

**Characterizing variability in *in vivo*
Raman spectra of different anatomical
locations in the upper gastrointestinal
tract toward cancer detection**

Mads Sylvest Bergholt

Wei Zheng

Kan Lin

Khek Yu Ho

Ming Teh

Khay Guan Yeoh

Jimmy Bok Yan So

Zhiwei Huang

Characterizing variability in *in vivo* Raman spectra of different anatomical locations in the upper gastrointestinal tract toward cancer detection

Mads Sylvest Bergholt,^a Wei Zheng,^a Kan Lin,^a Khek Yu Ho,^b Ming Teh,^c Khay Guan Yeoh,^b Jimmy Bok Yan So,^d and Zhiwei Huang^a

^aNational University of Singapore, Optical Bioimaging Laboratory, Department of Bioengineering, Faculty of Engineering, Singapore 117576

^bNational University of Singapore and National University Hospital, Department of Medicine, Yong Loo Lin School of Medicine, Singapore 119260

^cNational University of Singapore and National University Hospital, Department of Pathology, Yong Loo Lin School of Medicine, Singapore 119074

^dNational University of Singapore and National University Hospital, Department of Surgery, Yong Loo Lin School of Medicine, Singapore 119074

Abstract. Raman spectroscopy is an optical vibrational technology capable of probing biomolecular changes of tissue associated with cancer transformation. This study aimed to characterize *in vivo* Raman spectroscopic properties of tissues belonging to different anatomical regions in the upper gastrointestinal (GI) tract and explore the implications for early detection of neoplastic lesions during clinical gastroscopy. A novel fiber-optic Raman endoscopy technique was utilized for real-time *in vivo* tissue Raman measurements of normal esophageal (distal, middle, and proximal), gastric (antrum, body, and cardia) as well as cancerous esophageal and gastric tissues from 107 patients who underwent endoscopic examinations. The non-negativity-constrained least squares minimization coupled with a reference database of Raman active biochemicals (i.e., actin, histones, collagen, DNA, and triolein) was employed for semiquantitative biomolecular modeling of tissue constituents in the upper GI. A total of 1189 *in vivo* Raman spectra were acquired from different locations in the upper GI. The Raman spectra among the distal, middle, and proximal sites of the esophagus showed no significant interanatomical variability. The interanatomical variability of Raman spectra among normal gastric tissue (antrum, body, and cardia) was subtle compared to cancerous tissue transformation, whereas biomolecular modeling revealed significant differences between the two organs, particularly in the gastroesophageal junction associated with proteins, DNA, and lipids. Cancerous tissues can be identified across interanatomical regions with accuracies of 89.3% [sensitivity of 92.6% (162/175); specificity of 88.6% (665/751)], and of 94.7% [sensitivity of 90.9% (30/33); specificity of 93.9% (216/230)] in the gastric and esophagus, respectively, using partial least squares-discriminant analysis together with the leave-one tissue site-out, cross validation. This work demonstrates that Raman endoscopy technique has promising clinical potential for real-time, *in vivo* diagnosis and detection of malignancies in the upper GI at the molecular level. © 2011 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.3556723]

Keywords: *in vivo* diagnosis; Raman spectroscopy; upper gastrointestinal tract; gastric; esophagus; endoscopy.

Paper 10600RR received Nov. 11, 2010; revised manuscript received Jan. 28, 2011; accepted for publication Jan. 31, 2011; published online Mar. 22, 2011.

1 Introduction

The patients suffering from gastrointestinal (GI) cancers have poor survival rates mainly due to the advanced stages upon the initial time of diagnosis. Esophageal and gastric malignancies in the upper GI tract are the two major causes of cancer-associated death in humans.^{1,2} Early diagnosis and localization of malignant lesions together with appropriate curative treatment [e.g., endoscopic mucosal resection, gastrectomy, esophagectomy, etc.] is critical to reducing the mortality rates of the patients.^{2,3} However, the identification and localization of pre-cancer and early flat cancerous lesions in the esophagus and gastric can be challenging for clinicians as conventional white-light

reflectance (WLR) endoscopy heavily relies on visual identification of gross morphological tissue changes, resulting in poor diagnostic accuracy. In recent years, optical imaging methods, such as autofluorescence imaging (AFI) technique capable of detecting the changes of endogenous fluorophores and morphological architectures of tissue, and the narrow-band imaging (NBI) technique which enhances visualization of irregular mucosal and vascular patterns, have shown promising diagnostic potential for *in vivo* detection of preneoplastic and early neoplastic lesions at endoscopy.^{4,5} While AFI and NBI imaging techniques provide high-detection sensitivities, these wide-field endoscopic imaging modalities still suffer from moderate diagnostic specificities due to interobserver dependence and the inherent lack of ability to reveal detailed biochemical information

