Spectroscopic diagnosis and imaging of invisible pre-cancer

Kamran Badizadegan,[a] Vadim Backman,[b] Charles W. Boone,[c] Christopher P. Crum,[d] Ramachandra R. Dasari,[c] Irene Georgakoudi,[c] Kristin Keefe,[e] Karl Munger,[f] Stanley M. Shapshay,[g] Ellen E. Sheets[e] and Michael S. Feld*[c]

[a] *Department of Pathology, Children's Hospital, Boston, MA, USA*
[b] *Biomedical Engineering Department, Northwestern University, Evanston, IL, USA*
[c] *MIT Laser Biomedical Research Center, GR Harrison Spectroscopy Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: msfeld@mit.edu*
[d] *Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA*
[e] *Division of Gynecologic Oncology, Brigham and Women's Hospital, Boston, MA, USA*
[f] *Department of Pathology, Harvard Medical School, Boston, MA, USA*
[g] *Department of Otolaryngology, Boston University School of Medicine, Boston, MA, USA*

Received 15th May 2003, Accepted 3rd July 2003
First published as an Advance Article on the web 23rd September 2003

The theme of this paper is the use of optical spectroscopy to diagnose invisible pre-cancer in patients undergoing endoscopy and similar medical procedures. We describe three techniques that provide diagnostic information and two instruments to implement them, the FastEEM for studying small regions of tissue and the LSS (light scattering spectroscopy) imaging system for wide-area surveillance. The FastEEM is an optical fiber clinical device that collects spectra of reflected light and fluorescence at multiple excitation wavelengths from the tissue, all in a fraction of a second. Quantitative information is obtained in real time, without removing the tissue and without the need for staining and fixation. Three types of spectral information are extracted—intrinsic fluorescence, diffuse reflectance and elastic light scattering. Each of the three analyses is based on a biophysical model, and each provides complementary quantitative physical and chemical information about cellular/tissue structures. This information is used to make a combined spectral diagnosis, a method we call tri-modal spectroscopy (TMS). Promising clinical studies are being carried out on patients undergoing routine pre-cancer surveillance in the oral cavity, the uterine cervix and the gastrointestinal tract. The LSS imaging system provides wide-area spectroscopic images of the epithelium, typically 2 cm in each dimension, depicting the size distribution and chromatin content of the cell nuclei, which are key parameters in diagnosing pre-cancer. This instrument is in preclinical stages of development, although a laboratory prototype has been used to create diagnostic images in resected colon polyp samples. The combination of the TMS/FastEEM and LSS imaging instrument will constitute a powerful new diagnostic tool, with LSS imaging to provide wide area surveillance and the TMS probe to provide detailed information on suspect tissue sites.

I. Introduction

It is well established that cancer can be most effectively treated when diagnosed at an early stage. Therefore, development, evaluation and validation of new tools for early detection of cancer and pre-cancer is an important priority. Most cancers (85–90%) originate in the cell-rich epithelial tissues that line the internal and external surfaces of the body. Over the past few years, our laboratory has been developing novel spectroscopic and imaging tools to diagnose and monitor the progression of pre-cancerous and early cancerous epithelial lesions (the pre-cancerous state is broadly known as *dysplasia* or *in-situ epithelial neoplasia*.) This new technology is based on fundamental physical principles, and has been successfully employed in a number of organs in patients undergoing routine diagnostic procedures. In this paper, we address three of these techniques: light scattering spectroscopy (LSS), for studying the size distribution and chromatin content of epithelial cell nuclei; intrinsic fluorescence spectroscopy (IFS), for monitoring of changes in tissue metabolism and biochemistry; and diffuse reflectance spectroscopy (DRS), for characterizing tissue scattering and oxygenation. All of these provide quantitative information about the state of the tissue, without the need for tissue removal.

Two instruments have been developed to implement these new techniques, the FastEEM for studying small regions of tissue and the LSS imaging system for wide-area surveillance. The FastEEM[1] samples a 1 mm^3 tissue volume *via* an optical fiber probe, collecting fluorescence excitation–emission matrices (EEMs) and white light reflectance in a fraction of a second (hence, FastEEM). (An EEM displays fluorescence emission spectra as a function of excitation wavelength, a 3-D representation that fully characterizes the fluorescence properties of a tissue sample.) Using these data, the techniques of IFS, DRS and LSS can be implemented to provide the three kinds of biological information mentioned above, which are then combined to obtain a clinical diagnosis, a method called tri-modal spectroscopy (TMS). Using the FastEEM, we have successfully employed TMS to study dysplasia in a number of organs, including dysplasia in Barrett's esophagus and cervical neoplasia at Boston's Brigham and Women's Hospital (BWH),[2,3] and oral dysplasia at Boston University Medical Center (BUMC).[4]

The second instrument, the LSS imaging system, is in an earlier stage of development. In this device, polarized light at a succession of wavelengths illuminates an epithelial surface, and the polarized back-scattered light is used to construct spectroscopic images of the epithelium, typically 2 cm in each dimension, depicting the size distribution and chromatin content of the cell nuclei over a wide area. Increased nuclear size, density (crowding), standard deviation (pleomorphism) and chromatin content (hyperchromasia) are key features used by pathologists to make a diagnosis of dysplasia. Wide area imaging is important because it provides an overview of the entire surface of an organ, highlighting suspicious areas for more detailed study. Dysplasia is often invisible or difficult to detect under gross clinical examination and, at present, final diagnosis of suspicious sites requires tissue removal, fixation, staining and microscopic examination by a trained pathologist, which is time consuming, labor intensive and qualitative. Current wide area surveillance is either based on random biopsy or tissue sampling based on qualitative changes in color, both of which are greatly prone to sampling error. The LSS imaging device is therefore of significant potential diagnostic utility. We have developed a laboratory prototype LSS imaging system, and demonstrated its ability to create diagnostic images of dysplasia in resected colon adenoma samples[5] and are now constructing a prototype system for clinical use. We have chosen the oral cavity and uterine cervix as model systems, because they share many histopathological features and can be directly viewed without the complexities of an endoscope, a considerable simplification in this stage of developing the technology.

This paper discusses the scientific basis of our spectral diagnosis techniques, the instrumentation used to collect data, our preliminary clinical results, and our plans to develop a integrated instrument. The combination of the TMS/FastEEM contact probe and LSS imaging in an integrated instrument will constitute a powerful new tool for diagnosing dysplasia, with LSS imaging to provide wide area surveillance and the TMS probe to provide detailed information on suspect tissue sites. This instrument will provide objective, quantitative diagnostic information in real time, without the need for tissue removal. In addition, the spectroscopic imaging data obtained may be stored electronically and downloaded at a later time to be compared with more recent data obtained during follow-up visits to the clinic by the patient. Diagnostic spectroscopic images may

thus be used prospectively to follow patients during and post-therapy to detect very early recurrence of neoplastic lesions.

