

Probing local tissue changes in the oral cavity for early detection of cancer using oblique polarized reflectance spectroscopy: a pilot clinical trial

Linda T. Nieman

The University of Texas M. D. Anderson Cancer Center
Department of Biomedical Engineering*
1515 Holcombe Boulevard
Houston, Texas 77030

Chih-Wen Kan

The University of Texas at Austin
Department of Biomedical Engineering*
1 University Station
Austin, Texas 78712-1084

Ann Gillenwater

The University of Texas M. D. Anderson Cancer Center
Department of Head and Neck Surgery
1515 Holcombe Boulevard
Houston, Texas 77030

Mia K. Markey

The University of Texas at Austin
Department of Biomedical Engineering*
1 University Station
Austin, Texas 78712-1084

Konstantin Sokolov

The University of Texas M. D. Anderson Cancer Center
Department of Biomedical Engineering*
and Department of Imaging Physics
1515 Holcombe Boulevard
Houston, Texas 77030
E-mail: kostia@mail.utexas.edu

1 Introduction

Biomedical optics has shown great promise in extracting biochemical and morphologic information from precancerous and cancerous tissue *in vivo* that has been traditionally obtained exclusively by microscopic examination of excised tissue.¹⁻³ Optical techniques can provide clinicians greater ability to noninvasively detect and monitor precancerous lesions during screening and treatment, particularly for extensive lesions that require multiple biopsies to adequately assess disease stage. The high turnover rate of cells in the epithelium makes it the most common site for cancer to emerge.⁴ Epithelial cancers are well characterized to develop in a multistep process from the accumulation of genetic mutations over

Abstract. We report the results of an oral cavity pilot clinical trial to detect early precancer and cancer using a fiber optic probe with obliquely oriented collection fibers that preferentially probe local tissue morphology and heterogeneity using oblique polarized reflectance spectroscopy (OPRS). We extract epithelial cell nuclear sizes and 10 spectral features. These features are analyzed independently and in combination to assess the best metrics for separation of diagnostic classes. Without stratifying the data according to anatomical location or level of keratinization, OPRS is found to be sensitive to four diagnostic categories: normal, benign, mild dysplasia, high-grade dysplasia, and carcinoma. Using linear discriminant analysis, separation of normal from high-grade dysplasia and carcinoma yield a sensitivity and specificity of 90 and 86%, respectively. Discrimination of morphologically similar lesions such as normal from mild dysplasia is achieved with a sensitivity of 75% and specificity of 73%. Separation of visually indistinguishable benign lesions from high-grade dysplasia and carcinoma is achieved with good sensitivity (100%) and specificity (85%), while separation of benign from mild dysplasia gives a sensitivity of 92% and a specificity of 69%. These promising results suggest that OPRS has the potential to aid screening and diagnosis of oral precancer and cancer. © 2008 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2907450]

Keywords: elastic light scattering; polarized light; reflectance spectroscopy; cancer diagnosis; oral cavity.

Paper 07286R received Jul. 30, 2007; revised manuscript received Oct. 9, 2007; accepted for publication Nov. 16, 2007; published online May 1, 2008.

time, resulting from carcinogenic exposure. For the majority of cancers, survival rate and quality of life are greatly improved when cancer and its precursors are detected early. An epithelial cancer where early detection is crucial to successful clinical outcome is oral cancer.⁵ If caught early, the 5-yr survival rate increases dramatically from 26% for distant staging to 82% for local staging.⁶ Even in developed countries such as the United States, where dental exams are routine, oral cancer is often not discovered until it has infiltrated surrounding organs or tissues.⁶ Indeed, U.S. 5-yr survival rates have changed little in the last 30 yr, remaining⁷ at approximately 55%.

Current oral cancer screening methods are limited by the variety of tissue architecture and by the similarity of appearance of benign inflammatory conditions to premalignant and malignant lesions. Further complications arise in high-risk patients, who often have carcinogenic exposure that covers the entire mucosal lining. After successful treatment of a cancerous tumor, secondary tumors can develop in adjacent areas over time. Monitoring these high-risk patients requires mul-

*The University of Texas Department of Biomedical Engineering is an inter-institutional department combining the strengths of The University of Texas at Austin, The University of Texas M.D. Anderson Cancer Center, and The University of Texas Health Science Center at Houston.
Address all correspondence to Konstantin Sokolov, Imaging Physics, UT M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Unit 56 - Box 0056-3T, Houston, TX 77030, United States of America; Tel: 512-471-7440; Fax: 512-471-0616; E-mail: kostia@mail.utexas.edu

multiple biopsies taken routinely over many years. Clearly, non-invasive optical modalities to detect the early stages of oral cancer have the potential to reduce patient pain, morbidity, and mortality. To this end, significant efforts have been directed toward exploring optical imaging and optical spectroscopy techniques to aid screening and diagnosis of the early stages of oral cancer.^{8–20} A comprehensive overview of the new emerging methods for the detection and treatment of oral carcinoma has recently been given in Refs. 21 and 22.

Optical imaging modalities such as confocal microscopy, optical coherence tomography (OCT), and nonlinear optical microscopy have shown great potential for oral lesion discrimination. While these nascent approaches alone or in combination with exogenous dyes, vital stains, or nanoparticles are highly promising, they require some degree of equipment complexity and operator training. Autofluorescence imaging and spectroscopy has been shown to give improved lesion contrast, which has been attributed in part to porphyrin fluorescence. However, it has been argued that porphyrin fluorescence is not a good diagnostic indicator, as it is synthesized by bacteria not only on ulcerating tumors, but also on the dorsums of normal tongues and on gingival plaques.¹¹ The presence of porphyrin fluorescence can therefore obscure detection of lesion grade using fluorescence techniques. Spectroscopic approaches that combine tissue autofluorescence and scattering showed promising sensitivity and specificity in pilot clinical trials, however, it required stratification of tissue sites according to level of keratinization before statistical analysis.¹⁶

Here we present the results of the first pilot clinical trial that assesses the optical technique of oblique polarized reflectance spectroscopy (OPRS) to discriminate oral precancers and cancers from normal or benign tissue. OPRS is a noninvasive optical modality that employs polarized light illumination and polarization sensitive detection. This method is very simple and robust and, thus, provides an attractive approach for earlier cancer detection in the oral cavity including screening in a high-risk population. OPRS is based on the following concept: the electric field orientation, or polarization, of the incident light remains unchanged after interaction with the optically dilute epithelial layer. In contrast, photons that propagate deeper to the optically dense stroma are remitted from tissue with their polarization state randomized. In OPRS, two scattered signals are collected: one with polarization parallel and the other with polarization perpendicular relative to the illumination polarization. The small epithelial signal can therefore be isolated by subtracting the perpendicular polarization intensity from the parallel polarization intensity.^{23–27} Combining polarization-sensitive detection with an oblique endoscopic collection geometry further resolves scattering signals from the upper epithelial layer and from the lower stromal layer.²⁸ This additional depth-dependent optical information carried by polarized light has the potential to enhance discrimination of the varying grades of dysplasia and carcinoma from visually indistinguishable benign lesions.

The two scattering signals collected by OPRS can be used separately or in combination to yield diagnostically relevant parameters. In this pilot clinical study, we extracted 10 spectroscopic features and nuclear size of epithelial cells with the goal to identify key parameters for detection and monitoring of precancerous lesions in the oral cavity. A thorough statisti-

cal analysis of the spectral parameters and extracted nuclear size was performed using linear discriminant analysis (LDA) and evaluated using receiver operating characteristic (ROC) analysis. OPRS was found to be sensitive to four clinically relevant histological groups: normal, benign, mild dysplasia, and severe dysplasia (defined as tissue requiring surgical excision for treatment). We demonstrated that the features that provide the best discrimination differ according to diagnostic category. This result emphasizes that a combination of features is required to efficiently tackle the multitask problem of cancer detection and diagnoses.