Address all correspondence to: National University of Singapore, Zhiwei Huang, Optical Bioimaging Laboratory, Department of Bioengineering, Faculty of Engineering, 9 Engineering Drive 1, Singapore 117576. Tel: +65-6516-8856; Fax: +65-6872-3069; E-mail: biehw@nus.edu.sg.

about the tissue.^{4,5} Therefore, it is of imperative clinical value to develop molecular sensitive optical diagnostic technologies that can assist in guiding endoscopists for the targeted biopsies of suspicious lesions in the GI tract for improving early disease diagnosis and characterization during routine endoscopic inspections.

Near-infrared (NIR) Raman spectroscopy has shown promising potential for the diagnosis and characterization of neoplastic progression of tissues with high diagnostic specificities at the biomolecular level.^{6–25} NIR Raman spectroscopy is a nondestructive, inelastic light scattering technique in which the incident laser light is shifted in frequencies depending on the specific vibrational frequencies of molecules in tissue. The Raman spectra of biological tissues reflect specific biochemical structures and conformations (i.e., biochemical signatures) of tissue, providing the unique opportunity to distinguish between different tissue types.^{9–11} To date, most Raman spectroscopic studies of the GI tract have been limited to *in vitro* tissue Raman and microscopic studies in conjunction with sophisticated multivariate analysis to render diagnostic algorithms [e.g., principal components analysis, linear discriminant analysis, classification and regression trees, etc.] for tissue characterization and classification.^{10,14,16} For instance, the diagnostic sensitivities and specificities in the range of ~85–95% and ~90–98%, respectively, have been reported for differentiation between different pathologic types (e.g., intestinal metaplasia, helicobacter pylori infection, dysplasia, and adenocarcinoma) of gastric tissues *in vitro* using NIR Raman spectroscopy associated with multivariate analysis.^{15–17} NIR Raman spectroscopic diagnosis of Barrett's esophagus, dysplasia, and neoplasia of the esophagus *in vitro* have also been reported as well.^{18–20} The *in vivo* clinical utility of Raman spectroscopy has been limited by slow acquisition times and the need for developing miniaturized fiber probes for efficient tissue excitation and Raman photon collection.⁸ Very recently, we have successfully developed a high throughput image-guided (i.e., WLR/NBI/AFI) Raman endoscopic technique coupled with a 1.8 mm fiber optic Raman probe that fits into the instrumental channel of conventional medical endoscopes for realizing real-time, *in vivo* tissue Raman measurements.⁸ Consequently, the *in vivo* Raman endoscopic detection of gastric dysplasia,²² neoplasia,^{23,24} and differential diagnosis between benign and malignant ulcers²⁵ have been demonstrated with success. However, as the histological profiles and morphologies of distinctive anatomical regions in the upper GI tract (i.e., esophagus and gastric) are highly functionally specialized and exhibit significant variations in architectural properties and cell types (e.g., tissue thickness variability, distinct glandular types, secretion products, vascularity, etc.),^{26,27} the extent of *in vivo* Raman spectral interorgan and interanatomical variabilities of tissues in the upper GI tract has not yet been examined in detail in literature. Hence, the main aim of this study was to evaluate the magnitude of interorgan and interanatomical variability of *in vivo* normal tissue Raman endoscopic spectra in the esophagus and gastric as well as to assess the implication for early diagnosis of neoplastic lesions. The semiquantitative biomolecular modeling [i.e., non-negativity-constrained least squares minimization (NNCLSM)]^{23,27} using representative biochemical basis spectra was employed for estimation of the most prominent Raman active constituents of tissue in the upper GI tract.