II. Optical spectroscopy as a diagnostic tool

Optical spectroscopy is a potentially important tool for tissue diagnosis. Spectroscopic measurements can be implemented *in vivo*, thus providing information about tissue in its native state, free of artifacts introduced by tissue excision and processing. Different spectroscopic techniques can be used to provide a variety of information about tissue morphology and biochemistry. The capability to obtain *in vivo* information about specific biochemical/morphological changes that take place during the development or regression of neoplasia can provide a rich source of diagnostic information, and can further the understanding of the biological processes involved, as well.

Light propagating in biological tissue can undergo a variety of interactions. Elastic scattering (*i.e.*, scattering without wavelength change) is the predominant mechanism, and light impinging on the tissue can be scattered once (single scattering) or multiple times (diffusive scattering) before returning to the surface to be detected. Light can also be absorbed by chromophores such as hemoglobin without being re-emitted (absorption), or by fluorophores such as NAD(P)H and re-emitted at longer wavelengths (fluorescence) before being detected. In addition, inelastic processes that shift wavelength (such as Raman scattering) can occur. The physical processes relevant to the work described below are fluorescence, diffuse scattering and single scattering.

Fluorescence

Fluorescence can provide information about the biochemical state of the tissue and the changes that occur during disease development. Promising *in vivo* results to diagnose cancer have been reported in a number of tissues, including the cervix,[6] the lung,[7,8] the gastrointestinal tract,[9–12] the oral cavity,[13,14] the skin[15] and the bladder.[16,17] These results have employed statistical or empirical analyses of the observed tissue fluorescence, the spectral features of which can be significantly distorted by the interplay of absorption and diffusive scattering, which is ubiquitous in biological tissue, limiting the ability to extract quantitative biochemical information. We have developed the technique of intrinsic fluorescence spectroscopy to disentangle these artifacts and extract the undistorted (*intrinsic*) tissue fluorescence.[18–20] As mentioned above, IFS provides an accurate method for characterizing the biochemical changes in tissue fluorescence associated with dysplasia and understanding their origins.[2–4,21]

Diffusive scattering

Light propagating in tissue is rapidly diffused due to multiple scattering. Diffusely reflected (or transmitted) light can be used to detect changes in the scattering and absorption properties of tissue. Diagnostic applications study the spectrum of diffusive back-scattered CW white light delivered and collected by means of a small ($\sim$1 mm diameter) optical fiber probe. Since this source–detector configuration samples the tissue to a depth of $\sim$1 mm or less, such techniques predominantly probe the subepithelial region of tissue, and are well suited for studying stromal changes in epithelial lesions. Promising clinical results have been reported for distinguishing normal and pre-cancerous tissues in the skin,[22] the breast,[23] the gastrointestinal tract[24,25] and the bladder.[25,26] The above studies have employed empirical methods of spectral analysis, correlating features in the diffuse reflectance spectrum with disease state. We have developed an analytical method for analyzing diffuse reflectance spectra, based on diffusion theory, in which information about the tissue scatterers and absorbers is extracted from the diffuse reflectance spectrum.[27] This method, diffuse reflectance spectroscopy (DRS), provides quantitative information about tissue composition, such as hemoglobin concentration, collagen matrix composition, *etc.*, from which diagnostic information is obtained.

Single scattering

As shown by our group, the back-scattered light also contains spectral information about single scattering from the epithelium, which we study *via* light scattering spectroscopy.[28,29] This spectral

component provides quantitative information about the size distribution and refractive index of scatterers in the epithelium. It comprises only a small fraction (2–5%) of the overall back-reflected light, and special methods are required to extract it. LSS spectra collected over small back-scattering angles (a few degrees) contain information about particles that are large compared to a wavelength of light, and analysis of such spectra provides information about the size distribution and chromatin content of epithelial cell nuclei which, as mentioned above, are hallmarks of dysplasia. Successful clinical results have been obtained in the oral cavity and the uterine cervix, as well as the gastrointestinal tract and the urinary bladder.[2–4,30,31] In addition, our recent studies indicate that light scattering spectroscopic measurements at large and small angles can provide information about sub-nuclear structure and overall cellular organization.[32]

Tri-modal spectroscopy (TMS)

Analysis of fluorescence EEMs and reflectance spectra has been performed to assess the potential of IFS, DRS and LSS as diagnostic tools for detecting pre-cancerous and cancerous lesions in Barrett's esophagus, the oral cavity and the uterine cervix. Successful diagnostic results were obtained in all three organs. These studies provide evidence that the three techniques provide accurate diagnostic information, and that tri-modal spectroscopy, the combined use of IFS, DRS and LSS, results in superior sensitivity and specificity compared to any one of the techniques alone. Our work with Barrett's esophagus patients was recently published.[2] In this section, we summarize more recent results in the cervix[3] and the oral cavity.[4]

III. Experimental methods

Oral cavity study protocol

Reflectance and fluorescence spectra at multiple excitation and emission wavelengths were collected from 53 tissue sites in 15 patients with known upper aerodigestive tract malignancies, and from 38 tissue sites in 8 healthy volunteers. Five sets of fluorescence and reflectance spectra were acquired in less than five seconds by bringing an optical fiber probe in gentle contact with the tissue. Immediately following spectral acquisition, a biopsy was taken from the same tissue site. Either a temporary pressure mark left by the probe on the tissue or an ink tattoo was used to ensure that the biopsied and spectroscopically examined tissue sites coincide. The biopsied tissue was examined and classified by an experienced oral pathologist.

Uterine cervix study protocol

Data were collected from 44 patients undergoing colposcopy following an abnormal Papanicolaou smear. The cervix was examined under colposcopic vision at 15× magnification. Following application of 3% acetic acid to the cervix, spectra were acquired from one or two colposcopically normal ectocervical sites located outside of the transformation zone and all colposcopically abnormal sites. Immediately following spectral acquisition, the suspicious sites were biopsied. The biopsy samples were examined and classified by an experienced pathologist, and the classification was then correlated with the results of spectroscopy. Reflectance and fluorescence spectra were collected from 50 colposcopically normal squamous ectocervical tissue sites (NSE), which were not biopsied, 5 biopsied sites classified histopathologically as benign or mature squamous epithelium (MSE), 16 biopsied sites classified as squamous metaplasia (SQM), and 13 biopsied sites classified as squamous intraepithelial lesions (SIL), two of which were low-grade (LSIL) and the remaining eleven were high-grade (HSIL). The biopsied MSE and SQM sites are referred to collectively as biopsied non-SILs.

