for each cluster. For those sites meeting these criteria, each row represents a separate clustering assignment; therefore, sites that were assigned to the same cluster are listed together, while those assigned to separate clusters are listed in different rows. For all but one parameter, oxygen saturation, there is at least one site or group of sites that form a well-defined cluster. For 15 of the 19 parameters studied, the HP was identified as being distinct from the majority of the other sites, either alone or in combination with another site(s). The GI also frequently demonstrated significant clustering (10 parameters). For 8 of the 19 parameters, the HP was assigned to the same cluster as the GI. Only in the case of one parameter [NADH/tryptophan (308 nm)] did the DT group with these sites. The FM dis-

Table 2 Summary of the mean, standard deviation (σ), and relative standard deviation (RSD) for each parameter for each of the nine sites, as well as the combined group of sites.

Parameter		BM	DT	FM	GI	HP	LT	RT	SP	VT	Combined
A											
	n	95	120	109	55	51	108	33	42	72	685
	Mean	1.7	1.4	1.5	1.5	1.4	1.5	1.7	1.6	1.5	1.5
	σ	0.1	0.1	0.2	0.2	0.2	0.1	0.2	0.1	0.1	0.2
	RSD [%]	8.6	8.8	14	12	11	7.9	11	9.4	9.1	12
B											
	n	100	124	110	51	52	110	33	43	74	697
	Mean	0.4	0.3	0.2	0.6	0.6	0.1	0.4	0.3	0.2	0.3
	σ	0.2	0.2	0.1	0.1	0.2	0.1	0.2	0.2	0.1	0.2
	RSD [%]	47	48	65	20	28	69	48	54	68	64
C											
	n	98	121	110	55	51	109	33	42	74	693
	Mean	0.010	0.012	0.02	0.004	0.002	0.018	0.008	0.009	0.02	0.013
	σ	0.006	0.007	0.01	0.004	0.002	0.006	0.008	0.005	0.01	0.009
	RSD [%]	61	62	53	96	112	32	96	59	46	72
cHb											
	n	88	113	103	53	48	102	32	38	72	649
	Mean	1.2	1	2	0.6	0.4	3	1.0	0.6	2	1
	σ	0.6	1	1	0.3	0.1	1	0.8	0.3	1	1
	RSD [%]	45	70	80	58	29	55	81	48	61	82
α											
	n	99	123	111	56	53	112	33	42	74	703
	Mean	0.6	0.7	0.6	0.6	0.7	0.7	0.6	0.7	0.6	0.6
	σ	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.1	0.2	0.2
	RSD [%]	24	20	32	16	13	22	25	21	29	24
Vessel radius [mm]											
	n	96	115	106	51	49	106	32	39	74	668
	Mean	0.004	0.005	0.005	0.0032	0.0028	0.006	0.004	0.0030	0.006	0.005
	σ	0.001	0.002	0.002	0.0006	0.0002	0.002	0.002	0.0003	0.002	0.002
	RSD [%]	31	40	46	18	7	33	44	11	36	44
$c\beta C$											
	n	96	123	107	55	53	106	31	43	73	687
	Mean	0.003	0.002	0.003	0.0012	0.0014	0.003	0.002	0.004	0.003	0.003
	σ	0.001	0.001	0.002	0.0008	0.0008	0.001	0.001	0.002	0.002	0.002
	RSD [%]	45	71	57	64	58	46	50	48	54	65

Table 2 (*Continued.*)

Parameter		BM	DT	FM	GI	HP	LT	RT	SP	VT	Combined
NADH/collagen (308 nm)											
	n	100	120	102	56	46	105	33	41	73	676
	Mean	0.4	0.39	0.6	0.34	0.40	0.39	0.3	0.5	0.5	0.4
	σ	0.1	0.09	0.1	0.08	0.06	0.09	0.1	0.1	0.1	0.1
	RSD [%]	30	23	17	24	15	23	36	24	21	29
NADH/tryptophan (308 nm)											
	n	98	119	110	51	52	107	33	39	72	681
	Mean	0.22	0.14	0.4	0.08	0.12	0.18	0.13	0.25	0.3	0.2
	σ	0.08	0.05	0.1	0.03	0.04	0.06	0.07	0.08	0.1	0.1
	RSD [%]	36	39	30	38	31	32	59	30	38	58
NADH/Coll401 (340 nm)											
	n	100	118	109	54	51	109	33	41	73	688
	Mean	0.6	0.8	0.8	1.1	1.3	0.4	0.5	0.8	0.7	0.8
	σ	0.2	0.3	0.2	0.6	0.4	0.2	0.2	0.2	0.2	0.4
	RSD [%]	32	36	27	53	29	44	46	29	27	49
NADH/Coll427 (340 nm)											
	n	98	120	106	52	49	108	31	42	71	677
	Mean	0.6	0.6	0.6	0.7	0.8	0.5	0.6	0.6	0.5	0.6
	σ	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.2
	RSD [%]	43	37	38	22	22	42	30	27	25	38
NADH/total collagen (340 nm)											
	n	99	117	109	58	49	109	32	42	70	685
	Mean	0.3	0.4	0.3	0.4	0.5	0.22	0.2	0.4	0.28	0.3
	σ	0.1	0.1	0.1	0.2	0.1	0.09	0.1	0.1	0.06	0.1
	RSD [%]	36	31	31	36	23	42	39	28	22	40
Coll401/Coll427 (340 nm)											
	n	96	120	106	57	50	111	30	42	71	683
	Mean	0.9	0.8	0.8	0.7	0.6	1.1	1.2	0.8	0.8	0.8
	σ	0.2	0.3	0.2	0.3	0.1	0.3	0.3	0.1	0.1	0.3
	RSD [%]	27	34	25	41	19	29	25	16	19	34
Tryptophan (308 nm)											
	n	100	120	109	51	54	108	33	39	72	686
	Mean	0.28	0.35	0.20	0.46	0.39	0.30	0.37	0.25	0.23	0.3
	σ	0.06	0.07	0.04	0.06	0.06	0.05	0.09	0.05	0.06	0.1
	RSD [%]	23	20	19	12	14	17	24	19	25	32