2 Materials and Methods

2.1 Clinical Measurement

A pilot clinical study was conducted with informed consent on 27 patients over the age of 18 that were referred to the Head and Neck Clinic at The University of Texas M. D. Anderson Cancer Center (MDACC) with oral mucosa lesions suspicious for dysplasia or carcinoma. A medical doctor performed a standard oral cavity examination, followed by spectroscopic measurements which were typically performed on one to two visually abnormal sites and one visually normal site. In some cases, more than one measurement was performed on the same tissue site without removing the probe. All measurements from the same tissue site were averaged to give a single spectrum.

A calibration spectrum was acquired before or after each patient evaluation using a diffuse reflectance substrate standard (Labsphere, Inc.). Data from three patients were removed from the analysis because of improper handling of the endoscope or malfunctioning of the clinical device. Biopsies were taken of all measured tissue sites. The biopsied tissue was sectioned into 4 μm transverse slices and mounted onto microscope slides. The slides were stained with a hematoxylin and eosin (H&E) stain for standard histological analysis. Paired normal and abnormal slides were reviewed by a trained pathologist at MDACC. Detailed descriptions were made of each slide indicating the extent of dysplasia, inflammation, keratinization, and hyperplasia.

2.2 Instrumentation

Although detailed extensively in Refs. 28 and 29, we briefly describe the instrument used in this clinical study for convenience. The illumination source was a broadband white light Xe pulsed lamp with ca. 4- μs pulse widths and a wavelength range of 400 to 700 nm. Light was delivered to the tissue site of interest through a single optical fiber with a core diameter of 200 μm and 0.22 numerical aperture (NA). The power delivered to the tissue was approximately 100 μW , well below the acceptable threshold limit value given by the American Conference of Governmental Industrial Hygienists.³⁰ Two identical optical fibers were placed on either side of the illumination fiber for collection of the remitted light. Two pieces of polarizing film with an extinction transmittance of 0.002% were adhered to the distal end of the fiber optic probe. These polarizers set the orientation of the illumination and collection polarization states and were oriented orthogonal to each other. The two collection fibers had polarizing film with transmission axes either parallel or orthogonal to the illumination polarization state. A protective fused silica window of defined

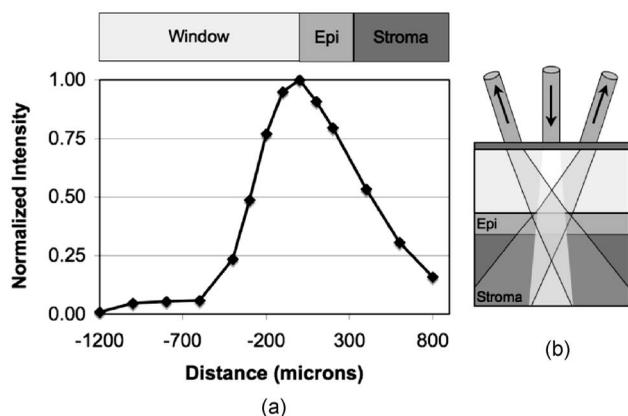


Fig. 1 (a) Overlap of illumination and collection acceptance cones as a function of distance from the distal end of the illumination and collection fibers. Distance zero corresponds to the epithelium surface. Measurements were made in water using a reflectance standard (Labsphere, Inc.). (b) Illustration of distal probe geometry. The center fiber delivers light to the tissue; the outer two fibers collect the scattered light with polarization parallel and orthogonal to the incident polarization.

thickness was placed over the fibers and polarizing film. The illumination fiber was oriented normal to the silica window surface and the collection fibers were oriented at ca. 37 deg with respect to the illumination fiber. With this geometry, the collection fiber's acceptance cones cross each other and with the illumination beam in the superficial tissue layer as shown in Fig. 1. The collection efficiency of a single obliquely oriented beveled fiber peaks at the maximum overlap of the illumination beam with the collection acceptance cone. Figure 1(a) is a plot of the collected intensity of a single collection fiber from a diffusely reflecting substrate (Labsphere, Inc.). The probe used in this study has a fused silica window that acts as a spacer such that the maximum overlap is at the tissue surface. The window and tissue thicknesses with respect to the maximum collection depth are illustrated above the plot in Fig. 1(a). This overlap drops to 50% at a depth of 300 to 400 μm , the typical thickness of oral epithelium. Hence, superficial traveling photons approximately 0.07 mm³ below the tissue surface are collected with greater efficiency than deeper traveling photons. The collection fibers deliver the remitted light to a grating spectrograph coupled to a gated intensified photodiode array detector. Operating in gated mode enables all measurements to be conducted under room light illumination so that OPRS measurements could blend seamlessly into a clinical examination or preoperative surgical procedure.

2.3 Data Analysis

2.3.1 Preprocessing

The collected parallel and perpendicular spectra were dark subtracted and then divided by the sum of the scattered light collected through both collection channels from a diffuse reflectance standard (Labsphere, Inc.) to correct for the wavelength-dependent response of the detection system and the spectral profile of the source. The parallel and perpendicular signals were studied alone and in the following combina-

tions: the ratio of parallel to perpendicular, parallel minus perpendicular, and the sum of the parallel and perpendicular signals. The sum of the parallel and perpendicular spectra is equivalent to the diffuse reflectance spectrum and their difference is defined as the depolarization ratio. Spectra were down-sampled using an averaging window with a spectral width of 5 nm to reduce data size and computation time.

The spectra were normalized to remove interpatient variation. Three approaches were tested: (1) no normalization, (2) division of entire spectra by the intensity value at 420 nm, and (3) additive dc offset applied to the entire spectra such that value at 420 nm was equal to the mean for all spectra within the same spectral type (e.g., parallel or ratio of parallel and perpendicular signals etc.). This normalization was intended to preserve the relative intensity scale between different spectral types. Normalization method 2 gave the highest area under the ROC curve for all classification tasks considered and hence was used throughout this study.

2.3.2 Determination of the Most Discriminatory Wavelength

The most discriminatory wavelengths were determined using the area under the ROC curve, which is a commonly used summary statistic to assess the effectiveness of a two-outcome classification process.³¹ The area under the nonparametric curve (AUC) was computed using the trapezoid rule. On the occasion when ties existed, the best wavelength was chosen manually such that wavelength variations were minimized.

2.3.3 Features

Two features were extracted from each spectrum of the five spectral types (parallel, perpendicular, diffuse, depolarization ratio, parallel/perpendicular), one being the mean intensity across the entire spectrum and the other being the intensity at the most discriminatory wavelength, yielding 10 spectral features per measurement site. In addition, nuclear size was extracted from the depolarization ratio spectrum using a Mie theory based algorithm described in Refs. 27 and 32. Figure 2(a) shows example depolarization ratio spectra for each diagnostic category and their corresponding fit. Figure 2(b) compares the mean nuclear size per diagnostic category extracted from OPRS measurements and from direct measurements of the biopsied tissue histology slides. Hence, a total of 11 features were extracted from each measurement site for use in eight binary classification tasks: (1) normal versus severe dysplasia (SD), (2) normal versus mild dysplasia (MD), (3) normal versus MD and SD combined, (4) MD versus SD, (5) benign versus SD, (6) benign versus MD, (7) benign versus MD and SD combined, and (8) benign versus normal.