Spectral data acquisition

The FastEEM[1] was used to collect real-time *in vivo* EEMs using 11 laser excitation wavelengths between 337 and 610 nm, and diffuse reflectance spectra over the range 350–700 nm (Fig. 1). Laser excitation light is generated using a 337 nm pulsed nitrogen laser to pump 10 cells containing

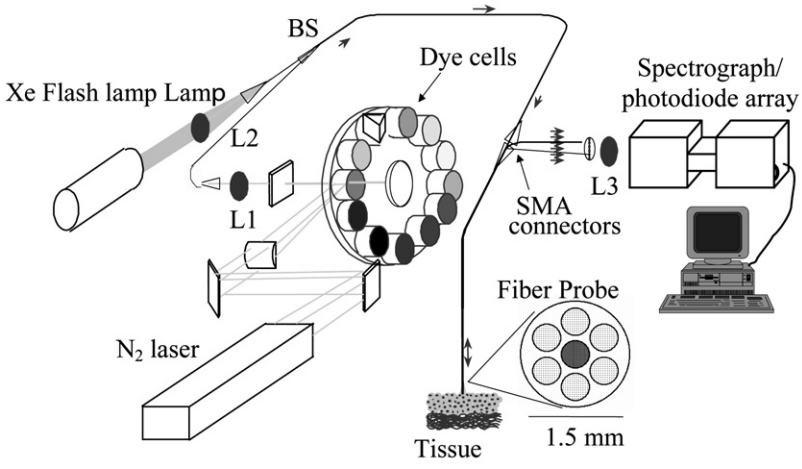

Fig. 1 FastEEM, schematic diagram.

different dyes, mounted on a rotating wheel. A signal to trigger the N_2 laser is generated whenever the dye wheel reaches an appropriate position. The diffuse reflectance spectra are acquired using a broadband white light pulse from a xenon flash lamp. The resulting 12 pulses of light are coupled into a specially designed and fabricated 1.5 mm diameter optical fiber probe, consisting of a single light delivery fiber surrounded by six collection fibers (200 μm core diameter, 0.22 NA), capped at the distal tip by an optical shield (1 mm long fused silica spacer). This configuration provides a reproducible light delivery–collection geometry when the tip is brought into contact with the tissue.[1] Reflected light and tissue autofluorescence are collected by the probe, coupled to a CP 200 spectrograph with *f*-matched optics, and detected with an intensified photodiode array. Data is recorded with OMAvision (EG&G) software. Data acquisition and storage are computer controlled. The instrument is built on a 4 foot high cart that can be moved to the various locations within a hospital. Detector gating enables data to be collected in room light, important for working in a hospital environment. Clinical results using this device in the uterine cervix and the oral cavity are presented below.

Spectral data analysis

The fluorescence and reflectance spectra acquired from the oral cavity and the uterine cervix were analyzed to provide information from intrinsic fluorescence (IFS), diffuse reflectance (DRS) and single scattering (LSS).

IFS. Light propagating in tissue undergoes multiple scattering and absorption, which can severely distort the intensity and lineshape of measured tissue fluorescence spectra. To extract the undistorted (intrinsic) tissue fluorescence, the measured fluorescence spectra were analyzed in combination with information from the corresponding reflectance spectra using a photon migration-based model.[18,19,20] The underlying principle of this model is that the fluorescence and reflectance spectra collected from a specific site using the same light delivery/collection geometry undergo similar distortions. In contrast to the observed fluorescence spectra, which consist of nonlinear contributions from tissue fluorescence, scattering and absorption, intrinsic fluorescence spectra can be described by a linear combination of the spectra corresponding to the biochemicals whose fluorescence is excited at a particular wavelength. By extracting the intrinsic tissue fluorescence, changes in tissue biochemistry can be identified more accurately, providing quantitative information about biochemical tissue composition. As shown previously,[21] at wavelengths of interest collagen and NAD(P)H are the two major contributing fluorophores. Thus, we fit a linear

combination of the fluorescence spectra of collagen and NAD(P)H to the tissue intrinsic fluorescence EEMs using a linear-least squares fitting algorithm.

DRS. The measured reflectance spectra were analyzed using a model based on diffusion theory that expresses the reflected light as a function of the wavelength-dependent absorption (μ_a) and reduced scattering (μ_s') coefficients of the tissue sampled.[27] This analysis provides information about the architecture and morphology of mainly the connective tissue, *i.e.*, the lamina propria and the submucosa, as the optical fiber probe preferentially samples tissue within 500–700 μm of the surface. A straight line was fit to the extracted wavelength-dependent values for μ_s'. Its slope and intercept were used as diagnostic parameters.

LSS. A small fraction (2–5%) of the back-reflected signal is due to light that is singly backscattered by the cell nuclei of the epithelium.[28,29] This light scattering spectrum is extracted from the reflectance data by subtracting the diffuse component provided by the model of ref. 27. The intensity of the LSS spectrum varies in wavelength in an oscillatory manner. The frequency and depth of these oscillations depend on the size and number density of the scatterers (cell nuclei). These variations were analyzed using a model based on the theory of light scattering to determine the number and size of the epithelial cell nuclei.[28,29]

Statistical analysis

Logistic regression and "leave-one-out" cross-validation were employed to determine and validate the diagnostic potential of the quantitative parameters extracted from the IFS, DRS and LSS spectra. For IFS, the relative NAD(P)H and collagen fluorescence contributions to the overall intrinsic tissue fluorescence were used as diagnostic parameters; for DRS, the slope and intercept of the line fit to $\mu_s'(\lambda)$ were used; for LSS, the number density, percent enlargement (*i.e.*, percentage of nuclei greater than 10 μm) and size standard deviation of epithelial cell nuclei extracted from the analysis were used. To evaluate the sensitivity and specificity of each technique to separate dysplastic from non-dysplastic lesions in an unbiased manner, we employed "leave-one-out" cross-validation.[33] Sensitivity and specificity values were determined by comparing the resulting spectroscopic classifications with those of histopathology.

Tri-modal spectroscopy (TMS)

IFS, DRS and LSS provide complementary information about each tissue site from spectra taken at the same time. We combined the diagnostic results from the three techniques to determine the extent to which this would improve diagnostic performance. This was done by assigning each site to the diagnostic category for which at least two of the three spectroscopic techniques agreed. The sensitivity and specificity of TMS was then determined by comparing the spectroscopic and histopathological classifications.

IV. Results

IFS

Fig. 2(a) shows a typical fluorescence spectrum from a normal oral gingival epithelium site excited with 337 nm light (thick curve). The two peaks might be attributed to the presence of two fluorophores. However, note that the decrease between these peaks occurs at the hemoglobin Soret absorption peaks (Hb, 435 nm, HbO_2, 415 nm). The effects of hemoglobin absorption are clearly present in the corresponding reflectance spectrum (Fig. 2(b)). Similar results were obtained from the cervix data.[21] An example of intrinsic (undistorted) fluorescence extracted from the measured fluorescence and reflectance, is shown in the thin curve of Fig. 2(c). Note that this spectrum exhibits a single (asymmetric) peak. In studying spectra from different areas of the oral cavity, we noted that keratinized epithelia (*e.g.*, the hard palate) and non-keratinized epithelia (*e.g.*, the buccal mucosa) exhibit distinct intrinsic fluorescence features. The development of different algorithms for these

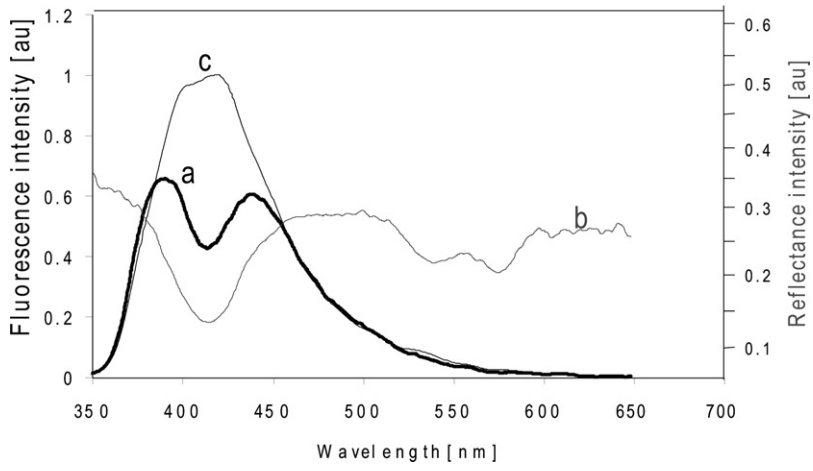

Fig. 2 Oral gingival fluorescence excited at 337 nm. (a) Observed and (c) intrinsic fluorescence; (b) reflectance.

two types of epithelia resulted in better sensitivity and specificity for separating pre-cancerous and cancerous lesions from normal tissues.