Table 2 (*Continued.*)

Parameter		BM	DT	FM	GI	HP	LT	RT	SP	VT	Combined
Collagen (308 nm)											
	n	91	114	103	55	49	104	32	39	70	657
	Mean	0.43	0.40	0.43	0.37	0.37	0.44	0.42	0.42	0.44	0.42
	σ	0.04	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.04
	RSD [%]	8.8	8.2	7.9	7.2	5.3	7.6	7.1	8.1	5.7	9.5
NADH (308 nm)											
	n	100	121	106	56	52	109	33	38	74	689
	Mean	0.28	0.24	0.36	0.19	0.23	0.26	0.21	0.31	0.32	0.27
	σ	0.06	0.05	0.05	0.06	0.04	0.05	0.07	0.05	0.06	0.08
	RSD [%]	22	22	13	31	19	20	36	15	20	28
NADH (340 nm)											
	n	100	118	106	58	49	109	32	42	73	687
	Mean	0.29	0.34	0.32	0.37	0.40	0.23	0.27	0.32	0.28	0.31
	σ	0.08	0.07	0.06	0.08	0.05	0.08	0.07	0.06	0.05	0.08
	RSD [%]	27	21	19	23	13	33	27	19	18	27
Coll401 (340 nm)											
	n	100	122	111	58	52	110	32	39	74	698
	Mean	0.32	0.28	0.29	0.25	0.22	0.38	0.39	0.30	0.30	0.30
	σ	0.05	0.06	0.04	0.09	0.04	0.07	0.07	0.03	0.04	0.07
	RSD [%]	16	23	14	35	18	19	19	11	15	25
Coll427 (340 nm)											
	n	92	118	103	54	54	104	32	43	71	671
	Mean	0.37	0.38	0.39	0.38	0.37	0.37	0.34	0.37	0.41	0.38
	σ	0.06	0.07	0.06	0.04	0.05	0.05	0.05	0.04	0.04	0.06
	RSD [%]	16	18	16	11	13	13	14	10	10	15

played unique properties for all of the non-normalized fluorescence parameters at 308-nm excitation, and two of the three area-normalized fluorescence parameters. The sites least likely to show significant clustering were the BM and DT. Each of these sites only met the criteria for Δ for a single parameter.

On the basis of the clustering results summarized in Table 3, the BM and FM would be expected to show some degree of separation based on any of the following parameters: A, cHb, vessel radius, NADH/collagen (308 nm), NADH/tryptophan (308 nm), tryptophan (308 nm), and NADH (308 nm). In each case, either the BM or FM met the criteria for Δ, and if both met the criteria, they were assigned to separate clusters. Figure 3(a) shows a binary scatter plot demonstrating a clear separation for these sites based on A and NADH/tryptophan (308 nm). The diagnostic decision line is also shown. The sensitivity, specificity, and AUC for separating the two sites based on these parameters were 81%, 83%, and 0.89, respectively. Similar results were obtained using either tryptophan (308 nm) and A (sensitivity, specificity, AUC of 87%, 79%, and 0.84, respectively) or NADH (308 nm) and A (sensitivity, specificity, AUC of 83%, 78%, and 0.87, respectively). As shown in Fig. 3(b), two different aspects of the tongue, the DT and VT, can be differentiated based on NADH/tryptophan (308 nm) and cβC with a sensitivity, specificity, and AUC of 92%, 78%, and 0.90 respectively. A sensitivity, specificity, and AUC of 92%, 81%, and 0.91 respectively, were obtained using cβC and tryptophan (308 nm). A sensitivity, specificity,

Table 3 Summary of the results from the *k*-means cluster analysis. For each parameter, sites in which Δ exceeded 30.0% for all three separations (*k*=2, *k*=3, and *k*=4) or at least two separations are listed. For sites meeting these criteria, those assigned to the same cluster are listed in the same row, while sites assigned to different clusters are listed in separate rows. In the last column, the range of values for Δ observed for a specific cluster is listed.

Parameter	Sites	Range for Δ [%]
A	BM	30.9–38.4
B	GI,HP	36.5–56.1
	LT,VT	30.4–44.5
C	GI,HP	38.6–66.7
cHb	GI,HP,RT,SP	34.4–75.0
	BM	30.7–40.9
α	No distinct sites	
Vessel radius	BM,GI,HP,SP	32.3–75.0
cβC	GI,HP	38.4–48.2
NADH/collagen (308 nm)	GI	30.4–41.1
	HP	31.5–34.1
	FM	32.4–48.0
NADH/tryptophan (308 nm)	DT,GI,HP,RT	34.9–69.1
	FM	36.7–43.6
NADH/Coll401 (340 nm)	HP	42.2–47.5
	LT,RT	34.8–42.7
NADH/Coll427 (340 nm)	LT,VT	30.6–39.8
NADH/total collagen (340 nm)	HP	39.8–50.5
	LT	33.6–38.1
Coll401/Coll427 (340 nm)	HP	48.0–53.7
	RT	40.0–41.7
Tryptophan (308 nm)	GI,HP	35.2–65.2
	FM,SP,VT	30.8–63.9
Collagen (308 nm)	GI,HP	46.4–56.5
NADH (308 nm)	HP	34.6–34.6
	GI	37.5–47.0
	FM	40.1–53.5
	SP	34.2–40.8
NADH (340 nm)	HP	39.8–50.3
	LT	30.9–36.5