2.3.4 Selection of the Most Discriminatory Features

In many classification tasks, a combination of features yields better discrimination than can be achieved with any single feature. Since there are 11 features in each of the two-class classification problems, $2^{11} - 1 = 2047$ different combinations of features are possible, for example: feature 1 alone; features 1 and 2 combined; features 1 and 3 combined; features 2 and 3 combined; features 1, 2, and 3 combined; etc. We exhaustively searched through these 2047 combinations to identify the feature or combination of features that best discriminates

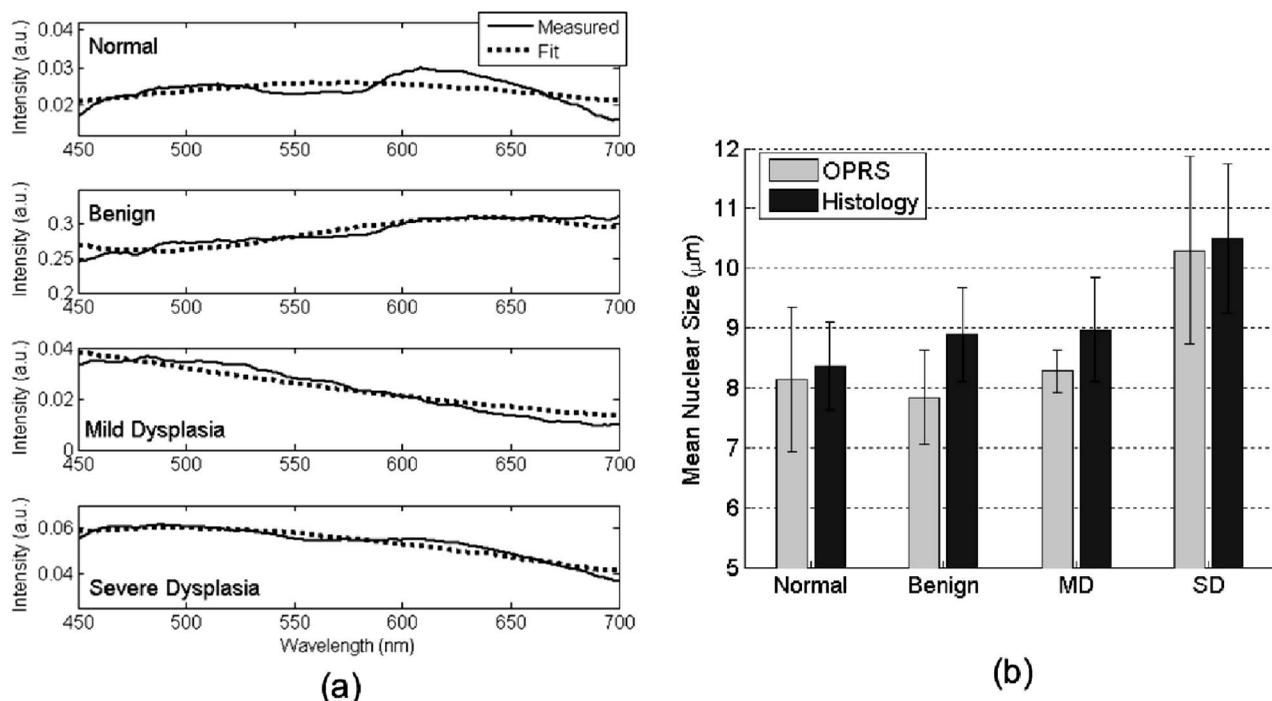


Fig. 2 (a) Examples of experimental depolarization ratio spectra (solid curves) for each diagnostic category and corresponding theoretical simulations (dotted curves) obtained using Mie scattering calculations and (b) bar graph of mean nuclear size per diagnostic category. Light gray bars indicate nuclear size extracted using OPRS measurements and dark gray bars indicate nuclear size obtained directly from stained and sliced biopsied tissue mounted on microscope slides. The standard deviations show variations in nuclear sizes between different patient sites.

between two diagnostic classes using³³ LDA. LDA was chosen because it works well with small datasets and it preserves the physical origins of features. Leave-one-out cross validation was employed to train and test all LDA models.

The performance of a feature combination was evaluated using the area (AUC) under the nonparametric ROC curve generated from the LDA decision variable. All possible LDA models were compared. The best LDA model was defined as that which had the highest area under the ROC curve. LDA calculations were carried out using the classify function in MATLAB® R7 Statistics Toolbox (The MathWorks, Natick, Massachusetts). For each binary classification task, several of the 2047 LDA models can have statistically equivalent discrimination. A bootstrapping technique was used to estimate the mean difference in the AUC between LDA models and the two-sided *p* value of that difference.³⁴ Any *p* values below the conventional threshold of 0.05 were regarded as statistically significant. The top *N* models with AUCs statistically indistinguishable from the maximum AUC observed were considered to be comparable models.

We hypothesized that dominant features would appear with higher frequency, while irrelevant features would appear randomly. Consequently, the fractional occurrence of individual features within the top *N* LDA models within each diagnostic category was counted. Those features that appeared with a frequency of 0.5 or greater are considered to be of diagnostic importance.

2.3.5 Precautions Taken to Avoid Overtraining

Overtraining is a concern any time one develops a classification model with a small dataset. To reduce this risk, we used

Table 1 Distribution of anatomical sites within the oral mucosa measured with OPRS.

Location	Clinical appearance		Diagnosis for Abnormals		
	Normal	Abnormal	Benign (B)	MD	SD ^a
Tongue	7	22	8	7	7
Buccal	13	7	3	3	1
Floor of mouth	1	2	0	1	1
Gingiva	1	2	1	0	1
Soft Palate	0	2	1	1	0
Total measured	22	35	13	12	10

^aSevere dysplasia is defined as tissue that requires surgery for treatment; it includes tissue that has the histopathological diagnosis of moderate to severe dysplasia or carcinoma. The physical appearance was noted prior to OPRS measurement as either normal or abnormal. Biopsies were taken of all 57 measured sites. Histopathology of abnormal biopsies is categorized as benign, MD, or SD.

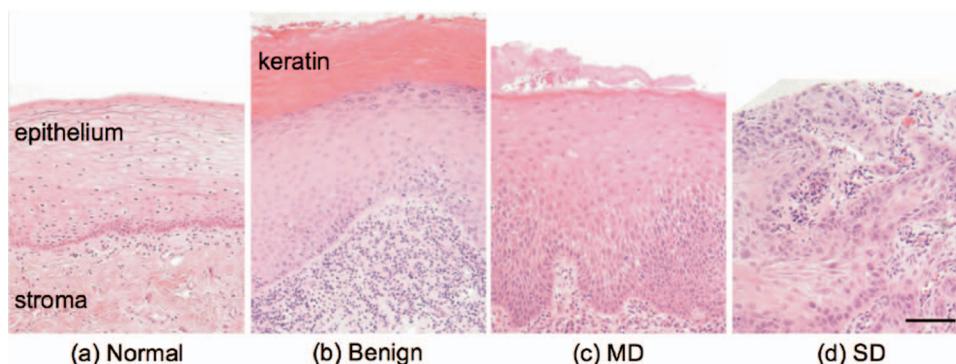


Fig. 3 Representative images of biopsied tissue diagnosed as (a) normal, (b) benign, (c) MD, or (d) SD. Tissue was stained with H&E for standard histopathological analysis. Scale bar is 100 μm .

leave-one-out cross validation to generate the LDA models. As a further check, a permutation test was applied in which the pathology definition of each patient measurement was randomly shuffled while the prevalence of disease was kept constant.³⁵ The shuffling was repeated 1000 times for each binary classification task, so that the mean and standard error of the area under ROC curves could be calculated and compared with the observed area under ROC curves.