2 Materials and Methods

2.1 Raman Endoscopy Instrumentation

The integrated Raman spectroscopy and trimodal wide-field imaging system developed for *in vivo* tissue measurements and characterization at endoscopy has been described in detail elsewhere.^{8,21–23} Briefly, the Raman endoscopy system consists of a spectrum stabilized 785 nm diode laser (maximum output: 300 mW, B&W TEK Inc., Newark, DE), a transmissive imaging spectrograph (Holospec f/1.8, Kaiser Optical Systems) equipped with a liquid-nitrogen cooled, NIR-optimized, back-illuminated, and deep depletion charge-coupled device (CCD) camera (1340×400 pixels at 20×20 μm per pixel; Spec-10: 400BR/LN, Princeton Instruments), and a specially designed 1.8-mm Raman endoscopic probe for both laser light delivery and *in vivo* tissue Raman signal collection. The novel Raman probe is composed of 32 collection fibers surrounding the central light delivery fiber with two stages of optical filtering incorporated at the proximal and distal ends of the probe for maximizing the collection of tissue Raman signals while reducing the interference of Rayleigh scattered light, fiber fluorescence, and silica Raman signals. The Raman probe can pass down through the instrument channel of most medical endoscopes and be directed to all anatomical tissue sites under the guidance of wide-field endoscopic imaging (WLR/AFI/NBI) modalities.⁸ Control of the *in vivo* Raman endoscopy system was implemented by a personal computer (PC) using a custom-designed software that triggers on-line data acquisition and analysis (e.g., CCD dark-noise subtraction, wavelength calibration, system spectral response calibration, signal saturation detection, cosmic ray rejection, etc.), as well as real-time display of *in vivo* tissue Raman spectra during clinical endoscopic measurements. The atomic emission lines from a mercury-argon spectral lamp (HG-1, Ocean Optics, Inc., Dunedin, FL) are used for wavelength calibration of tissue Raman spectra, and all wavelength-calibrated spectra were corrected for the wavelength-dependence of the system using a tungsten-halogen calibration lamp (RS-10, EG&G Gamma Scientific, San Diego, CA). The spectral resolution of the Raman system is 9 cm^{-1} and the system acquires Raman spectra in the wavenumber range of 800–1800 cm^{-1} from *in vivo* GI tissue within 0.5 s using the 785 nm excitation power of 1.5 W/cm^2 , which is less than the maximum permissible skin exposure limit set out by the American National Standards Institute.²⁸ Our further calculations based on the finite difference thermal model and the optical properties of the GI tissue^{29,30} indicate that even without consideration of other cooling effects (e.g., perfusion and evaporation in tissue), the maximum tissue temperature rise is only about 0.15 °C after 1 min of 785-nm laser radiation with an incident power of 30 mW on a tissue spot size of 200 μm during tissue Raman measurements. This temperature rise estimated is far below the level to generate cytotoxicity in cells and tissue,³¹ suggesting that the laser power density used in this study is safe for *in vivo* tissue Raman measurements in the upper GI. The trimodal endoscopy imaging system primarily comprises a 300 W short-arc xenon light source, a gastrointestinal videoscope (GIF-FQ260Z, Olympus), and a video system processor (CV-260SL, Olympus). The light reflected or autofluorescence emitted from tissue are detected by two monochrome CCD chips mounted behind the two objective lens placed next to each other at the distal tip of the GI videoscope: one CCD for

WLR/NBI and the other one for AFI. With this unique image-guided Raman endoscopy system, wide-field endoscopic images (WLR/AFI/NBI) and the corresponding *in vivo* Raman spectra of the tissue imaged can be simultaneously acquired, displayed, and recorded in the video system processor and the PC, respectively. The raw Raman spectra (800–1800 cm^{-1}) measured from *in vivo* GI tissue represented a composition of weak Raman signals, intense autofluorescence background and noise. Thus, the raw spectra were preprocessed by a first-order Savitsky-Golay smoothing filter (window width of five pixels, which corresponded to the system spectral resolution) to reduce noise.⁸ A fifth-order polynomial was found to be optimal for fitting the autofluorescence background in the noise-smoothed spectrum,⁸ and this polynomial was then subtracted from the raw spectrum to yield the tissue Raman spectrum alone. Each background-subtracted Raman spectrum was also normalized to the integrated area under the curve from 800–1800 cm^{-1} , enabling a better comparison of the spectral shapes and relative Raman band intensities among different organs and anatomical regions. All the spectral preprocessing is completed online and the Raman spectrum and outcome (e.g., of diagnostic decision algorithms) can be displayed real-time in a comprehensible graphical user interface during clinical Raman measurements at endoscopy.