Table 3 (*Continued.*)

Parameter	Sites	Range for Δ [%]
Coll401 (340 nm)	HP	48.1–57.1
	RT	34.4–40.6
Coll427 (340 nm)	VT	34.3–35.9

and AUC of 74%, 94%, and 0.89, respectively, were obtained using cβC and NADH (308 nm). Figure 3(c) shows a similar binary scatter plot for the GI and HP of cHb versus B, two parameters for which these sites were assigned to the same cluster. As the plot shows, the two sites overlap considerably (AUC=0.53).

Figures 4(a)–4(c) show pictures of the normal tissue architecture for the GI, HP, and BM. As can be seen for the former two sites, their strong spectroscopic similarity reflects a shared morphology. In contrast, the BM [Fig. 4(c)] displays a markedly different organization, in particular, the absence of overlying keratin.

3.5 Validation of Normal Tissue Discrimination

We next evaluated how accurately the results in Table 2 captured the parameter distributions for normal, healthy tissue ("normal limits") for each site and can be used to characterize potentially malignant, but clinically normal sites. The complete set of tissue data was divided into a set of calibration data and a set of validation data. The calibration data consisted of 584 spectra collected from 65 subjects and was used to derive the normal limits for each parameter for each site based on the 95% confidence interval (CI). The IQR exclusion criteria were applied to the calibration data in the same way as for analysis of the complete set of tissue data. The validation data consisted of 126 spectra collected from the remaining 14 subjects. These data were used to test the reliability of the calibration data to correctly categorize normal tissue. We used the calibration data to determine the parameter distributions and 95% CIs for each parameter for all nine sites. Then, for each parameter extracted from a sample in the validation data, we tested whether it lay within the 95% CI for that specific parameter. The site from which the sample was measured was also taken into account. Based on the results of this test we calculated the percentage of the validation data samples that were identified as normal, healthy tissue (within the 95% CI) for each parameter. The average age (±standard deviation) for subjects in the calibration and validation data was 41.0±16 and 43.8±18 years, respectively. Table 1 provides a summary of the number of spectra collected from each site for both sets of data. The percentage of all the samples in the validation data falling within the normal limits developed from the calibration data is shown in Table 4.

4 Discussion

4.1 Overview

The oral cavity is comprised of a number of different anatomic sites, and exhibits a wide variety of architectural properties. One notable example is that while most of the mucosa
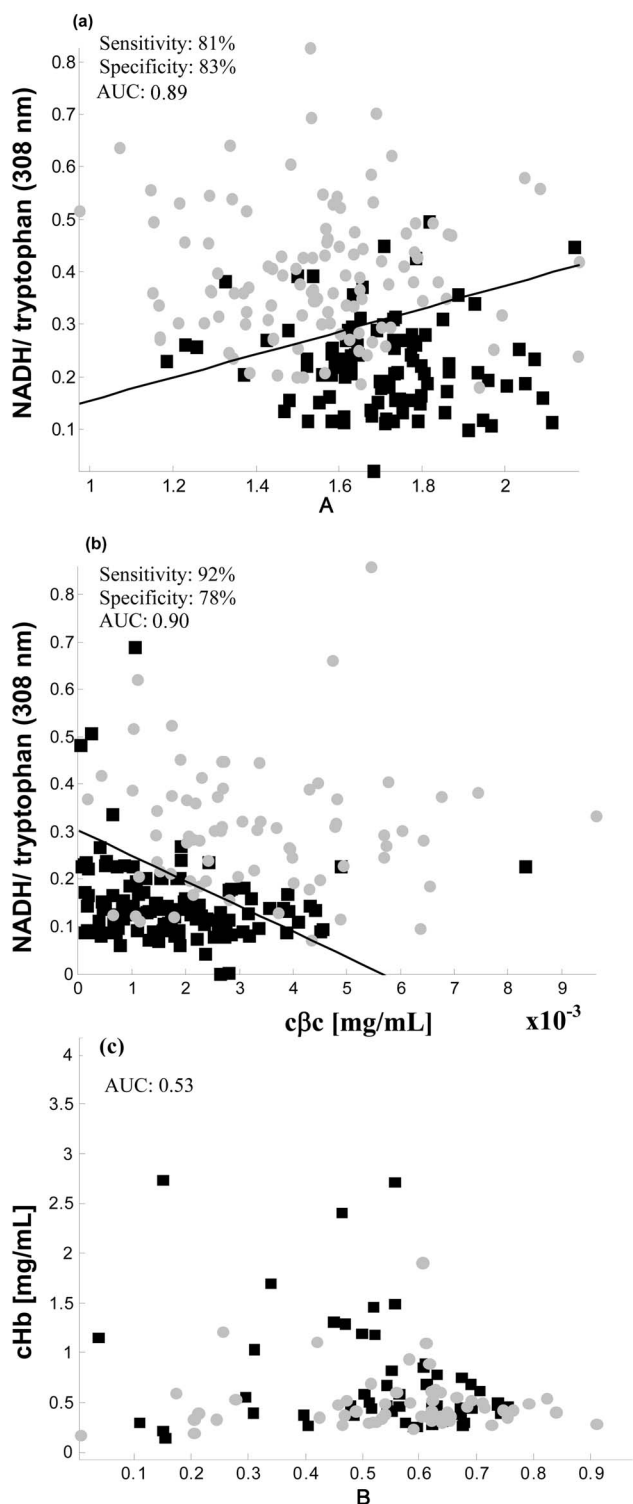
Fig. 3 Binary scatter plots comparing two sites. (a) Plot of NADH/tryptophan (308 nm) versus A for the FM (gray circles) and BM (black squares). (b) Plot of NADH/tryptophan (308 nm) versus $c\beta C$ for the DT (black squares) and VT (gray circles). (c) Plot of cHb versus parameter B for the GI (black squares) and HP (gray circles). The decision line providing the best sensitivity and specificity for separating the two sites, is shown in (a) and (b).