3 Results

3.1 Sample Distribution

Table 1 summarizes the distribution of sites from the oral cavity that were measured, then subsequently biopsied. A total of 57 sites in 24 patients were measured and analyzed in this

study. Table 1 is divided into two categories based on the examining physician's visual impression at the time of biopsy: (1) normal and (2) abnormal. All visually normal sites were confirmed to be normal by histological analysis. The clinically appearing "abnormals" were further subdivided into three categories according to their histological diagnosis: benign (B), MD, or SD. We defined SD as tissue that requires surgical excision. In this study, all but one SD site were classified as carcinoma (the other was classified as moderate dysplasia). Figure 3 shows images taken of representative histopathology slides from this study illustrating the morphology of the different diagnostic categories. All binary combinations of the four diagnostic classes were used to test classification of OPRS data.

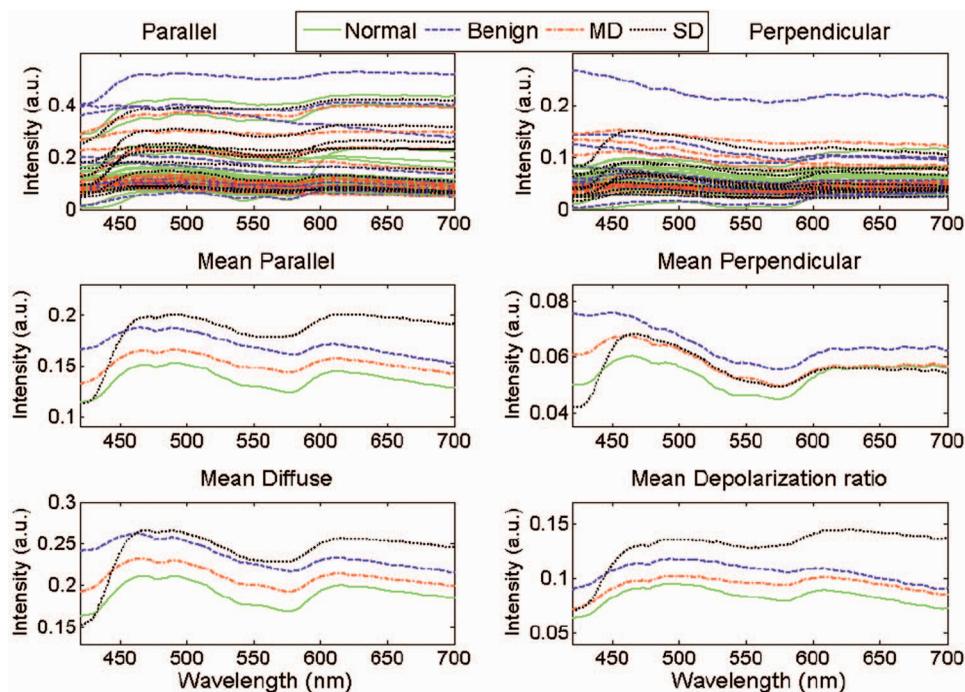


Fig. 4 Measured spectra. The first row shows all the spectra collected from the two collection channels. The second and third rows show the mean spectra according to diagnostic class for the parallel, perpendicular, diffuse reflectance, and depolarization ratio. Spectra from normal tissue are shown as solid green curves, benign tissue shown as dashed blue curves, mild dysplasia as red dash-dotted curves, and severe dysplasia as black dotted curves.

Table 2 The wavelengths, given in nanometers, that provided the maximum separation between two diagnostic classes.

Spectral Type	Normal from SD	Normal from MD	Normal from MD and SD	MD from SD	Benign from SD	Benign from MD	Benign from MD and SD	Benign from Normal
Parallel	581	465	655	581	539	597	497	465
Perpendicular	454	576	454	454	454	602	491	486
Diffuse	539	465	718	586	539	497	497	497
Depolarization ratio	602	465	628	454	539	497	581	465
Parallel/perpendicular	633	565	612	633	644	512	644	512

3.2 Measurement Repeatability

To assess the reproducibility of our measurements, 51 of the 57 tissue sites had two subsequent measurements performed without removing the probe. We calculated a mean intensity difference over all wavelengths between the two spectra from the same tissue site and their average spectrum. An average value for the intensity difference for all 51 sites was ca. 10%. The repeatability of spectral shape is also important especially when performing nuclear size extraction. Therefore, we ran the nuclear extraction algorithm on a subset of the 51 patient sites that had duplicate measurements. Fifteen sites were chosen randomly without reference to the diagnosis. The extracted nuclear size for both repeat measurements was compared to the nuclear size found using the mean spectrum. We found that the average nuclear size difference between the measured spectra and the mean spectrum was 0.45 μm . This is well below the size difference (ca. 2 μm) between diagnostic categories of normal and SD. However, these data indicate that probe handling is an important issue in clinical trials that can lead to an increase in variations of optical measurements in vivo. Translation of the probe between measurements can cause differences in the spectra as the tissue volume sampled will have changed. The tissue volumes may have differing scattering characteristics that may affect collection of parallel and perpendicular components and extraction of the depolarization ratio spectrum. In addition, it has been shown that the pressure applied to the tissue can significantly affect fluorescence spectral intensity.³⁶ Further work must be performed in this area to fully assess the effect of probe pressure and handling on spectroscopic measurements.

3.3 Polarized Reflectance Spectra

Figure 4 shows the measured spectra before normalization. The diagnostic category is indicated by the color of the curves: solid green, normal; dashed blue, benign; dash-dotted red, mild dysplasia; and dotted black, severe dysplasia. The first row shows all measured spectra from the parallel and perpendicular collection channels. The second and third rows show the averaged spectra per diagnostic class for the parallel, perpendicular, diffuse reflectance, and the depolarization ratio. Analysis of the mean spectra showed a good, albeit qualitative, separation of all diagnostic classes.

A qualitative look at the mean spectra reveals differences in the total intensity, which is modulated by hemoglobin absorption. Within a narrow wavelength band, the mean spectral differences can be quite large. For example, the mean perpendicular spectra have the largest separation between diagnostic classes for wavelengths shorter than 450 nm. Similarly, the mean depolarization ratio spectra have the largest separation between normal and SD in the red or long-wavelength region. As a whole, the mean spectra hint at the possibility of using select wavelength regions for improved diagnostic discrimination, thus prompting an analysis of the optimal wavelength for maximum separation of two diagnostic classes for each spectral type.

3.4 Diagnostically Relevant Features

Table 2 shows the best wavelengths determined for the eight two-outcome diagnostic classification tasks and five spectral types. All spectra used in this analysis were normalized to one at 420 nm, as described in Sec. 2.3.1. Table 3 lists the AUCs for individual features and the best combination of features, as determined by exhaustive LDA search, for each classification task considered. An AUC of 1 corresponds to perfect classification while an AUC of 0.5 corresponds to chance performance. Comparison of the individual feature AUC and the LDA AUC per classification task shows that the LDA model typically performs better. Note that the performance of individual features varies quite drastically, depending on the classification task. These results demonstrate that a combination of features is needed to efficiently tackle the multitask classification problem involved in cancer detection and diagnosis.