Significant differences are observed in the intrinsic tissue fluorescence of normal, squamous metaplastic (SQM) and squamous intraepithelial lesion (SIL) sites of the uterine cervix,[21] and of normal, dysplastic and cancerous sites of the oral cavity (Fig. 3). At 337 nm excitation, the lineshape of the dysplastic and cancerous sites broadens and shifts to the red region of the spectrum compared to the normal tissue sites. The intrinsic fluorescence intensity of dysplastic and cancerous tissue sites is generally lower than that of normal sites. Similar intensity trends are observed at longer excitation wavelengths, but with increasing excitation wavelength the lineshape differences gradually disappear.

To extract quantitative biochemical information about tissue composition, the intrinsic tissue fluorescence EEMs were decomposed into a linear sum of contributions due to NAD(P)H and collagen, the two major tissue fluorophores in the visible range of the spectrum.[21] The contributions of these fluorophores to the intrinsic tissue fluorescence of oral cavity tissue sites are shown in

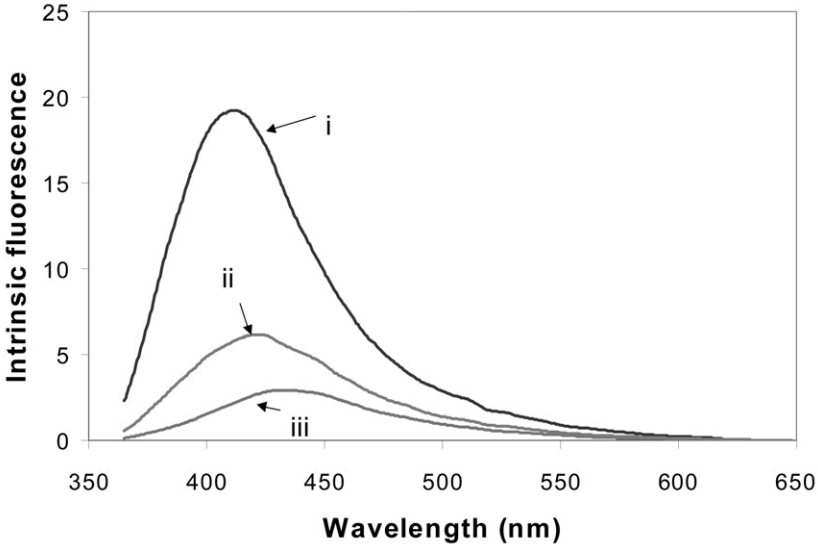

Fig. 3 Intrinsic fluorescence spectra, oral tissue excited at 337 nm. Normal (i), dysplastic (ii), cancer (iii).

Fig. 4. A number of interesting results arise from this analysis. In the oral cavity, collagen fluorescence decreases gradually from normal to dysplastic to cancerous tissue sites (Fig. 4). In the cervix, a significant decrease is also observed in collagen fluorescence of the SQM and SIL sites compared to that of the normal squamous epithelium. These changes could be the result of differences in the levels of expression of enzymes such as matrix metalloproteinases (MMPs), a class of enzymes responsible for stromal matrix degradation. Differences in the levels and/or patterns of expression of MMP-2 and other proteases involved in carcinogenesis have been reported between normal squamous epithelium, squamous metaplasia and SILs.[34] Additionally, an increase in the NADH fluorescence is noted for the cancerous/dysplastic oral tissue sites as compared to the normal sites (Fig. 4), and in the cervical SILs compared to that of the SQM sites. This increase could result from an increased number of epithelial cells and/or their metabolic activity.[35] We find that the diagnostic thresholds are slightly different for keratinized and non-keratinized oral cavity epithelial tissues.

DRS

As discussed above, the reflectance spectra can be analyzed using a mathematical model based on the diffusion approximation of light propagation in tissue.[27] A typical reflectance spectrum is shown in Fig. 2(b). The model provides excellent fits, from which the wavelength dependent scattering parameter, μ_s', and the hemoglobin concentration and saturation can be obtained. The extracted curves of $\mu_s'(\lambda)$, modeled using Mie theory, are fit to straight lines. Fig. 5 displays the slopes and intercepts at λ nm of these lines for the non-keratinized tissue sites. A trend towards smaller and flatter $\mu_s'(\lambda)$ is observed, particularly between the normal and dysplastic/cancerous oral cavity tissue sites and the normal and SIL cervical sites. The diagnostic thresholds (decision lines) based on logistic regression analysis are shown. Differences in the scattering properties of keratinized and non-keratinized oral cavity tissue sites were also observed in the extracted values of μ_s'. They may be related to the scattering properties of keratin and the underlying lamina propria.

The progressive decrease in μ_s' during the development of dysplasia/cancer could be due to the decrease in the volume of connective tissue sampled by the probe, which may decrease as the number of epithelial cells increases. As a result, the density of scatterers within the probed tissue

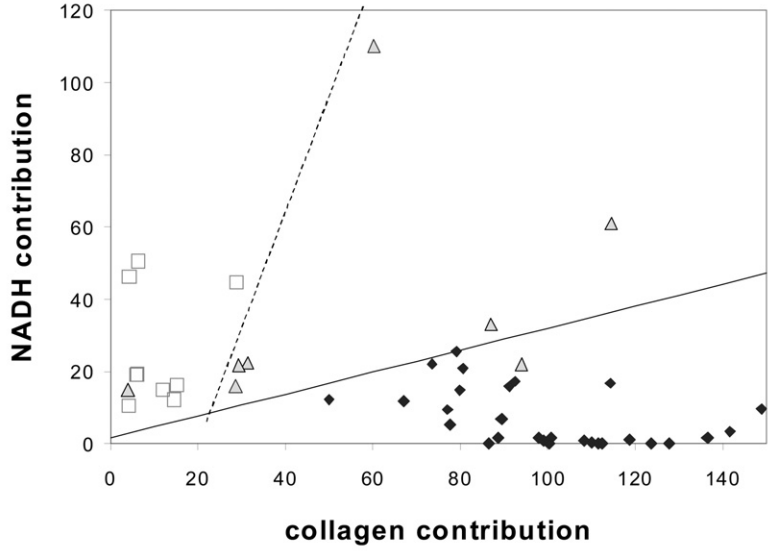

Fig. 4 NADH and collagen contributions to the intrinsic fluorescence from non-keratinized oral tissue. Normal (diamonds); dysplastic (triangles); cancerous (squares). The normal/abnormal (solid line) and dysplastic/cancer (dotted line) decision lines are indicated.