Fig. 4 Photos showing the normal microscopic architecture of (a) the HP and (b) the GI sites, which were frequently found to be spectroscopically similar to each other. The contrasting architecture of the BM is shown in (c). Overlying keratin is indicated by **k**.

of the oral cavity is nonkeratinized, the GI, HP and some regions of the DT are keratinized. Other key differences include the thickness of the epithelium (i.e., BM≫FM), the composition of the submucosa (i.e., bone, muscle), the density and size of the vessels, as well as the presence of other unique features, such as glands (i.e., BM, HP, SP) and papillae (DT). We expect these variations to also be reflected in the spectral information we collect from the tissue.

4.2 Reflectance and Fluorescence Modeling

In order to fully capture the major spectral features necessary for accurately modeling the reflectance and fluorescence spectra of healthy oral tissue, we incorporated the absorber

Table 4 Summary of the percentage of the 126 spectra in the validation data collected from healthy volunteers in which the listed parameters fell within the 95% CI ("normal limits") for a specific parameter for a given site. The normal limits were developed based on the results of the calibration data.

Extracted parameter	Percentage of data within 95%CI
A	95.2
B	93.7
C	93.7
cHb	88.1
Oxygen saturation	95.2
Vessel radius	92.1
cβC	88.1
NADH/collagen (308 nm)	96.8
NADH/tryptophan (308 nm)	93.7
NADH/Coll401 (340 nm)	92.1
NADH/Coll427 (340 nm)	88.9
NADH/total collagen (340 nm)	92.9
Coll401/Coll427 (340 nm)	97.6
Tryptophan (308 nm)	93.7
Collagen (308 nm)	93.7
NADH (308 nm)	96.0
NADH (340 nm)	94.4
Coll401 (340 nm)	95.2
Coll427 (340 nm)	96.8

β-carotene. Although β-carotene has not been previously noted in reflectance spectra collected from the oral cavity, there are numerous reports in the literature of its presence in oral cells/tissue, as well as in plasma.[32–34] This absorber is found at much lower concentrations than hemoglobin, but its extinction coefficient is >200 times that of hemoglobin at 450 nm.[26] Distortions of the hemoglobin spectrum due to vessel packaging effects were also taken into account in the reflectance model. This correction has been used in other investigations studying tissue reflectance.[18,21]

We model the fluorescence emission produced at 340-nm excitation by a linear combination of three fluorophores: NADH and two additional spectra, both of which we attribute to collagen(s), likely in combination with elastin. To determine if keratin contributed to the measured fluorescence, we separated the keratinized sites (GI, HP, DT) from the nonkeratinized sites and performed MCR analysis separately on these two groups. The basis spectra extracted for both groups were

identical. Furthermore, the keratinized sites did not demonstrate larger contributions of either of the two collagen components, as compared to nonkeratinized sites. The necessity for two distinct collagen spectra may also result from differences in the types of collagen being excited, the layered tissue architecture, depth-related effects (which can shift the proportions of light penetrating to the basement membrane or stromal layers), or a combination thereof. Spectra collected from pure collagen show, for example, that collagen IV (major type in the basement membrane) has a different emission profile from collagen I (major type in the stroma).[35] This supports the fact that variations in the collagen signal exist, although the emission of the extracted spectral components and chemicals do not appear to be equivalent. A study of freshly excised *ex vivo* cervical tissue by Chang et al. found that the stromal fluorescence differed greatly among various specimens and that data from their large patient data set were best modeled using a linear combination of three separate collagen fluorescence spectra.[36] The authors attributed two of the components to different types of collagen cross-links, the structural components responsible for collagen fluorescence. Differences between the spectra of the pure chemicals and the results from MCR are likely due to the inability of the dry chemical to recapitulate the actual protein *in vivo*, where the fluorescence emission may be influenced by the tissue microenvironment, the contribution of other types of collagens, and the presence of other fluorophores that strongly overlap in emission, such as elastin and possibly keratin.

4.3 Model-Based Parameter Distributions

The detailed summary of the parameters provided in Table 2 reveals a number of findings about the relationship between the extracted parameters and the structural characteristics of the sites. The GI, DT, and HP are sites that are normally keratinized; however, only the GI and HP display unique features as compared to the nonkeratinized sites. The keratinized sites are distinguished by high values for B, tryptophan (308 nm), and NADH (340 nm), and low values for C, cHb, effective vessel radius, cβC, Coll401/Coll427, NADH/total collagen (340 nm), collagen (308 nm), and NADH (308 nm). The high B value for the GI and HP may reflect strong scattering by the superficial keratin layer, resulting in a rapid fall in the scattering as a function of wavelength. In their depth-resolved studies of fluorescence from epithelial tissue, Wu and Qu observed a similar trend in the fluorescence intensity.[37] They found that the presence of keratin caused the fluorescence intensity to rapidly decay with increasing sampling depth, which was not observed for nonkeratinized tissues. The DT, which is only partially keratinized, exhibited a B value comparable to the nonkeratinized sites.[38] The presence of keratin appears to diminish signals from small, more superficial cellular scatterers, resulting in low C values for the HP and GI.