Discrimination of normal from SD has a relatively high LDA AUC. This is consistent with our expectations since normal and SD are two extremes of the continuum of histopathological status. Likewise, histopathologically similar tissue such as normal and MD is more difficult to discriminate. The similarity of normal and MD tissue can be seen in Fig. 3, where the areas of MD (indicated by increased nuclear density) encompass a small section at the basal layer. It is not uncommon for MD to be focally located amid normal tissue. In an OPRS measurement, the optical signal from MD is weighted by the surrounding normal tissue. Similarly, the AUC for the classification task of normal from MD and SD

Table 3 Area under the ROC curve for individual features and the best LDA combination of features.

<i>Individual Features</i>	Area under ROC							
	<i>Normal from</i>			<i>MD from</i>	<i>Benign from</i>			
	<i>SD</i>	<i>MD</i>	<i>MD and SD</i>	<i>SD</i>	<i>SD</i>	<i>MD</i>	<i>MD and SD</i>	<i>Normal</i>
Nuclear size	0.79	0.69	0.74	0.65	0.75	0.62	0.68	0.56
Mean parallel	0.81	0.50	0.64	0.82	0.88	0.65	0.76	0.69
Mean perpendicular	0.69	0.60	0.53	0.77	0.82	0.73	0.77	0.81
Mean diffuse	0.79	0.53	0.62	0.82	0.86	0.69	0.77	0.74
Mean depolarization ratio	0.84	0.53	0.67	0.82	0.90	0.60	0.74	0.57
Mean par/per	0.73	0.66	0.69	0.69	0.70	0.55	0.62	0.63
Parallel, x nm	0.84	0.60	0.68	0.87	0.90	0.68	0.78	0.77
Perpendicular, x nm	0.78	0.65	0.57	0.86	0.89	0.74	0.78	0.83
Diffuse, x nm	0.82	0.60	0.64	0.86	0.90	0.69	0.78	0.78
Depolarization ratio, x nm	0.85	0.61	0.71	0.86	0.91	0.64	0.76	0.72
Par/per, x nm	0.77	0.67	0.71	0.71	0.71	0.58	0.62	0.70
Best LDA combination	0.89	0.72	0.74	0.87	0.91	0.76	0.78	0.84

Eight binary classification tasks, indicated by the column headings, were considered. Features identified by x nm correspond to the intensity at the most discriminatory wavelength. A value of unity indicates perfect performance, while 0.5 is chance performance.

combined yield a similar AUC to that of normal from MD. In comparison, clear diagnostic potential is shown for the classification task of MD versus SD.

One of the most clinically challenging classification tasks requires the ability to distinguish between dysplastic tissue and benign tissue, which has the outward appearance of dysplasia or carcinoma but is histologically normal. Therefore, it is encouraging that some features in our study show clear separation between benign and SD sites. Similar to what is observed when normal tissue is used as the reference, there is less discriminatory power for the classification tasks differentiating benign from MD or benign from MD and SD combined.

Due to the limited amount of data in this pilot study, there is not enough statistical power to distinguish between small differences in ROC area. For example, in the classification task of normal from SD, 36 LDA feature sets with the highest ROC areas were found to be statistically indistinguishable. Rather than attempting to identify a single “best” model, which is impossible to do in a pilot study, we identified the features that most frequently appear in the set of statistically indistinguishable best performing models. Table 4 shows the frequency with which each feature occurs in the top N feature sets with statistically similar ROC areas for each classification task. Dashed lines delineate features that appear with a frequency of 0.5 or greater. We suggest that these features warrant the greatest attention in future large clinical trials.

3.5 Check for Overtraining

With a small data set, overtraining is always a concern. As noted in the methods section, leave-one-out cross validation was used for classifier training to reduce this risk. As an additional check, a permutation test was conducted where the diagnostic state of the measured site was randomly assigned for a given classifier task while preserving the number of patients within each class. The results are shown in Fig. 5. The mean \pm standard deviation of the AUCs for a randomly shuffled task is shown as a gray cross-hair with error bars, while the real AUC of the top LDA model is shown as a filled black circle with a single bar indicating the extent of AUC values within the top N comparable LDA models. The randomly assigned permutations have AUCs that are clustered around 0.5 or chance performance, whereas the real AUCs are well above the error bars. The p values of the permutation tests showed significant difference between the real AUC and the randomly shuffled AUCs. For example, the permutation test for classifying normal and SD has a zero p value. These results demonstrate that the LDA model is capturing meaningful differences between the diagnostic classes as opposed to merely magnifying chance differences in the feature values.

4 Discussion

Given the variety of oral mucosa and the resulting spectral diversity that can confound classification, it is noteworthy that

Table 4 For each classification task, the frequency of appearance of individual features within the best performing statistically comparable LDA feature sets is shown.

	<i>Normal from SD</i>	<i>Normal from MD</i>	<i>Normal from MD and SD</i>	<i>MD from SD</i>
Depol ratio, x nm	0.69	Nuclear size	0.64	Nuclear size
Mean depol ratio	0.56	Perpendicular, x nm	0.32	Depol ratio, x nm
Nuclear size	0.53	Par/per, x nm	0.26	Mean par/per
Parallel, x nm	0.53	Mean depol ratio	0.25	Par/per, x nm
Mean diffuse	0.44	Parallel, x nm	0.22	Mean depol ratio
Mean perpendicular	0.44	Mean perpendicular	0.19	Mean perpendicular
Par/per, x nm	0.36	Depol ratio, x nm	0.19	Mean parallel
Perpendicular, x nm	0.31	Diffuse, x nm	0.19	Mean diffuse
Mean parallel	0.25	Mean parallel	0.18	Parallel, x nm
Diffuse, x nm	0.25	Mean par/per	0.17	Diffuse, x nm
Mean par/per	0.17	Mean diffuse	0.16	Perpendicular, x nm
	<i>Benign from SD</i>	<i>Benign from MD</i>	<i>Benign from MD and SD</i>	<i>Benign from Normal</i>
Mean perpendicular	0.54	Par/per, x nm	0.90	Mean perpendicular
Mean parallel	0.50	Parallel, x nm	0.87	Perpendicular, x nm
Parallel, x nm	0.47	Diffuse, x nm	0.87	Diffuse, x nm
Mean diffuse	0.46	Depol ratio, x nm	0.75	Parallel, x nm
Diffuse, x nm	0.46	Mean diffuse	0.66	Mean depol ratio
Perpendicular, x nm	0.46	Mean parallel	0.57	Depol ratio, x nm
Nuclear size	0.37	Mean depol ratio	0.56	Mean parallel
Depol ratio, x nm	0.37	Mean par/per	0.49	Mean diffuse
Mean depol ratio	0.34	Mean perpendicular	0.38	Nuclear size
Mean par/per	0.13	Perpendicular, x nm	0.34	Mean par/per
Par/per, x nm	0.11	Nuclear size	0.32	Depol ratio, x nm
				Nuclear size

Features identified by *x nm* correspond to the intensity at the most discriminatory wavelength. Features are sorted such that those that occur with a frequency greater than 0.5 appear above the dashed line. These features are considered to be the most diagnostically relevant.

statistical significance was obtained across diagnostic classes without the need to stratify the data according to tissue location (i.e., buccal, tongue, etc.) or tissue keratinization. The detailed interpretation of these findings as they relate to tissue morphology is somewhat difficult, although a qualitative understanding can be obtained, thus guiding future studies and probe designs.