2.2 Patients

The present study is part of an ongoing nationwide cancer screening program aiming at early diagnosis and treatment of upper GI malignancies run by the Singapore gastric cancer epidemiology, clinical and genetic program.³² This study was conducted with approval by the Institutional Review Board of the National Healthcare Group of Singapore. All patients signed an informed consent permitting the collection of *in vivo* upper GI Raman spectra in the endoscope centre at the National University Hospital, Singapore. In this study, 107 patients (61 men and 46 women with a median age of 66 years old) were enrolled for *in vivo* Raman endoscopy under the guidance of multimodal wide-field imaging modalities (WLR/AFI/NBI).⁸ The Raman probe was passed into the endoscope instrument channel and placed in gentle contact with the esophageal and gastric mucosal surfaces; and the probe positioning against the mucosa was verified on the endoscopy monitor by the clinicians in-charge during endoscopic examinations. As a result, a total of 1189 *in vivo* Raman spectra were acquired from the upper GI tract including normal (antrum sites with greater and lesser curvatures, body sites with greater and lesser curvatures, and cardia) and cancerous gastric tissues, as well as normal (distal end, middle sites, and proximal sites) and cancerous esophageal tissues. Table 1 summarizes the number of *in vivo* Raman spectra acquired and the corresponding integration time required for each tissue type in the upper GI. Immediately after Raman endoscopic acquisitions, the biopsies were taken from the tissue sites measured (with suction markings) and fixed in 10% formalin solution for histopathological confirmation of normal or malignant tissue types by a senior gastrointestinal pathologist who was blinded to the Raman spectroscopy results. Note that the cancerous gastric tissues [adenocarcinomas ($n = 175$)] were found in different sites (e.g., antrum/body/cardia) of the stomach; whereas

Table 1 The number of *in vivo* Raman spectra acquired and the corresponding integration time required for each tissue type in the gastric and esophagus.

Tissue type	Number of Raman spectra acquired	Integration time (s)	Patient numbers
<i>Gastric</i>			
Antrum (normal)	377	0.5	62
Body (normal)	268	0.2	42
Cardia (normal)	106	0.3	26
Cancer	175	0.5	35
<i>Esophagus</i>			
Distal (normal)	57	0.4	13
Middle (normal)	66	0.4	15
Proximal (normal)	107	0.4	19
Cancer	33	0.5	15

cancerous esophageal tissues [including adenocarcinomas ($n = 26$) and squamous cell carcinoma ($n = 6$)] were localized in the distal and middle parts of the esophagus, respectively. The histopathological examinations served as the gold standard for evaluating the diagnostic sensitivity and specificity of *in vivo* Raman endoscopy technique.

2.3 Biomolecular Modeling

The biomolecular and histological intuitive characterization of esophageal and gastric tissue is essential for implementation of the Raman endoscopy technique as a clinical tool for *in vivo* diagnosis of cancer and precancer in the upper GI. The biochemical spectral fitting by means of Raman spectra of cell constituents (e.g., cytoplasm, nucleus, etc.) and pure biochemicals has been extensively employed to gain better understanding of distinctive tissue constituents associated with pathologic changes.^{18,23} In this work, a semiquantitative model of *in vivo* tissue Raman spectra is rendered based on *a priori* insight of inter-/intra-cellular constituents using a linear combination [i.e., NNCLSM]²³ of basis Raman spectra that represent the main biochemical constituents in GI tract. Of over 35 basis reference Raman spectra obtained from different biomolecules associated with GI tissue (e.g., actin, albumin, pepsin, pepsinogen, B-NADH, RNA, DNA, myosin, hemoglobin, collagen I, collagen II, collagen V, mucin 1, mucin 2, mucin 3, flavins, elastin, phosphatidylcholine, cholesterol, glucose, glycogen, triolein, histones, beta-carotene, etc.),^{6–11,13,15–23,26,27} our NNCLSM modeling indicate that the following five biochemicals, i.e., actin (A3653), histones (H6005), collagen type I (C9879), DNA (P4522), and triolein (T7140) (Sigma, St Louis, MO), were the most significant Raman-active biochemical constituents that can effectively characterize gastric and esophageal (normal and cancerous) tissue with very small fit-residuals.²³ For instance, DNA represents nucleic acids within the cell nucleus; triolein represents typical lipid signals; actin and histones resembles

Table 2 Tentative molecular vibrational and biochemical assignments (Refs. 9–11 and 14–17) of *in vivo* Raman spectra of esophageal and gastric tissues (ν , stretching mode; ν_s , symmetric stretching mode; δ , bending mode).