The GI, HP, and SP demonstrated very low concentrations of hemoglobin as compared to other sites. In the case of the former two sites, these results may indicate that keratin prevents light from reaching the underlying stroma where the blood vessels are located. During data collection from the GI, visible blanching of the mucosa was often observed when the probe contacted the tissue, which may also contribute to the

lower cHb values obtained for this site. The FM, LT, and VT,—sites with relatively thin epithelia as compared to other sites—demonstrated the highest cHb. The oxygen saturation values we obtain ($\sim$60%) are comparable to those measured in studies of the oral cavity and other tissues.[17,21] The mean effective vessel radius was largest for areas on the tongue (DT, VT, LT) and the FM. Blood vessels are easily observable at the macroscopic level on the VT and FM.

Because β-carotene is present in adipose tissue and in blood, we would expect the sites that demonstrated a high cHb and contain submucosal fat to exhibit the highest concentrations of this absorber. The highest concentration was found in the SP, which meets both these criteria. The FM, LT, VT, and BM were tied for the next highest value. The former three sites were those that displayed the highest cHb and the BM and FM both contain adipose tissue in the submucosa. Similar to the results for cHb, the GI and HP exhibit the lowest concentration of the absorber.

At 308-nm excitation, light is readily absorbed by a number of tissue components. As a result, the light does not penetrate deeply into the tissue, and the emission signal is heavily influenced by superficial components. The FM displayed a mean value significantly higher ($p < 0.05$) than all the other sites for the NADH/tryptophan (308 nm) and NADH/collagen (308 nm) ratios, as well as the NADH (308 nm) and tryptophan (308 nm) parameters. The FM is distinguished by a very thin epithelium (190 ± 40 μm), as compared to 580 ± 90 μm for the BM, and 310 ± 50 μm for the HP.[30] At 340-nm excitation, the site with the highest values for the ratio of NADH to Coll401, Coll427, and the total collagen contribution, respectively, was the HP. The GI had the second highest value for the former two ratios, but was tied for second with the DT and SP for the latter ratio. These ratios are affected by a complex combination of epithelial and stromal properties. The GI and HP demonstrated the lowest mean Coll401/Coll427 ratio of all the sites. In a study by Collier et al. examining cervical tissue by confocal microscopy, the authors found that keratin (which can be present both in normal and precancerous tissue) introduces a significant source of variability in the depth penetration of the incident light and, therefore, the extent to which scattering from different layers contribute to the observed signal.[39] Although there is a significant amount of collagen in the deeper stroma (composed mostly of collagen I), this is offset by the higher probability of light reaching and returning from more shallow structures, such as the basement membrane (composed mostly of collagen IV). Most sites display Coll401/Coll427 ratios of <1 [also Coll401 (340 nm) < Coll427 (340 nm)], indicating that emission from Coll427 (340 nm), which is similar in profile to collagen IV, is preferentially collected over Coll401 (340 nm), which most closely resembles collagen I. The presence of keratin on the GI and HP likely further reduces penetration of light to the deeper stroma thus resulting in the low Coll401 (340 nm) and Coll401/Coll427 values.

For most parameters, the HP and, to a lesser extent, the SP were also unique in that the RSDs of the parameter distributions for these sites were significantly smaller than for most other sites. We attribute this to the fact that these sites are relatively hard and noncompressible and, therefore, less prone to error from variations in probe pressure. In addition to the

natural variation among various sites as a result of the architecture of the mucosa, regional variations in the morphology within a given site, poor accessibility with the probe (i.e., the RT, some locations along the palate), the intrinsic movement of the tissue (i.e., the tongue), and the highly compressible nature of the softer tissue sites also contribute to the width of the parameter distributions. Analysis of data collected at different points along the palate and tongue, respectively, (data not shown) demonstrated not only the variation within a site, but also that areas of transition between sites can produce either gradual or marked shifts in a parameter, depending on its sensitivity to the specific structural differences between the two sites. These transitions can be appreciated based on gross visual inspection of oral tissue, as well as microscopic examination.[30] In future studies focused on disease detection, the area-normalization procedure should be performed for the analysis of the fluorescence data because this resulted in smaller spreads in the parameter distributions compared to non-normalized fluorescence data. The sources of the variations in the data will be discussed in further detail in the subsequent discussion of results for the validation of normal tissue discrimination.

4.4 k-Means Cluster Analysis

The degree of clustering based on anatomy can be very significant, as shown by the range of values for Δ (Table 3). The keratinized sites, the GI and HP, frequently clustered together, while the partially keratinized DT only displays significant clustering with these sites for a single parameter [NADH/tryptophan (308 nm)]. In addition to differences in the extent of keratinization, the DT is distinct from the GI and HP because it contains muscle in the submucosa rather than bone.