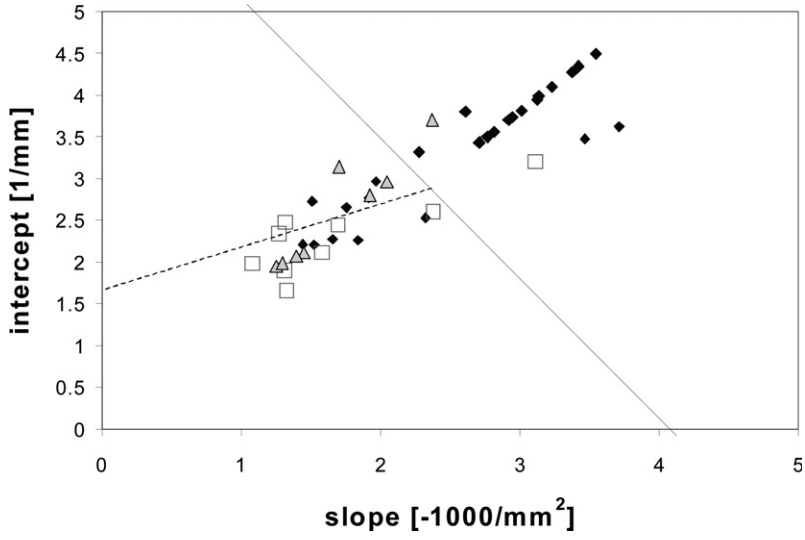

Fig. 5 Reduced scattering coefficients of non-keratinized oral tissue sites, slope–intercept scatter plot. Normal (diamonds); dysplastic (triangles); cancerous (squares). The normal/abnormal (solid line) and dysplastic/cancer (dotted line) decision lines are indicated.

volume will decrease, since the light undergoes most of its scattering within the connective tissue layer. Alternatively, this decrease could represent a decrease in the density of the collagenous network effected by increased activity of MMPs or other proteolytic enzymes.

LSS

To extract the light scattering spectrum from a particular site, the modeled diffusely reflected light component was subtracted from the measured reflectance spectrum. Fig. 6(a) shows the LSS spectra extracted from cervical reflectance spectra from representative SQM (solid line) and SIL (dashed line) sites. The LSS spectra were analyzed using a model based on elastic light scattering to obtain the size distribution of the epithelial cell nuclei.[28,30] The resulting nuclear size distribution curves are shown in Fig. 6(b). The width (standard deviation) of the SIL nuclear size distribution and the total number of nuclei per unit area are larger than those of the SQM site, and these were used as diagnostic parameters. For the oral cavity data, the

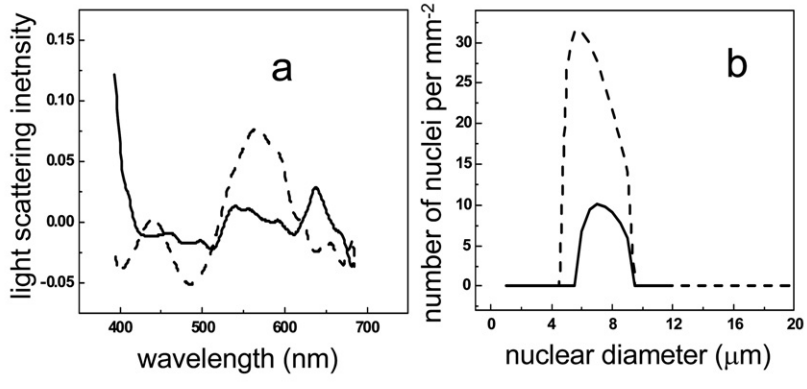

Fig. 6 LSS data for cervical tissue.(a) LSS spectra and (b) nuclear size distributions for SQM (solid) and SIL (dashed) sites.

percent of nuclei larger than 10 μm, and the width of the size distribution were the most useful diagnostic parameters (Fig. 7).

TMS

IFS, DRS and LSS provide complementary information about different aspects of tissue biochemistry and morphology. Thus, combining this information enhances the sensitivity and specificity with which we classify normal and pre-cancerous/cancerous tissues. Because the sample size is relatively small, we combined the information from the three spectroscopic techniques (TMS) simply by assigning a classification to each tissue site that is consistent with the results of at least two of three methods.

Results for the oral cavity data are shown in Table 1. The analysis yielded an overall sensitivity and specificity of 96% and 96%, respectively, for TMS identification of abnormal *vs.* normal oral tissue.

The results of the analysis for the cervical data are shown in Table 2. TMS analysis yielded an overall sensitivity and specificity of 92% and 90%, respectively, for identifying squamous intraepithelial lesions.

In both cases it is seen that combining the three spectroscopic modalities provides an improved overall sensitivity and specificity. TMS is more accurate than any of the three techniques taken alone. In addition, the use of several parameters, each associated with a different biological feature, provides improved understanding of the changes in tissue morphology and biochemistry accompanying epithelial neoplasia.

V. LSS imaging

The FastEEM samples millimeter-size regions of tissue with an optical fiber contact probe. Such a technique would be particularly valuable if combined with a clinical technique for wide area surveillance. Recently, we have extended the capabilities of LSS to an imaging mode that can map variations in epithelial cell nuclei of living tissues over large surfaces.[5] The resulting parametric/functional images provide direct quantitative measures of nuclear enlargement and chromatin content.

In the instrument, a highly collimated beam of plane-polarized light from a 75 W xenon arc lamp illuminates a 1.3 cm × 1.3 cm area of epithelium. (The size of the illuminated area can easily be increased or reduced.) Before reaching the sample, the light is filtered by one of 11 narrow-band

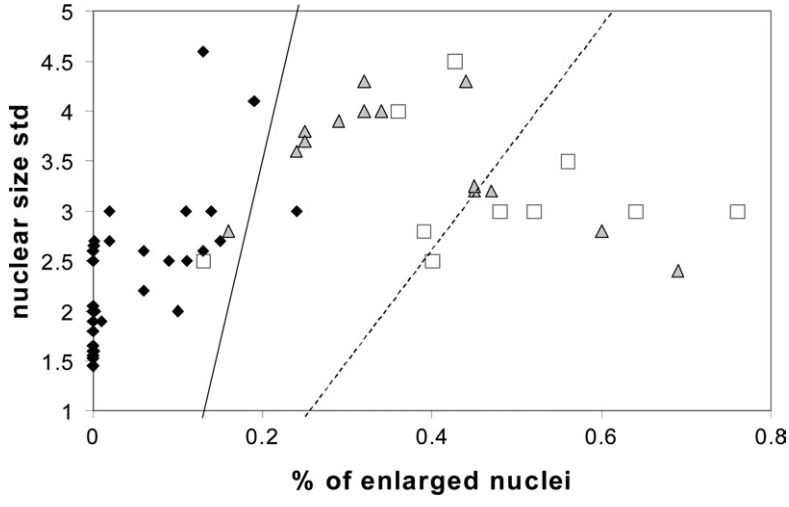

Fig. 7 LSS binary decision plots of oral tissue sites. Normal (diamonds); dysplastic (triangles); cancer (squares). The normal/abnormal (solid line) and dysplastic/cancer (dotted line) decision lines are indicated.