In Fig. 3, the progression of normal tissue to premalignancy then to malignancy can be viewed in terms of the physical alteration of normal tissue from a homogeneous two-layer structure to a very irregular single layer tissue. It has been shown that changes in spectral profiles that accompany progression to carcinoma can be related to changes in both the

epithelium and the stroma such as increased microvascularization and scattering alterations.^{2,37,38}

The frequency of appearance of features in Table 4 points toward their diagnostic importance. The features that occur most consistently are nuclear size, the intensity ratio of parallel to perpendicular channels, and the mean perpendicular signal. Nuclear size appears with the highest frequency in classification tasks involving normal tissue while the mean perpendicular feature and the ratio of parallel to perpendicular feature dominate cases that involve benign tissue. Clearly nuclear size is a measure of the morphological changes that occur in superficial (i.e., epithelial) tissue. The regular appearance of nuclear size is consistent with histopathology where

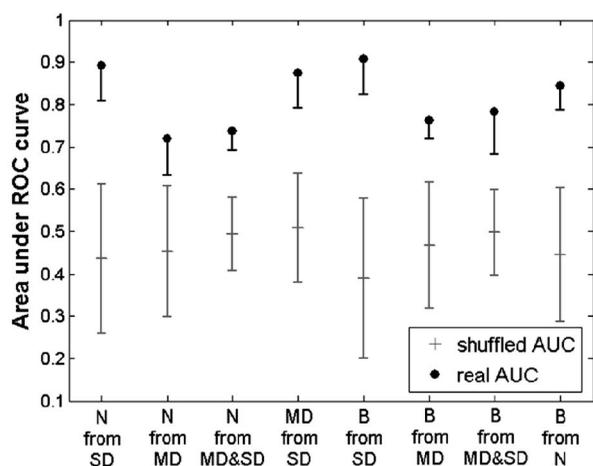


Fig. 5 Permutation test to check for overtraining. The diagnosis assignment was shuffled for each classification task while preserving the number of patients within each class. The mean and standard deviation of the area under the ROC curve, AUC, obtained using the shuffled data sets is shown as gray cross hairs. The AUC for the true LDA feature sets is shown in black. The highest LDA AUC is shown as a filled black circle, while a lower black whisker indicates the lowest statistically equivalent LDA AUC. The true LDA AUCs are well above the permuted AUCs, which are grouped around chance performance, demonstrating that the data is not overtrained.

epithelial cell nuclei are well documented to enlarge with the progression of cancer.⁴ In Fig. 2(b) both the extracted and measured nuclear size shows an increase with disease progression.

The ratio of the parallel to the perpendicular signal, on the other hand, can be interpreted as the ratio of shallow to deep tissue changes. This can be seen if the parallel/perpendicular signal is rewritten as $[(I_{\text{par}} - I_{\text{per}})/I_{\text{per}}] + 1$, where $(I_{\text{par}} - I_{\text{per}})$ represents photons that have undergone few scattering events, while I_{per} represents photons that have had many scattering interactions. The mean perpendicular feature is, therefore, a measure of the interaction of photons in tissue below the epithelium where increases in capillary density will manifest as hemoglobin modulation of the perpendicular scattering spectrum. These alterations in the perpendicular spectrum will then affect the parallel/perpendicular ratio.

Recent work in other organ sites indicates that changes in blood content and oxygenation that occur below the epithelium can be related to tumor development and, potentially, to premalignant lesion formation. Siegel et al. reported increased blood supply in subepithelial mucosa before the development of dysplasia in adenomatous human colon biopsies and rat colons treated with a carcinogen.³⁹ Zonios et al. found an increase in hemoglobin concentration in adenomatous colon polyps, but not hemoglobin oxygenation.⁴⁰ In contrast, Bard et al. found that endobroncho tumors were characterized by lower blood oxygenation.⁴¹ In another study, Fawzy et al. demonstrated that malignant lung lesions had differences in blood volume fraction and oxygen saturation when compared to normal or benign lesions.⁴² In their investigation, the blood volume fraction was significantly higher in malignant lesions than benign lesions.

These findings can be extended to the oral mucosa where 85% of all precancerous lesions have the clinical appearance

of white patches or leukoplakia.⁴³ Reviews of the prognosis of oral premalignant lesions in several countries including the USA, India, Hungary, Netherlands, and Norway by Silverman et al.⁴⁴ and Rebeil⁴⁵ has shown that the rate of malignant transformation of leukoplakia can range from less than 1% to 18%, where the highest transformation rate was found in the USA. Less frequently encountered is a red patch or erythroplakia, which is nearly always associated with dysplasia or carcinoma at the time of identification.⁵ Both erythroplakia and their mixtures with leukoplakia are at a higher risk for malignancy. The clinical description of erythroplakia as a red patch indicates an increase in blood perfusion. The implication that capillary density is correlated with precancer progression in the oral cavity is a subject of interest for future work.

Adding to the interpretation of the mean perpendicular feature is the fact that the polarization change that gives rise to the perpendicular signal can be viewed as a diffusion process where depolarization increases with increasing optical-tissue interaction.²⁴ Consequently, areas of increased scattering in superficial regions of tissue such as keratin or dysplasia will also contribute to the perpendicular signal. The oblique collection geometry of our probe makes it more sensitive to superficial tissue changes as the collection efficiency is greatest in the first 300 to 400 μm of tissue. This sensitivity enhances the effects of increased superficial scattering. It is well known that the nonuniformity of appearance of oral cavity lesions, arising from architectural and morphological changes, is correlated with transformation to invasive cancer.⁴⁶⁻⁴⁸ Work in the breast, cervix, and bronchus has also indicated that the local variation of tissue can potentially yield diagnostically relevant information.^{41,49-51} We hypothesize that it is this combination of blood absorption from deeper tissue and localized increases in epithelial scattering, resulting from changes in the local tissue morphology, that causes the perpendicular and the parallel/perpendicular features to appear with such high frequency in the top LDA models for classification tasks involving benign tissue. Further study of these high-frequency features is necessary to fully assess their physical meaning and their impact on diagnostic classification.

Another important outcome of the statistical analysis is the ability of OPRS to separate benign lesions from all other lesion types despite having the same or similar outward appearance. Many optical techniques are capable of discerning normal from malignant oral mucosa with a high degree of sensitivity and specificity, but discrimination of benign lesions from precancer and cancer is more elusive. A few groups have reported discrimination of the differing grades of abnormal human oral cavity tissue *in vivo* using imaging techniques. Wang et al. was able to separate benign from dysplastic and cancerous buccal mucosa autofluorescence (sensitivity of 81%, specificity of 96%) using a partial least squares artificial neural network analysis.⁵² Onizawa et al. were able to separate benign from cancerous oral cavity tissue with a sensitivity and specificity of 91% and 84%, and also benign from dysplasia plus cancer with a sensitivity of 94% and a specificity of 96% using UV flash photography.⁸ Kulapaditharom and Boonkitticharoen reported a sensitivity and specificity of 100% and 73%, respectively, for separation of benign from dysplastic plus malignant tissue.⁵³ Although these imaging

Table 5 Sensitivity and specificity of OPRS for all binary combinations of four distinct histological groups: normal, benign, MD, and SD.