Despite the large RSD values and apparent overlap of the parameter distributions, clear spectral contrasts between sites exist based solely on anatomy [Figs. 3(a) and 3(b)]. The sensitivity and specificity values ranged between 78 and 92%, and for both distinctions, an AUC of $\sim$0.90 was achieved. The plots show that excellent results can be obtained when only two parameters are used for the distinction. However, by combining several parameters, the distinction may be further enhanced. By combining C with A and tryptophan (308 nm), a sensitivity, specificity, and AUC of 94%, 80%, and 0.88, respectively, can be achieved in discriminating between the BM and FM (sensitivity, specificity, and AUC of 87%, 79%, and 0.84, respectively, with only the latter two parameters). As another example, the FM and VT are poorly distinguished (sensitivity, specificity, and AUC of 80%, 47%, and 0.68, respectively) based on parameter A and tryptophan (308 nm). However, with the addition of NADH (340 nm), a sensitivity, specificity, and AUC of 72%, 72%, and 0.73, respectively, can be achieved.

The GI and HP could not be distinguished based on cHb and B, two parameters for which they are assigned to the same cluster [Fig. 3(c)]. Figure 4 shows the similarity of the GI and HP histology [Figs. 4(a) and 4(b)], and their contrast to the thicker, nonkeratinized epithelium of the BM [Fig. 4(c)]. The BM and GI were statistically different for 15/19 parameters, and the BM and HP were statistically different for 16/19 parameters. With only two parameters (cHb and A), the BM and HP could be distinguished with a sensitivity, speci-

ficity, and AUC of 89%, 95%, and 0.93, respectively (data not shown).

4.5 *Validation of Normal Tissue Discrimination*

Using the normal limits derived from the parameter distributions for the calibration data, we tested our ability to correctly classify parameters extracted from healthy tissue in a second set of validation data. We obtained excellent results, with few false positives (abnormal results), indicating that we have successfully captured the distributions for each parameter (Table 4). All of the fluorescence parameters except NADH/Coll427 (340 nm) demonstrate excellent accuracy in classifying normal tissue. The concentrations of hemoglobin and β-carotene are somewhat less accurate for identifying normal tissue. This is not surprising in that, within any given site (particularly those in which vessels may be located nonuniformly throughout the site), the probe may sample more or less hemoglobin or β-carotene depending on the proximity and density of the blood vessels.

There are a number of factors that contribute to the spread observed for a given site for the extracted parameters. These include variations among and within individuals, variations in the anatomy within a site, as well as probe-related factors. Because the data analyzed in the calibration and validation sets were collected from a large number of subjects, the effect of probe pressure can contribute to the observed spread for each site. We performed two tests evaluating the effect of variable probe pressure and repeat measurements in the same location, respectively, and found that the spread produced in these tests was smaller than that observed in the HV data.

5 Conclusion

Understanding the normal spectral contrasts associated with healthy oral tissue is essential for guiding the analysis and interpretation of data collected from lesions, and the development of spectroscopic-based tools for cancer diagnosis. Because the oral cavity is complex, in that it includes several different anatomic sites, the goal of this work was to characterize the normal spectral contrasts. There were significant differences in the means and spreads of the parameter distributions among the nine anatomic sites evaluated. Our findings also demonstrated the strong correlation between the extracted spectral parameters and the physical properties of the tissue. Keratinized sites (HP, GI) were distinct from nonkeratinized sites for a number of parameters. We have identified which parameters are affected by keratin and how the parameters are affected. Even among nonkeratinized sites, certain sites could be distinguished with >80% sensitivity and specificity using only two parameters. The results of the *k*-means cluster analysis also demonstrated the magnitude of the contrasts among the various anatomic sites. Finally, we have developed thresholds for the range of normal values for each parameter for each site.

These results provide an important foundation for our ongoing patient studies. Our understanding of the impact of keratin on the spectral parameters will be valuable in our assessment of patient data, as many lesions (benign, dysplastic, or malignant) exhibit hyperkeratosis (leukoplakia). Using the thresholds for the range of normal values for each parameter for each site, we have the potential to determine whether we can detect malignancy in the absence of visible findings.

The findings in this work also highlight a number of major points to be considered in the application of spectroscopy to the detection of oral cancer. First, when comparing lesion categories, differences in the distribution of white (hyperkeratotic) lesions in the two categories could produce spectral contrast. As a result, the apparent discrimination between the categories could mistakenly be attributed to the disease process. This was shown by the frequent distinction of the HP and GI from the remaining nonkeratinized sites. The impact of this would be particularly pronounced when data collected from healthy volunteers are directly compared to data collected from patients with lesions, since oral lesions are often characterized by hyperkeratosis. When small numbers of benign samples are grouped with large numbers of samples collected from healthy volunteers without visible lesions, this effect is also likely to occur. Second, this work also emphasizes that by combining several anatomic sites, differences in anatomy can skew distributions, such that in comparing malignant tissue samples to benign lesions (or healthy) tissue samples, each of which are comprised of different proportions of each site, a separation that is completely unrelated to malignancy may arise. Finally, combining sites that have statistically different mean values can result in a larger spread in each parameter, and thus make it more difficult for a change in a spectral parameter associated with the presence of cancer to become significant. The combination of a site with a large parameter spread (large RSD) with a site with a smaller parameter spread (small RSD) would also make it more difficult to appreciate changes with disease for the site with the smaller parameter spread. These effects are clearly demonstrated by the group statistics for the combination of nine sites as compared to a single site.

In future studies of lesions, we may be able to combine certain sites with similar spectral properties. The number of biopsy cases demonstrating dysplasia or cancer is usually limited, and combining sites would increase the number of samples from which a diagnostic algorithm could be developed, and also extend the number of sites to which it could be applied. The disadvantage is that this may incur a decrease in sensitivity and specificity for cancer due to an increase in the overall spread for the parameter distributions by combining sites. This is being explored in our ongoing diagnostic study of patients with oral lesions.