Table 1 Sensitivity and specificity ofvarious spectroscopic techniques to differentiate abnormal (dysplasia and cancer) from normal oral tissue

	Non-keratinized		Keratinized		Total	
	Abnormal *vs.* normal		Abnormal *vs.* normal		Abnormal *vs.* normal	
Modality	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
IFS	94%	96%	86%	96%	90%	96%
DRS	88%	68%	64%	81%	72%	74%
LSS	93%	100%	90%	91%	92%	97%
TMS	100%	96%	92%	96%	96%	96%

(4 nm bandwidth) filters spanning the range 450–700 nm, inserted in the beam path. The backscattered light is collected by a pair of equifocal achromatic lenses separated by twice their focal lengths and relayed to an analyzer and a CCD detector positioned one focal length away from the outer lens. This arrangement gives 1:1 imaging and a collection solid angle of 0.5°. The CCD array is 512×512 pixels, each pixel having dimensions 25×25 μm^2. Two images are collected at each of the eleven illumination wavelengths, one with the analyzer parallel to the polarization of the incident beam and the other perpendicular to it. Subtracting the two images pixel by pixel eliminates the unwanted depolarized component of the elastic scattered light emerging from the tissue that would otherwise dwarf single scattering, giving the contribution due to single backscattering, primarily from epithelial cell nuclei. The pixels are then binned 5 on a side to form 100×100 resolution elements, each 125 μm on a side. Each element contains an 11 point spectrum which can be analyzed using Mie theory to give the size distribution of the epithelial cell nuclei, their population density, and their refractive index relative to that of the cytoplasm. We then obtain diagnostic images of the spatial distribution of dysplasia in terms of the percentage of enlarged nuclei and the distribution of nuclear solid mass.

We have studied freshly resected colon tissue samples from patients undergoing colectomy. We imaged an adenoma surrounded by non-dysplastic colon mucosa using the LSS imaging apparatus (Fig. 8). For each pixel (25 μm × 25 μm) of the imaged field (1.3 cm × 1.3 cm), a spectrum of light backscattered by the nuclei was first distinguished from the rest of the reflected light using polarization discrimination and then analyzed using the Mie theory-based computerized algorithm. The parameters obtained were the size and refractive index of the nuclei in each pixel. The imaged field was divided into 125 μm × 125 μm regions, and the percentage of nuclei larger than 10 μm was obtained for each of these areas.

The resulting color-coded images are shown in Fig. 9. As expected, the nuclei are enlarged in the central, adenomatous region, but not in the surrounding non-dysplastic tissue (panel a). Panel b in the same figure is an LSS image of the spatial distribution of the mass of nuclear solids (chromatin and other constituents), displayed in units of pg (nucleus)$^{-1}$. These values were derived from the knowledge of the diameter and the refractive index of the nuclei, obtained from our LSS image. It is known that the nuclear solid mass correlates with chromatin content.[36,37] One can see that in the region of the polyp, the chromatin content is larger than in the surrounding normal tissue. This condition is an indication of dysplasia, as well.

Table 2 *In vivo* performance of the various spectroscopic techniques for separating SILs from non-SILs

	Biopsied non-SILs *vs.* SILS		All non-SILs *vs.* SILs	
Modailty	Sensitivity	Specificity	Sensitivity	Specificity
IFS	62%	67%	62%	92%
DRS	69%	57%	62%	82%
LSS	77%	71%	77%	83%
TMS	92%	71%	92%	90%

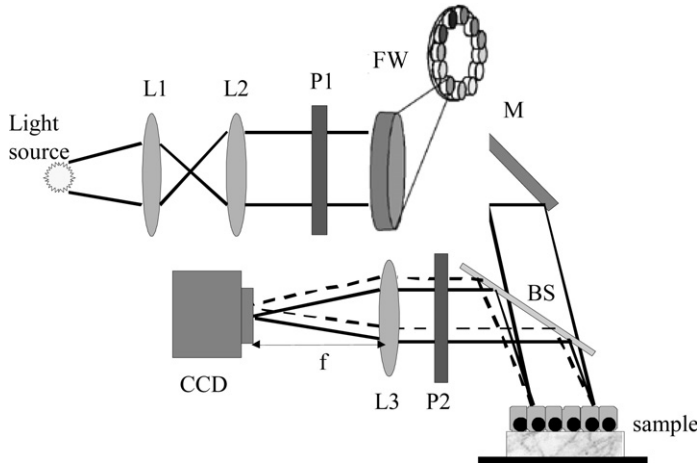

Fig. 8 Schematic diagram of LSS imaging apparatus. L1, L2: telescope lenses; P1, P2: polarizers; FW: filter wheel; M: mirror; BS: beam splitter; L3: imaging lens.

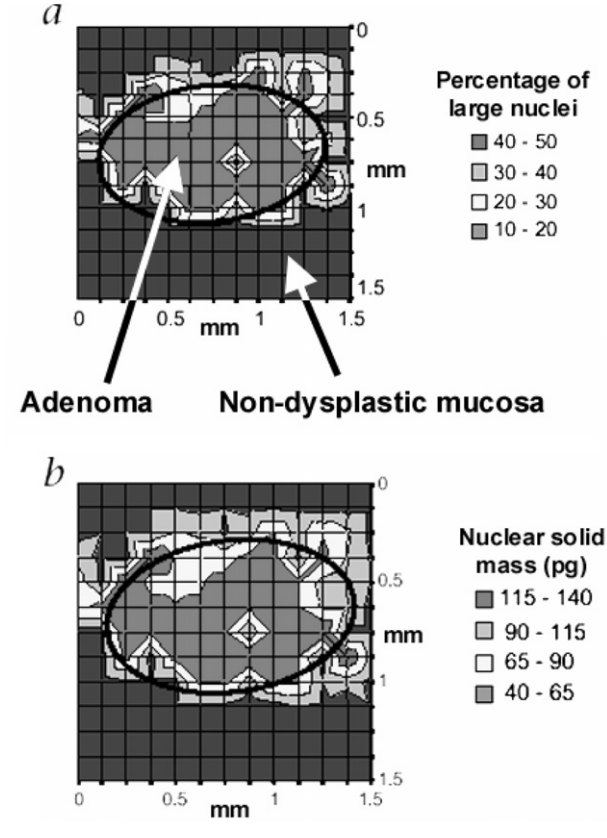

Fig. 9 LSS images of a colon tissue adenoma. Spatial distributions are shown for: (a) percentage of enlarged nuclei; (b) nuclear solid mass. The polyp borders are marked by ellipses.

The accuracy of measuring the size and relative refractive index was found to be 25 nm and 0.001, respectively. We note that sub-wavelength size resolution is possible because the scatterers act as interferometers and data is collected at multiple wavelengths. Hence, exceedingly small quantities can be measured, and measuring accuracy can be very high.