Raman peaks (cm^{-1})	Vibrational assignments	Biochemicals
853	ν (C–C)	proteins (collagen)
936	ν (C–C) in α conformation	proteins
1004	ν_s (C–C) breathing	proteins (phenylalanine)
1078	ν (C–C)	lipids
1265	Amide III ν (C–N) δ (N–H)	proteins
1302	δ (CH ₂)	proteins/lipids
1335	CH ₃ CH ₂ wagging mode	DNA/proteins
1445	δ (CH ₂)	lipids/proteins
1618	ν (C = C)	porphyrins
1655	Amide I ν (C = O)	proteins/lipids
1745	ν (C = O)	lipids

proteins of different conformations and are the major components of the cytoskeleton and chromatin, respectively, whereas collagen type I is a substantial part of the extracellular matrix. Note that the basis Raman spectra of these biochemicals were measured in their native conditions without any further purification using the Raman Endoscopy instrumentation previously described⁸ for semiquantitative biochemical modeling. The five coefficients rendered using NNCLSM were further constrained to a sum of 100%. Data analysis was performed using in-house written scripts in the Matlab (Mathworks Inc., Natick, MA) programming environment.

3 Results

High quality *in vivo* Raman spectra can routinely be acquired from different anatomical locations or lesion sites within the esophagus and gastric in real-time (<0.5 s) under multimodal imaging (WLR, NBI, and AFI) guidance during clinical endoscopic examinations. Figure 1(a) shows the *in vivo* mean Raman spectra \pm 1 standard deviations (SD) of normal (antrum, body, and cardia) and cancer gastric tissues; whereas Fig. 1(b) shows the *in vivo* mean Raman spectra of cancer and normal (distal, middle, and proximal) esophageal tissues. The Raman signals of the esophageal and gastric mucosa were reproducible among different patients. Prominent Raman bands are observed in both organs in the following peak positions with tentative biomolecular assignments (Table 2):^{9–11,14–17} 853 cm^{-1} [ν (C–C) of proteins], 936 cm^{-1} [ν (C–C) of α -helix conformation for proteins], 1004 cm^{-1} [ν_s (C–C) ring breathing of phenylalanine], 1078 cm^{-1} (C–C stretching mode of lipids), 1265 cm^{-1} [Amide III ν (C–N) and δ (N–H) of proteins], 1302 and 1335 cm^{-1} [δ (CH₂) deformation and CH₂CH₃ twisting of proteins and nucleic acids], 1445 cm^{-1} [δ (CH₂) of proteins and lipids], 1618 cm^{-1}

Table 3 The diagnostic significance (p -value) of the distinctive Raman peaks identified [one-way ANOVA with Bonferroni correction at 5% (Ref. 33)] from *in vivo* normal gastric tissues (antrum, body, and cardia) based on the post-hoc Fisher's LSD test (Refs. 17 and 34) [* p -value < 0.01 (diagnostically significant based on pair-wise comparison)].

Raman peaks (cm^{-1})	p -values based on post-hoc Fisher's LSD test		
	Antrum versus body	Antrum versus cardia	Body versus cardia
936	$*9.1 \times 10^{-5}$	1.1×10^{-2}	4.4×10^{-1}
1245	$*2.0 \times 10^{-7}$	$*5.2 \times 10^{-5}$	8.9×10^{-1}
1335	$*1.1 \times 10^{-16}$	$*1.2 \times 10^{-10}$	2.1×10^{-2}
1618	$*6.7 \times 10^{-16}$	$*2.7 \times 10^{-3}$	9.5×10^{-2}
1665	$*6.7 \times 10^{-16}$	$*2.7 \times 10^{-3}$	9.5×10^{-2}
1745	$*3.1 \times 10^{-17}$	$*9.3 \times 10^{-12}$	4.0×10^{-1}

(C = C stretching mode of porphyrins), 1655 cm^{-1} [Amide I ν (C = O) of proteins] and 1745 cm^{-1} [ν (C = O) of lipids]. The difference spectra \pm 1 SD between different tissue types [Figs. 2(a) and 2(b)] clearly resolve the subtle interanatomical variations and the major biochemical changes associated with cancer transformation; whereas Fig. 2(c) reveals important biochemical differences between the two organs within the gastroesophageal (GE) junction (i.e., the distal end of the esophagus and cardia in the stomach). One-way analysis of variance (ANOVA) with Bonferroni correction at 5% (Ref. 33) was separately employed for each organ (normal tissue). Table 3 summarizes the six significant Raman peak positions identified as well as the corresponding p -values generated from the post-hoc Fisher's least significant difference (LSD) test^{17,34} of all pair-wise anatomical regions in normal gastric tissue (i.e., antrum versus body; body versus cardia; antrum versus cardia). Similar statistical analysis has not found significant interanatomical variability of normal tissues within the esophagus (i.e., distal, middle, and proximal); whereas significant spectral changes ($p < 1.0 \times 10^{-6}$) were found between the two organs (distal esophagus versus cardia), particularly at 1078, 1150, 1197, 1302, 1409, 1455, and 1745 cm^{-1} . One notes that no significant interanatomical variability ($p > 0.1$) of malignant tissues in the upper GI has been found by Raman endoscope technique, which is probably due to the reason that the sizes of the tumor tissues measured in the esophagus and gastric are much larger than the probing volume ($\sim 1 \text{ mm}^3$) of our endoscopic Raman probe design.^{8,22}