The results presented in this study provide strong evidence that a robust and accurate spectroscopic-based diagnostic algorithm for oral cancer will need to be applied in a site-specific manner, to ensure accurate evaluation of malignancy and to avoid the confounding effects of anatomy.

References

1. D. C. G. de Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, H. J. C. M. Sterenborg, and J. L. N. Roodenburg, "Clinical study for classification of benign, dysplastic, and malignant oral lesions using autofluorescence spectroscopy," *J. Biomed. Opt.* **9**(5), 940–950 (2004).
2. G. S. Nayak, S. Kamath, K. M. Pai, A. Sarkar, S. Ray, J. Kurien, L.

D'Almeida, B. R. Krishnanand, C. Santhosh, V. B. Kartha, and K. K. Mahato, "Principal component analysis and artificial neural network analysis of oral tissue fluorescence spectra: Classification of normal premalignant and malignant pathological conditions," *Biopolymers* **82**(2), 152–166 (2006).

3. M. G. Muller, T. A. Valdez, I. Georgakoudi, V. Backman, C. Fuentes, S. Kabani, N. Laver, Z. M. Wang, C. W. Boone, R. R. Dasari, S. M. Shapshay, and M. S. Feld, "Spectroscopic detection and evaluation of morphologic and biochemical changes in early human oral carcinoma," *Cancer (N.Y.)* **97**(7), 1681–1692 (2003).

4. N. Subhash, J. R. Mallia, S. S. Thomas, A. Mathews, and P. Sebastian, "Oral cancer detection using diffuse reflectance spectral ratio R540/R575 of oxygenated hemoglobin bands," *J. Biomed. Opt.* **11**(1), 014018 (2006).

5. D. L. Heintzelman, U. Utzinger, H. Fuchs, A. Zuluaga, K. Gossage, A. M. Gillenwater, R. Jacob, B. Kemp, and R. R. Richards-Kortum, "Optimal excitation wavelengths for *in vivo* detection of oral neoplasia using fluorescence spectroscopy," *Photochem. Photobiol.* **72**(1), 103–113 (2000).

6. A. Sharwani, W. Jerjes, V. Salih, B. Swinson, I. J. Bigio, M. El-Maaytah, and C. Hopper, "Assessment of oral premalignancy using elastic scattering spectroscopy," *Oral Oncol.* **42**(4), 343–349 (2006).

7. H. J. van Staveren, R. L. P. van Veen, O. C. Speelman, M. J. H. Witjes, W. M. Star, and J. L. N. Roodenburg, "Classification of clinical autofluorescence spectra of oral leukoplakia using an artificial neural network: A pilot study," *Oral Oncol.* **36**(3), 286–293 (2000).

8. H. M. Chen, C. P. Chiang, C. You, T. C. Hsiao, and C. Y. Wang, "Time-resolved autofluorescence spectroscopy for classifying normal and premalignant oral tissues," *Lasers Surg. Med.* **37**(1), 37–45 (2005).

9. V. R. Kolli, H. E. Savage, T. J. Yao, and S. P. Schantz, "Native cellular fluorescence of neoplastic upper aerodigestive mucosa," *Arch. Otolaryngol. Head Neck Surg.* **121**(11), 1287–1292 (1995).

10. A. Gillenwater, R. Jacob, R. Ganeshappa, B. Kemp, A. K. El-Naggar, J. L. Palmer, G. Clayman, M. F. Mitchell, and R. Richards-Kortum, "Noninvasive diagnosis of oral neoplasia based on fluorescence spectroscopy and native tissue autofluorescence," *Arch. Otolaryngol. Head Neck Surg.* **124**(11), 1251–1258 (1998).

11. J. K. Dhingra, D. F. Perrault, K. McMillan, E. E. Rebeiz, S. Kabani, R. Manoharan, I. Itzkan, M. S. Feld, and S. M. Shapshay, "Early diagnosis of upper aerodigestive tract cancer by autofluorescence," *Arch. Otolaryngol. Head Neck Surg.* **122**(11), 1181–1186 (1996).

12. D. C. G. de Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, D. J. C. M. Sterenborg, W. M. Star, and J. L. N. Roodenburg, "Autofluorescence characteristics of healthy oral mucosa at different anatomical sites," *Lasers Surg. Med.* **32**(5), 367–376 (2003).

13. M. G. Muller, I. Georgakoudi, Q. G. Zhang, J. Wu, and M. S. Feld, "Intrinsic fluorescence spectroscopy in turbid media: Disentangling effects of scattering and absorption," *Appl. Opt.* **40**(25), 4633–4646 (2001).

14. J. W. Tunnell, A. E. Desjardins, L. Galindo, I. Georgakoudi, S. A. McGee, J. Mirkovic, M. G. Mueller, J. Nazemi, F. T. Nguyen, A. Wax, Q. G. Zhang, R. R. Dasari, and M. S. Feld, "Instrumentation for multi-modal spectroscopic diagnosis of epithelial dysplasia," *Technol. Cancer Res. Treat.* **2**(6), 505–514 (2003).

15. G. Zonios, L. T. Perelman, V. M. Backman, R. Manoharan, M. Fitzmaurice, J. Van Dam, and M. S. Feld, "Diffuse reflectance spectroscopy of human adenomatous colon polyps *in vivo*," *Appl. Opt.* **38**(31), 6628–6637 (1999).