VI. Discussion

Epithelial neoplasms account for significant morbidity and mortality worldwide, and significant healthcare resources are spent each year in screening protocols for early detection of these neoplasms and their biological precursors. The current method for detecting dysplasia and early cancer in the oral cavity is visual screening followed by biopsy and histopathological evaluation of visible lesions. In the cervix, the Pap smear serves as a cytological screening technique, with positive results followed up by visual inspection with a low power microscope (colposcopy) and, again, by biopsy and histopathology of visible lesions. Visual inspection during colposcopy is aided by application of acetic acid, which often produces a non-specific ''whitening'' of various epithelial changes, including some dysplastic lesions. In the gastrointestinal tract, screening for dysplasia in the setting of Barrett's esophagus and chronic inflammatory bowel disease is based entirely on random sampling of the epithelium, because early dysplastic lesions are essentially invisible to the naked eye. Clearly, these clinical methodologies for detection of pre-cancers lack sensitivity and specificity, and often result in unnecessary biopsies or missed lesions.

The presence of dysplasia gives rise to a number of changes in the epithelium and the underlying stroma. Tissue structural and biochemical changes often include nuclear enlargement and crowding with increased variability in shape and size, loss of cellular maturation, and overall thickening of the epithelium.[38] The development of dysplasia and cancer is also commonly accompanied by inflammatory and architectural changes, including neovascularization.[39] The lamina propria undergoes reorganization and various proteases degrade parts of the basement membrane and underlying collagen network, facilitating cancer cell invasion and ultimately metastatic dissemination *via* the lymphatic and vascular system.[40]

The spectral changes produced by structural and biochemical changes in the epithelium and the subepithelial collagen network manifest themselves in the intrinsic fluorescence. NADH and collagen fluorescence in the oral cavity and cervix are modulated by changes in cell number and metabolic activity, and stromal fluorescence, respectively. In addition, some of the epithelium of the oral cavity tissue is covered with a layer of keratin, which scatters some of the incoming excitation light and reduces its penetration depth. This results in an NADH contribution that is relatively higher than that of collagen.

The increased cell and metabolic activity associated with progression of neoplasia also leads to an increase in NADH fluorescence. Additionally, due to the thickening of the epithelium, less excitation light can reach the underlying collagen network. It has also been shown that collagenase enzymes responsible for the degradation of collagen are usually present in tissue areas undergoing significant architectural changes.[41] This gives rise to the observed changes in NADH and collagen contributions to the tissue fluorescence spectrum with progression of disease.

The structural and biochemical changes in the epithelium and underlying stroma also manifest themselves in the diffuse reflectance spectra. The decrease in the scattering coefficient we observed with the presence of dysplasia may be explained by thickening of the epithelium and a slow destruction of collagen cross-links by proteases.

Microscopic tissue changes in the epithelium could be detected with LSS. Enlargement of epithelial cell nuclei and the increase in the standard deviation of the nuclear size distribution are well-established biomarkers of dysplastic cells and are commonly used by pathologists in making diagnostic decisions. Since LSS is sensitive to singly scattered light, the information obtained is limited to the epithelium, which complements the information from deeper tissue layers obtained with DRS. The direct backscattering detection geometry makes this technique more sensitive to structures of the size of cell nuclei, and interference from cell scattering can be neglected. Very good results were obtained for LSS.

The three techniques (IFS, DRS and LSS) provide several types of information for tissue characterization: biochemical, stromal scattering/absorption, and size distribution of the epithelial

nuclei. By combining the complementary tissue biochemical and morphological information thus provided, we achieve a more sensitive and specific diagnosis. We also note that the histopathology results, which were used as the standard, are themselves subject to strong intra-examiner and inter-examiner variations,[42] limiting the accuracy of our analysis. The good results we have obtained with our spectroscopic techniques indicate that TMS may be used as a guide to biopsy, thereby enhancing the physician's ability to detect lesions at an early stage.

The current study used modeling of fluorescence and reflectance spectra to extract quantitative biochemical and morphological features associated with cancer progression. Thus, complementary structural and biochemical information was extracted from a macroscopic tissue sample of approximately 1 mm^3 in dimension. Some of these morphological features are customarily employed qualitatively by pathologists in making their diagnoses. Thus, we could create a novel tissue classification scheme built on modeling spectra to derive the underlying causal features in malignant tissue progression. The inherent ability of spectroscopic techniques to provide independent algorithm-based diagnoses is in contrast to morphological imaging techniques such as optical coherence tomography and confocal microscopy that require "expert" analysis of the *in vivo* images to reach a diagnosis.

The successful imaging of colon adenomas using LSS demonstrates the promise of that technique for mapping large regions of tissue. Adenomas contain dysplastic tissue in polypoid form, making them easily recognizable and a good model to test the new imaging technology. Both nuclear enlargement and nuclear mass could be imaged, illustrating the power of the technique. Incorporation of LSS imaging and the TMS probe in an integrated instrument should provide a powerful new means of identifying dysplasia, by scanning wide areas of tissue for invisible lesions and then investigating suspicious sites in more detail with the TMS probe.

VII. Conclusion

The work presented here illustrates how optical spectroscopy can monitor morphological and biochemical features that accompany dysplasia and provide diagnostic information in a quantitative, objective, non-invasive manner, in real time, without the need for tissue removal or processing. This rich source of information also allows us to understand more fully the changes that take place during the onset, progression and regression of neoplasia, specially during chemoprevention or medical therapy. We showed that changes in NADH and collagen intrinsic fluorescence have excellent diagnostic potential, and that additional diagnostic changes in tissue hemoglobin concentration and stromal morphology are extracted from the diffuse reflectance spectra. Furthermore, complementary morphological information about epithelial cell nuclei is contained in the light scattering spectra. All tissue biochemical and structural information was obtained *in vivo* without the use of exogenous dyes, and without the introduction of artifacts due to tissue removal and processing. The use of the three spectroscopic techniques together in the TMS spectral probe provided the ability to accurately distinguish normal from abnormal tissue and dysplastic tissue from cancer. This probe, when combined with LSS imaging, should provide a powerful new tool for wide area surveillance of invisible dysplasia, guiding and perhaps eventually replacing biopsy during clinical screening and during surgery for detecting cancer margins in the operating room.

Acknowledgements

This research was carried out at the Massachusetts Institute of Technology Laser Biomedical Research Center and supported by National Institutes of Health grants P41-RR02594 and R01-CA72517.

References

1 R. Zangro, L. Silveira, R. Manoharan, I. Itzkan, R. Dasari, J. Van Dam and M. Feld, *Appl. Opt.*, 1996, **35**, 5211–5219.
2 I. Georgakoudi, B. Jacobson, J. Van Dam, V. Backman, M. Wallace, M Muller, Q. Zhang and K Badizadegan, *Gastroenterology*, 2001, **120**, 1620–1629.