To characterize the spectral contribution from important Raman active tissue constituents (e.g., proteins, DNA, and lipids) in gastric and esophageal tissue types, the NNCLSM was subsequently employed for semiquantitative modeling using the five basis reference spectra acquired as shown in Fig. 3 (i.e., actin, collagen type I, DNA, histones, and triolein). Figures 4(a)–4(d) show the comparisons of the mean measured Raman spectra and the reconstructed Raman spectra of different tissues (i.e., gastric normal/cancer and esophageal normal/cancer). The residual variations of the NNCLSM model rendered from each of the tissue sites in the gastric and

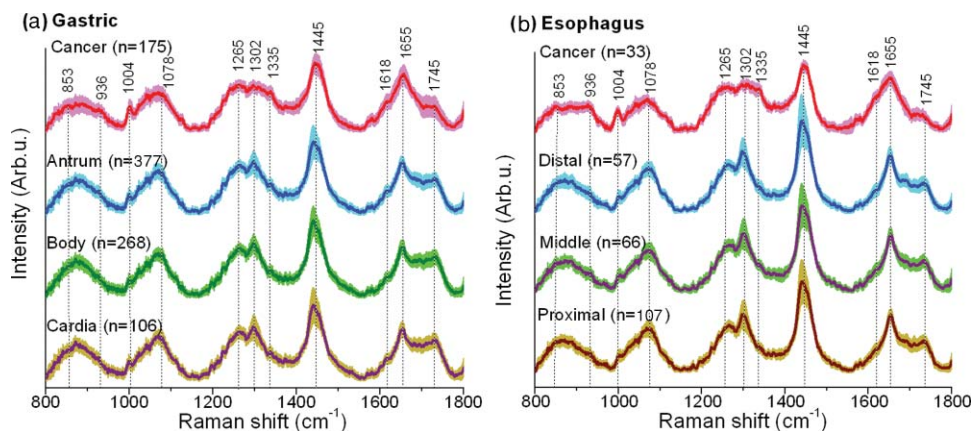


Fig. 1 (a) *In vivo* mean Raman spectra ± 1 SD of normal (antrum, body, and cardia) and neoplastic gastric tissues. (b) *In vivo* mean Raman spectra ± 1 SD of the normal (distal, middle, and proximal) and neoplastic esophageal tissues. Note that the mean Raman spectra are vertically shifted for better visualization.

esophagus are listed in Table 4. The fit-residuals between the reconstructed and the measured Raman spectra are of less than 10%, indicating that the five tissue biochemical constituents used in NNCLSM modeling can largely account for most diagnostic information contained in tissue Raman spectra of the upper GI. To further assess the validity of biomolecular modeling for tissue characterization, we rendered difference spectra ± 1 SD of the reconstructed Raman spectra for estimating the similarities with the measured difference spectra. The correlation evaluations of the modeled difference spectra with the respective graphs of Figs. 2(a)–2(c) confirmed that cancerous tissue transformation and interorgan variability can largely be resolved using biomolecular modeling (correlation coefficient $R > 0.82$); but the interanatomical variability is markedly delicate that biomolecular modeling could not adequately resolve these subtle spectral changes (correlation coefficient $R < 0.19$).

Figures 5(a) and 5(b) display the relative Raman spectral contribution of the five biochemicals of different tissue sites

in the gastric and esophagus together with the corresponding *p*-value reflecting subtle interanatomical Raman spectral changes while significant differences associated with tissue carcinogenesis processes.

4 Discussion

We report on our recently developed multimodal image-guided Raman endoscopic technique that offers the advantages of multiparametric measurements of intrinsic biomolecular constituents for real-time *in vivo* tissue diagnosis and characterization in the upper GI. With a specially designed miniaturized Raman endoscopic probe,⁸ which can fit into the instrument channel of medical endoscopes for effective excitation and collection of tissue Raman scattered photons, the clinicians are now able to nondestructively assess the endogenous biochemical and morphological information of internal tissues under the wide-field endoscopic imaging (WLR, NBI, and AFI) guidance. Unlike *in vitro* Raman spectroscopy,^{14–17} the image-guided Raman endoscopic technique can provide real-time biochemical assessment of tissues *in situ* (e.g., free of artifacts introduced by mechanically cutting, oxidation, hydration, loss of mucus layer, and vascular pressure, etc.) and therefore offer the Raman endoscopy as a possible routine diagnostic tool in the esophagus and gastric. However, as the interanatomical regions of the esophagus and particularly the gastric tissue are highly biological and functionally specialized (e.g., thickness variability, distinct glandular types, secretion products, vascularity, etc.),^{26,27} it is essential to figure out, to which extent, the effect of interorgan and interanatomical tissue sites' variations may affect the diagnostic algorithm developments of the Raman endoscope technique for the accurate diagnosis of early cancerous tissue in the esophagus and gastric.^{22–25}