16. J. R. Mourant, J. P. Freyer, A. H. Hielscher, A. A. Eick, D. Shen, and T. M. Johnson, "Mechanisms of light scattering from biological cells relevant to noninvasive optical-tissue diagnostics," *Appl. Opt.* **37**(16), 3586–3593 (1998).

17. P. R. Bargo, S. A. Prahl, T. T. Goodell, R. A. Sleven, G. Koval, G. Blair, and S. L. Jacques, "*In vivo* determination of optical properties of normal and tumor tissue with white light reflectance and an empirical light transport model during endoscopy," *J. Biomed. Opt.* **10**(3), 034018 (2005).

18. A. Amelink, H. J. C. M. Sterenborg, M. P. L. Bard, and S. A. Burgers, "*In vivo* measurement of the local optical properties of tissue by use of differential path-length spectroscopy," *Opt. Lett.* **29**(10), 1087–1089 (2004).

19. D. Hidovic-Rowe and E. Claridge, "Modelling and validation of spectral reflectance for the colon," *Phys. Med. Biol.* **50**(6), 1071–1093 (2005).

20. I. S. Saidi, S. L. Jacques, and F. K. Tittel, "Mie and Rayleigh modeling of visible-light scattering in neonatal skin," *Appl. Opt.* **34**(31), 7410–7418 (1995).

21. R. L. P. van Veen, W. Verkruysse, and H. J. C. M. Sterenborg, "Diffuse-reflectance spectroscopy from 500 to 1060 nm by correction for inhomogeneously distributed absorbers," *Opt. Lett.* **27**(4), 246–248 (2002).

22. W. Verkruysse, G. W. Lucassen, J. F. deBoer, D. J. Smithies, J. S. Nelson, and M. J. C. vanGemert, "Modelling light distributions of homogeneous versus discrete absorbers in light irradiated turbid media," *Phys. Med. Biol.* **42**(1), 51–65 (1997).

23. L. O. Svaasand, E. J. Fiskerstrand, G. Kopstad, L. T. Norvang, E. K. Svaasand, J. S. Nelson, and M. W. Berns, "Therapeutic response during pulsed laser treatment of port-wine stains: Dependence on vessel diameter and depth in dermis," *Lasers Med. Sci.* **10**(4), 235–243 (1995).

24. S. A. Prahl, "Tabulated molar extinction coefficient for hemoglobin in water," Oregon Medical Laser Center, ⟨http://omlc.ogi.edu/spectra/hemoglobin/summary.html⟩, **8** March 2008.

25. C. H. Best, N. B. Taylor, and J. B. West, *Best and Taylor's Physiological Basis of Medical Practice*, Williams & Wilkins, Baltimore (1991).

26. H. Du, R. C. A. Fuh, J. Z. Li, L. A. Corkan, and J. S. Lindsey, "PhotochemCAD: A computer-aided design and research tool in photochemistry," *Photochem. Photobiol.* **68**(2), 141–142 (1998).

27. Q. G. Zhang, M. G. Muller, J. Wu, and M. S. Feld, "Turbidity-free fluorescence spectroscopy of biological tissue," *Opt. Lett.* **25**(19), 1451–1453 (2000).

28. R. R. Wilcox, *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*, Springer, New York (2001).

29. N. J. Perkins and E. F. Schisterman, "The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve," *Am. J. Epidemiol.* **163**(7), 670–675 (2006).

30. T. A. Winning and G. C. Townsend, "Oral mucosal embryology and histology," *Clin. Dermatol.* **18**(5), 499–511 (2000).

31. S. S. Sternberg, *Histology for Pathologists*, Lippincott, Williams & Wilkins, New York (1997).

32. I. Paetau, D. Rao, E. R. Wiley, E. D. Brown, and B. A. Clevidence, "Carotenoids in human buccal mucosa cells after 4 wk of supplementation with tomato juice or lycopene supplements," *Am. J. Clin. Nutr.* **70**(4), 490–494 (1999).

33. Y. S. Peng, Y. M. Peng, D. L. McGee, and D. S. Alberts, "Carotenoids, tocopherols, and retinoids in human buccal mucosal cells—Intraindividual and interindividual variability and storage stability," *Am. J. Clin. Nutr.* **59**(3), 636–643 (1994).

34. R. Brandt, G. E. Kaugars, W. T. Riley, Q. Y. Dao, S. Silverman, J. G. L. Lovas, and B. P. Dezzutti, "Evaluation of serum and tissue-levels of beta-carotene," *Biochem. Med. Metab. Biol.* **51**(1), 55–60 (1994).

35. A. R. Ten Cate, *Oral Histology: Development, Structure, and Function*, Mosby, St. Louis (1998).

36. S. K. Chang, N. Marin, M. Follen, and R. Richards-Kortum, "Model-based analysis of clinical fluorescence spectroscopy for *in vivo* detection of cervical intraepithelial dysplasia," *J. Biomed. Opt.*, **11**(2), 024008 (2006).

37. Y. C. Wu and J. N. Y. Qu, "Autofluorescence spectroscopy of epithelial tissues," *J. Biomed. Opt.* **11**(5), 054023 (2006).

38. E. J. Reith and M. H. Ross, *Atlas of Descriptive Histology*, Harper and Row, New York, (1970).

39. T. Collier, M. Follen, A. Malpica, and R. Richards-Kortum, "Sources of scattering in cervical tissue: Determination of the scattering coefficient by confocal microscopy," *Appl. Opt.* **44**(11), 2072–2081 (2005).