3 I. Georgakoudi, E. Sheets, M. Muller, V. Backman, C. Crum, K. Badizadegan, R. Dasari and M. Feld, *Am. Obstet. Gynecol.*, 2002, **186a**, 374–382.
4 M. Muller, T. Valdez, I. Georgakoudi, V. Backman, C. Fuentes, S. Kabani, N. Laver, Z. Wang, C. Boone, R. Dasari, S. Shapshay and M. Feld, *Cancer*, 2003, **97**, 1681–1692.
5 R. Gurjar, V. Backman, L. Perelman, I. Georgakoudi, K. Badizadegan, I. Itzkan, R. Dasari and M. Feld, *Nat. Med.*, 2001, **7**, 1245–1248.
6 M. Mitchell, S. Cantor, N. Ramanujam, G. Tortolero-Luna and R. Richards-Kortum, *Obstet. Gynecol.*, 1999, **93**, 462–470.
7 S. Lam, T. Kennedy, M. Unger, Y. E. Miller, D. Gelmont, V. Rusch, B. Gipe, D. Howard, J. C. LeRiche, A. Coldman and A. F. Gazdar, *Chest*, 1998, **113**, 696–702.
8 M. Zellweger, P. Grosiean, D. Goujon, P. Monnier, H. van den Bergh and G. Wagnieres, *J. Biomed. Opt.*, 2001, **6**, 41–51.
9 A. Mycek, K. Schomacker and N. Nishioka, *Gastrointest. Endosc.*, 1998, **48**, 390–394.
10 M. Panjehpour, B. Overholt, T. Vo-Dinh, R. Haggitt, D. Edwards and F. Buckley, *Gastroenterology*, 1996, **111**, 93–101.
11 R. Cothren, M. Sivak, J. Van Dam, R. Petras, M. Fitzmaurice, J. Crawford, J. Wu, J. Brennan, R. Rava, R. Manoharan and M. Feld, *Gastrointest. Endosc.*, 1996, **44**, 168–176.
12 K. Schomacker, J. Frisoli, C. Compton, T. Flotte, J. Richter, T. Deutsch and N. Nishioka, *Gastroenterology*, 1992, **102**, 1155–1160.
13 C. Betz, M. Mehlmann, K. Rick, H. Stepp, G. Grevers, R. Baumgartner and A Leunig, *Lasers Surg. Med.*, 1999, **25**, 323–334.
14 A. Gillenwater, R. Jacob, R. Ganeshappa, B. Kemp, A. K. El-Naggar, J. L. Palmer, G. Clayman, M. F. Mitchell and R. Richards-Kortum, *Arch. Otolaryngol. Head Neck Surg.*, 1998, **124**, 125.
15 L. Brancaleon, J. Durkin, J. Menaker, G. Fallon and N. T. Kollias, *Photochem. Photobiol.*, 2001, **73**, 178–183.
16 M. Anidjar, O. Cussenet, S. Avrillier, E. Ettori, P. Teillac and A. Le Duc, *Ann. N. Y. Acad. Sci.*, 1998, **838**, 130–142.
17 F. Koenig, F. McGovern, H. Enquist, R. Larne, T. Deutsch and K. Schommacker, *Chest*, 1998, **113b**, 696–702.
18 M. Muller, I. Georgakoudi, Q. Zhang, J. Wu and M. Feld, *Appl. Opt.*, 2001, **40a**, 4633–4646.
19 Q. Zhang, M. Muller, J. Wu and M. Feld, *Opt. Lett.*, 2000, **25**, 1451–1453.
20 J. Wu, M. Feld and R. Rava, *Appl. Opt.*, 1993, **32**, 3585–3595.
21 I. Georgakoudi, B. Jacobson, J. Van Dam, V. Backman, C. Crum, K. Badizadegan, R. Dasari and M. Feld, *Cancer Res.*, 2002, **62b**, 682–687.
22 V. Wallace, J. Bamber, J. Crawford, R. Ott and P. Mortimer, *Phys. Med. Biol.*, 2000, **45b**, 2859–2871.
23 I. Bigio, S. Brown, G. Briggs, C. Kelley, S. Lakhani, D. Pickard, P. Ripley and I. Rose, *J. Biomed. Opt.*, 2000, **5**, 221–228.
24 Z. Ge, K. Schomacker and N. Nishioka, *Appl. Spectrosc.*, 1998, **52**, 833–839.
25 J. Mourant, J. Boyer, T. Johnson and J. Lacey, *Biomed. Opt.*, 1996, **1**, 1–8.
26 F. Koenig, R. Larne, H. Enquist, F. McGovern, K. Schomacker, N. Kollias and T. Deutsch, *Urology*, 1998, **51a**, 342–345.
27 G. Zonios, L. Perelman, V. Backman, R. Manoharan, M. Fitmaurice, J. Van Dam and M. Feld, *Appl. Opt.*, 1999, **38**, 6628–6637.
28 V. Backman, in *Physics*, MIT, Cambridge, 1998.
29 L. Perelman, V. Backman, M. Wallace, G. Zonios, R. Manoharan, A. Nusrat, S. Shields and M. Seiler, *Phys. Rev. Lett.*, 1998, **80**, 627–630.
30 V. Backman, M. Wallace, L. Perelman, J. Arendt, R. Gurjar, M. Muller, Q. Zhang and G. Zonios, *Nature*, 2000, **406**, 35–36.
31 M. Wallace, L. Perelman, V. Backman, J. Crawford, M. Fitzmaurice, M. Seiler, K. Badizadegan and S. Shields, *Gastroenterology*, 2000, **119a**, 677–682.
32 V. Backman, V. Gopal, M. Kalashnikov, K. Badizadegan, R. Gurjar, A. Wax, I. Georgakoudi, M. Muller, C. Boone, R. Dasari and M. Feld, *IEEE J. Sel. Top. Quantum Electron.*, 2001, **7**, 887–893.
33 M. Schumacher, N. Hollander and W. Sauerbrei, *Statistics Med.*, 1997, **16**, 2813–2827.
34 A. Talvensaari-Mattila, M. Apaja-Sarkkinen, M. Hoyhta, A. Westerland, U. Puistola and T. Turpeenniemi-Hujanin, *Gynecol. Oncol.*, 1999, **72**, 306–311.
35 B. Chance, P. Cohen, F. Jobsis and B. Schoener, *Science*, 1962, **137**, 499–508.
36 G. Brown, M. McEwen and M. Pratt, *Nature*, 1995, **176**, 161–162.
37 H. Davies, E. Deeley and E. Denby, *Exp. Cell Res. Suppl.*, 1957, **4**, 136–149.
38 C. Boone, J. W. Bacus, J. V. Bacus, V. Steele and G. Kelloff, *Proc. Soc. Exp. Biol. Med.*, 1997, **216**, 151–165.
39 A. Lee, L. Happerfield, L. Bobrow and R. Millis, *J. Clin. Pathol.*, 1997, **50**, 669–673.
40 B. Bodey, B. J. Bodey, S. Siegel and H. Kaiser, *In Vivo*, 2001, **15**, 57–64.
41 D. Ellis and I. Yannas, *Biomaterials*, 1996, **17**, 291–299.
42 L. Abbey, G. Kaugers and J. Gunsolley, *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.*, 1995, **80**, 188–191.