In this work, we evaluate for the first time *in vivo* Raman spectral properties of different anatomical regions of normal gastric and esophageal tissues, whereby we have identified distinct Raman bands in both organs that are highly associated with proteins, DNA, and lipids [Figs. 1(a) and 1(b)]. The difference spectra [Figs. 2(a) and 2(b)] resolve the subtle interanatomical variability, indicating that the overall biomolecular and biochemical constituents of different normal tissues are very similar, whereas neoplastic tissues exhibit distinctive Raman

Table 4 Residual variations of the NNCLSM model rendered from each of the tissue sites in the gastric and esophagus.

Tissue type	NNCLSM model residual variations
<i>Gastric</i>	
Antrum (normal)	$\pm 1.0 \times 10^{-3}$
Body (normal)	$\pm 1.0 \times 10^{-3}$
Cardia (normal)	$\pm 1.0 \times 10^{-3}$
Neoplasia	$\pm 1.0 \times 10^{-3}$
<i>Esophagus</i>	
Distal (normal)	$\pm 8.3 \times 10^{-4}$
Middle (normal)	$\pm 8.5 \times 10^{-4}$
Proximal (normal)	$\pm 8.8 \times 10^{-4}$
Neoplasia	$\pm 9.9 \times 10^{-4}$

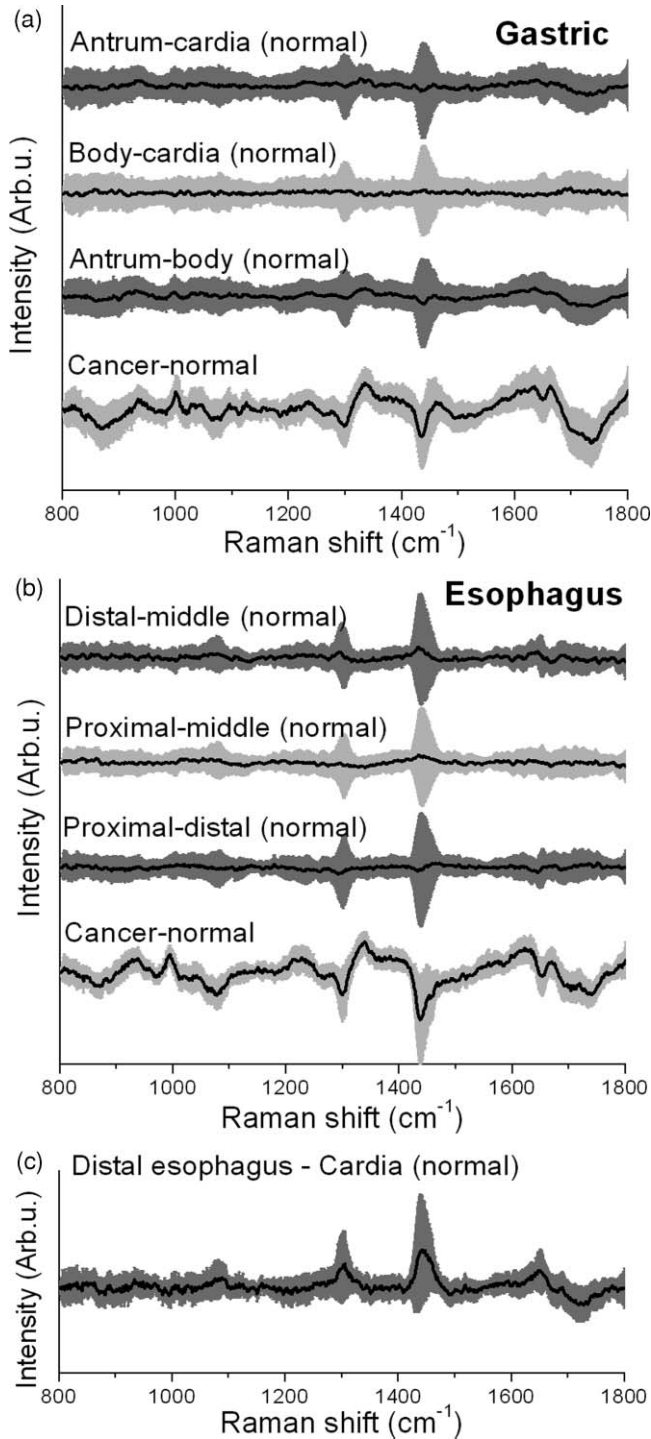


Fig. 2 Difference spectra ± 1 SD comparing different anatomical regions and tissue types of the (a) gastric and (b) esophagus tissue. (c) Difference spectrum of the distal end of the esophagus and the cardia in the stomach, resolving interorgan variability at the molecular level.

spectral profiles (e.g., Raman peaks shifting, bandwidths broadening, or narrowing, relative Raman peak intensity changes). Further Raman spectral analysis also reveals significant spectral differences of tissues within the GE junction [Fig. 2(c)], confirming the promising ability of Raman endoscopy for the sensitive characterization of morphological and biomolecular constituents in the upper GI tract. The Raman active biochemi-

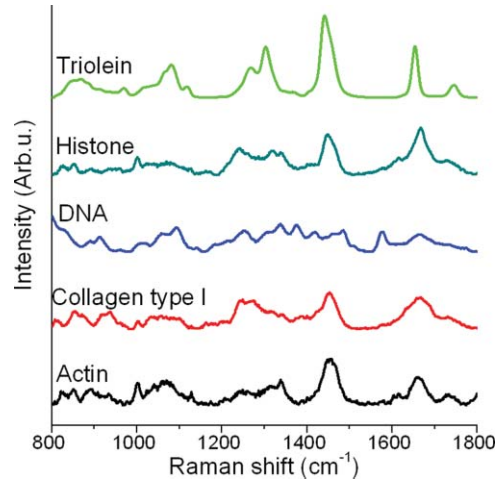


Fig. 3 The five basis reference Raman spectra (i.e., actin, collagen, DNA, histones, and triolein) are used for biochemical modeling of esophageal and gastric tissue.