	Normal from			MD from
	MD	MD and SD	SD	SD
Sensitivity (%)	75	73	90	80
Specificity (%)	73	64	86	83
	Benign from			Benign from
	MD	MD & SD	SD	Normal
Sensitivity (%)	92	86	100	85
Specificity (%)	69	61	85	73

studies showed good sensitivity and specificity they had certain limitations. The results by Wang et al. were limited by the isolation of their study to a single oral cavity location (buccal mucosa) and the similar history of carcinogenic exposure (areca quid chewing and smoking). The results achieved by Onizawa et al. and Kulapaditharom and Boonkitticharoen could be attributed to increased porphyrin fluorescence with dysplasia. However, Betz et al. indicated that porphyrin fluorescence is not a good indicator of disease as it was only present in one-third of tumors they studied and was also present on normal tongue and gingival plaques.¹¹

Overall spectroscopic studies have shown less optimistic results. Muller et al. found a sensitivity and specificity of 64 and 90%, respectively, for separating dysplasia from cancer using trimodal spectroscopy.¹⁶ De Veld et al. recently concluded that neither diffuse reflectance spectroscopy nor tissue autofluorescence (alone or in combination) could distinguish benign tissue from dysplastic and cancerous tissue based on a clinical trial of 134 abnormal lesions.⁵⁴ The results of our pilot clinical trial are more consistent with the cited imaging work. We have also found that the OPRS is capable of distinguishing benign tissue from normal, precancerous, and cancerous tissue with good sensitivity and specificity. Although it is more informative to report AUCs rather than a single point on the ROC curve, Table 5 lists the sensitivity and specificity of OPRS for all diagnostic classification tasks for easy comparison to work by others. In each case, the ROC threshold was chosen such that it gave both high sensitivity and high specificity. We suggest that multiple diagnostically relevant features that can be extracted from a single OPRS measurement are the key to its diagnostic performance.

Separation of normal tissue from SD yielded a sensitivity of 90% and a specificity of 86%. Discrimination of identical looking benign lesions from severe dysplasia was also achieved with high sensitivity (100%) and specificity (85%). Evaluation of all dysplasia (including carcinoma) against normal or benign tissue yielded sensitivities and specificities of 73 and 64% and 86 and 61%, respectively. Lesser diagnostic grades such as MD and benign lesions do not require surgical resection, but must be monitored for possible transformation

to malignancy. OPRS was able to discriminate MD from SD with a sensitivity of 80% and specificity of 83%. These compelling findings emphasize the need for a large prospective study to assess OPRS as an adjunct to clinical examination for the detection and monitoring of precancerous and cancerous tissue.

5 Conclusion

The oral cavity has a variety of tissue architectures in addition to a whole host of benign conditions such as leukoplakia, erythroplakia, and lichen planus that can mask precancer and cancer. Using a multipronged approach to discriminate the earliest stages of precancer could solve this problem. An ideal pared down system would be simple, low-cost, robust, and noninvasive. We believe that OPRS has the potential to fulfill these requirements. Our results from a preliminary pilot clinical trial have demonstrated the ability of OPRS to discriminate, with high sensitivity and specificity, normal tissue from high-grade dysplasia and cancer (SD). Further, OPRS can discriminate visually identical lesions such as benign from SD and benign from premalignant and malignant lesions. These promising results suggest that OPRS has the potential to augment current clinical practice for diagnosis and monitoring of oral premalignancies and malignancies.

Acknowledgments

Financial support from the Whitaker Foundation and the National Institute of Biomedical Imaging and BioEngineering (NIBIB) EB003540 is gratefully acknowledged.

References

1. R. Richards-Kortum and E. Sevick-Muraca, "Quantitative optical spectroscopy for tissue diagnosis," *Annu. Rev. Biochem.* **47**, 555–606 (1996).
2. I. J. Bigio and J. R. Mourant, "Ultraviolet and visible spectroscopies for tissue diagnostics: fluorescence spectroscopy and elastic-scattering spectroscopy," *Phys. Med. Biol.* **42**, 803–814 (1997).
3. K. Sokolov, M. Follen, and R. Richards-Kortum, "Optical spectroscopy for detection of neoplasia," *Curr. Opin. Chem. Biol.* **6**(5), 651–658 (2002).
4. R. Cotran, V. Kumar, and S. Robbins, *Pathologic Basis of Disease* (1994).
5. J. J. Sciubba, "Oral cancer: the importance of early diagnosis and treatment," *Am. J. Clin. Dermatol.* **2**(4), 239–251 (2001).
6. A. Jemal, R. Siegel, E. Ward, T. Murray, J. Xu, and M. J. Thun, "Cancer statistics, 2007," *Ca-Cancer J. Clin.* **57**, 43–66 (2007).
7. "SEER Statistics Review 1975–2003, Oral Cavity and Pharynx," in *Surveillance Epidemiology and End Results*, National Cancer Institute.
8. K. Onizawa, H. Saginoya, Y. Furuya, H. Yoshida, and H. Fukuda, "Usefulness of fluorescence photography for diagnosis of oral cancer," *Int. J. Oral Maxillofac Surg.* **28**(3), 206–210 (1999).
9. B. Kulapaditharom and V. Boonkitticharoen, "Laser-induced fluorescence imaging in localization of head and neck cancers," *Adv. Oto-Rhino-Laryngol.* **107**, 241–246 (1998).
10. E. Svistun, R. Alizadeh-Naderi, A. El-Naggar, R. Jacob, A. Gillenwater, and R. Richards-Kortum, "Vision enhancement system for detection of oral cavity neoplasia based on autofluorescence," *Head Neck* **26**, 205–215 (2004).
11. C. S. Betz, M. Mehlmann, K. Rick, H. Stepp, G. Grevers, R. Baumgartner, and A. Leunig, "Autofluorescence imaging and spectroscopy of normal and malignant mucosa in patients with head and neck cancer," *Lasers Surg. Med.* **25**, 323–334 (1999).
12. A. L. Clark, A. Gillenwater, R. Alizadeh-Naderi, A. K. El-Naggar, and R. Richards-Kortum, "Detection and diagnosis of oral neoplasia with an optical coherence microscope," *J. Biomed. Opt.* **9**(6), 1271–1280 (2004).

13. P. Wilder-Smith, T. Krasieva, W. G. Jung, J. Zhang, Z. Chen, K. Osann, and B. Tromberg, "Noninvasive imaging of oral premalignancy and malignancy," *J. Biomed. Opt.* **10**, 051601 (2005).
14. P. M. Lane, T. Gilhuly, P. Whitehead, Z. Haishan, C. F. Poh, S. Ng, P. M. Williams, Z. Lewei, M. P. Rosin, and C. E. MacAulay, "Simple device for the direct visualization of oral-cavity tissue fluorescence," *J. Biomed. Opt.* **11**(2), 24006–24241 (2006).
15. 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**, 1251–1258 (1998).
16. M. G. Muller, T. A. Valdez, I. Georgakoudi, V. Backman, C. Fuentes, S. Kabani, N. Laver, Z. 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* **97**(7), 1681–1692 (2003).
17. D. C. G. d. 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).
18. S. P. Schantz, V. Kolli, H. E. Savage, G. Yu, J. P. Shah, D. E. Harris, A. Katz, R. R. Alfano, and A. G. Huvos, "In vivo native cellular fluorescence and histological characteristics of head and neck cancer," *Clin. Cancer Res.* **4**(5), 1177–1182 (1998).
19. J. K. Dhingra, D. F. Perrault, Jr., 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).
20. 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).
21. A. Gillenwater, V. Papadimitrakopoulou, and R. Richards-Kortum, "Oral premalignancy: new methods of detection and treatment," *Curr. Oncol. Rep.* **8**(2), 146–154 (2006).
22. D. C. G. De Veld, M. J. H. Witjes, H. J. C. M. Sterenborg, and J. L. N. Roodenburg, "The status of in vivo autofluorescence spectroscopy and imaging for oral oncology," *Oral Oncol.* **41**(2), 117–131 (2005).
23. S. G. Demos and R. R. Alfano, "Temporal gating in highly scattering media by the degree of optical polarization," *Opt. Lett.* **21**(2), 161–163 (1996).
24. S. L. Jacques, J. R. Roman, and K. Lee, "Imaging superficial tissues with polarized light," *Lasers Surg. Med.* **26**, 119–129 (2000).
25. R. R. Anderson, "Polarized light examination and photography of the skin," *Arch. Dermatol.* **127**(7), 1000–1005 (1991).
26. V. Backman, R. Gurjar, K. Badizadegan, I. Itzkan, R. R. Dasari, L. T. Perelman, and M. S. Feld, "Polarized light scattering spectroscopy for quantitative measurement of epithelial cellular structures in situ," *IEEE J. Sel. Top. Quantum Electron.* **5**(4), 1019–1026 (1999).
27. K. Sokolov, R. Drezek, K. Gossage, and R. Richards-Kortum, "Reflectance spectroscopy with polarized light: is it sensitive to cellular and nuclear morphology," *Opt. Express* **5**(13), 302–317 (1999).
28. L. Nieman, A. Myakov, J. Aaron, and K. Sokolov, "Optical sectioning using a fiber probe with an angled illumination-collection geometry: evaluation in engineered tissue phantoms," *Appl. Opt.* **43**(6), 1308–1319 (2004).
29. K. Sokolov, L. T. Nieman, A. Myakov, and A. Gillenwater, "Polarized reflectance spectroscopy for precancer detection," *Technol. Cancer Res. Treat.* **3**(1), 1–14 (2004).
30. "Threshold limit values and biological exposure indices," American Conference of Governmental Industrial Hygienists (1996).
31. J. A. Hanley, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology* **143**(1), 29–36 (1982).
32. A. Myakov, L. Nieman, L. Wicky, U. Utzinger, R. Richards-Kortum, and K. Sokolov, "Fiber optic probe for polarized reflectance spectroscopy in vivo: design and performance," *J. Biomed. Opt.* **7**(3), 388–397 (2002).
33. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York (1973).
34. B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall, New York (1993).
35. N. Ancona, R. Maglietta, A. Piepoli, A. D'Addabbo, R. Cotugno, M. Savino, S. Liuni, M. Carella, G. Pesole, and F. Perri, "On the statistical assessment of classifiers using DNA microarray data," *BMC Bioinf.* **7**(1), 387 (2006).
36. A. Nath, K. Rivoire, C. Sung, D. Cox, E. N. Atkinson, M. Follen, and R. Richards-Kortum, "Effect of probe pressure on cervical fluorescence spectroscopy measurements," *J. Biomed. Opt.* **9**(3), 523–533 (2004).
37. R. K. Jain, "Antiangiogenic therapy for cancer: current and emerging concepts," *Oncology* **19**(4 Suppl 3), 7–16 (2005).
38. R. Drezek, M. Guillaud, T. Collier, I. Boiko, A. Malpica, C. MacAulay, M. Follen, and R. Richards-Kortum, "Light scattering from cervical cells throughout neoplastic progression: influence of nuclear morphology, DNA content, and chromatin texture," *J. Biomed. Opt.* **8**(1), 7–16 (2003).
39. M. P. Siegel, Y. L. Kim, H. K. Roy, R. K. Wali, and V. Backman, "Assessment of blood supply in superficial tissue by polarization-gated elastic light-scattering spectroscopy," *Appl. Opt.* **45**(2), 335–342 (2006).
40. G. Zonios, L. T. Perelman, V. 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).
41. M. P. L. Bard, A. Amelink, M. Skurichina, M. d. Bakker, S. A. Burgers, J. P. v. Meerbeeck, R. P. W. Duin, J. G. J. V. Aerts, H. C. Hoogsteden, and H. J. C. M. Sterenborg, "Improving the specificity of fluorescence bronchoscopy for the analysis of neoplastic lesions of the bronchial tree by combination with optical spectroscopy: preliminary communication," *Lung Cancer* **47**, 41–47 (2005).
42. Y. S. Fawzy, M. Petek, M. Terelj, and H. Zeng, "In vivo assessment and evaluation of lung tissue morphologic and physiological changes from non-contact endoscopic reflectance spectroscopy for improving lung cancer detection," *J. Biomed. Opt.* **11**(4), 044003 (2006).
43. J. J. Sciubba, "Oral leukoplakia," *Crit. Rev. Oral Biol. Med.* **6**(2), 147–160 (1995).
44. S. Silverman, Jr., M. Gorsky, and F. Lozada, "Oral leukoplakia and malignant transformation. A follow-up study of 257 patients," *Cancer* **53**(3), 563–568 (1984).
45. J. Reibel, "Prognosis of oral pre-malignant lesions: significance of clinical, histopathological, and molecular biological characteristics," *Crit. Rev. Oral Biol. Med.* **14**(1), 47–62 (2003).
46. P. Holmstrup, P. Vedtofte, J. Reibel, and K. Stoltze, "Long-term treatment outcome of oral premalignant lesions," *Oral Oncol.* **42**(5), 461–474 (2006).
47. L. Zhang, M. Williams, C. F. Poh, D. Laronde, J. B. Epstein, S. Durham, H. Nakamura, K. Berean, A. Hovan, N. D. Le, G. Hislop, R. Priddy, J. Hay, W. L. Lan, and M. P. Rosin, "Toluidine blue staining identifies high-risk primary oral premalignant lesions with poor outcome," *Cancer Res.* **65**(17), 8017–8021 (2005).
48. S. Silverman and P. B. Sugerman, "Oral premalignancies and squamous cell carcinoma," *Clin. Dermatol.* **18**(5), 563–568 (2000).
49. A. Amelink and H. J. C. M. Sterenborg, "Measurement of the local optical properties of turbid media by differential path-length spectroscopy," *Appl. Opt.* **43**(15), 3048–3054 (2004).
50. R. L. P. v. Veen, A. Amelink, M. Menke-Pluymers, C. V. D. Pol, and H. J. C. M. Sterenborg, "Optical biopsy of breast tissue using differential path-length spectroscopy," *Phys. Med. Biol.* **50**, 2573–2581 (2005).
51. I. Georgakoudi, E. E. Sheets, M. G. Muller, V. Backman, C. P. Crum, K. Badizadegan, R. R. Dasari, and M. S. Feld, "Trimodal spectroscopy for the detection and characterization of cervical precancers in vivo," *Am. J. Obstet. Gynecol.* **186**(3), 374–382 (2002).
52. C.-Y. Wang, T. Tsai, H.-M. Chen, C.-T. Chen, and C.-P. Chiang, "PLS-ANN based classification model for oral submucous fibrosis and oral carcinogenesis," *Lasers Surg. Med.* **32**(4), 318–326 (2003).
53. B. Kulapaditharom and V. Boonkitticharoen, "Performance characteristics of fluorescence endoscope in detection of head and neck cancers," *Ann. Otol. Rhinol. Laryngol.* **110**(1), 45–52 (2001).
54. D. C. G. d. Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, H. J. C. M. Sterenborg, and J. L. N. Roodenburg, "Autofluorescence and diffuse reflectance spectroscopy for oral oncology," *Lasers Surg. Med.* **36**(5), 356–364 (2005).