cal profiles of gastric and esophageal tissue types were further assessed by rendering semiquantitative models constructed from essential biochemical constituents of gastric and esophageal tissues. For instance, proteins were found to be the most prominent contributors in *in vivo* tissue Raman spectra, which are similar to the *in vitro* tissue studies.^{18,35,36} The difference spectra [Figs. 2(a) and 2(b)] also reflected large interpatient variations in the subregions 1078 [$\nu(\text{C}-\text{C})$], 1302 [$\delta(\text{CH}_2)$ deformation], and 1440 cm^{-1} [$\delta(\text{CH}_2)$], which are highly associated with signals from proteins and lipids, as confirmed by biochemical modeling (Fig. 5). Moreover, the statistical analysis (Table 3) indicated subtle but complex differences in signals originating from proteins, lipids, and DNA of the antrum as compared to the body and cardia sites of the stomach. These distinctive spectral features of the body likely reflect the intricate morphology³⁷ and biomolecular compositions of the body regions that constitute highly specialized cells (e.g., zymogenic and oxyntic cells in a densely packed architecture secreting acids, enzymes, etc.).²⁶ For instance, the relative less collagen of the body site compared to antrum as revealed by biomolecular modeling ($p = 6.2 \times 10^{-2}$) could be associated with the closely packed morphologic structures containing less connective tissue.²⁶ We also found that the *in vivo* Raman spectrum of cardia resembled the body region, which might be caused by the endoscopically ill defined and narrow zone of the cardia region and gradual transitional mucosa with cardia occasionally bearing some resemblance with body histology (e.g., presence of oxyntic cell types).^{26,38} But the difference spectra [Fig. 2(b)] and statistical analysis indicated no statistically significant interanatomical variability of normal tissues within the esophagus ($p > 0.1$), confirming that the relatively thick squamous epithelium (~ 10 – 20 cell layers) is approximately identical among the distal, middle, and proximal regions of the esophagus.²⁷ Our Monte Carlo simulations have indicated that the 785-nm laser light penetration depth from the Raman endoscopic probe developed is ~ 750 – $800 \mu\text{m}$ with a $200 \mu\text{m}$ beam size on the tissue surface, which is well within the lamina propria (the thickness of squamous epithelium is ~ 300 – $500 \mu\text{m}$) of the esophagus.²⁷ Since most of the collected Raman scattered light originates from the shallower layer of the mucosa (e.g., epithelium) in

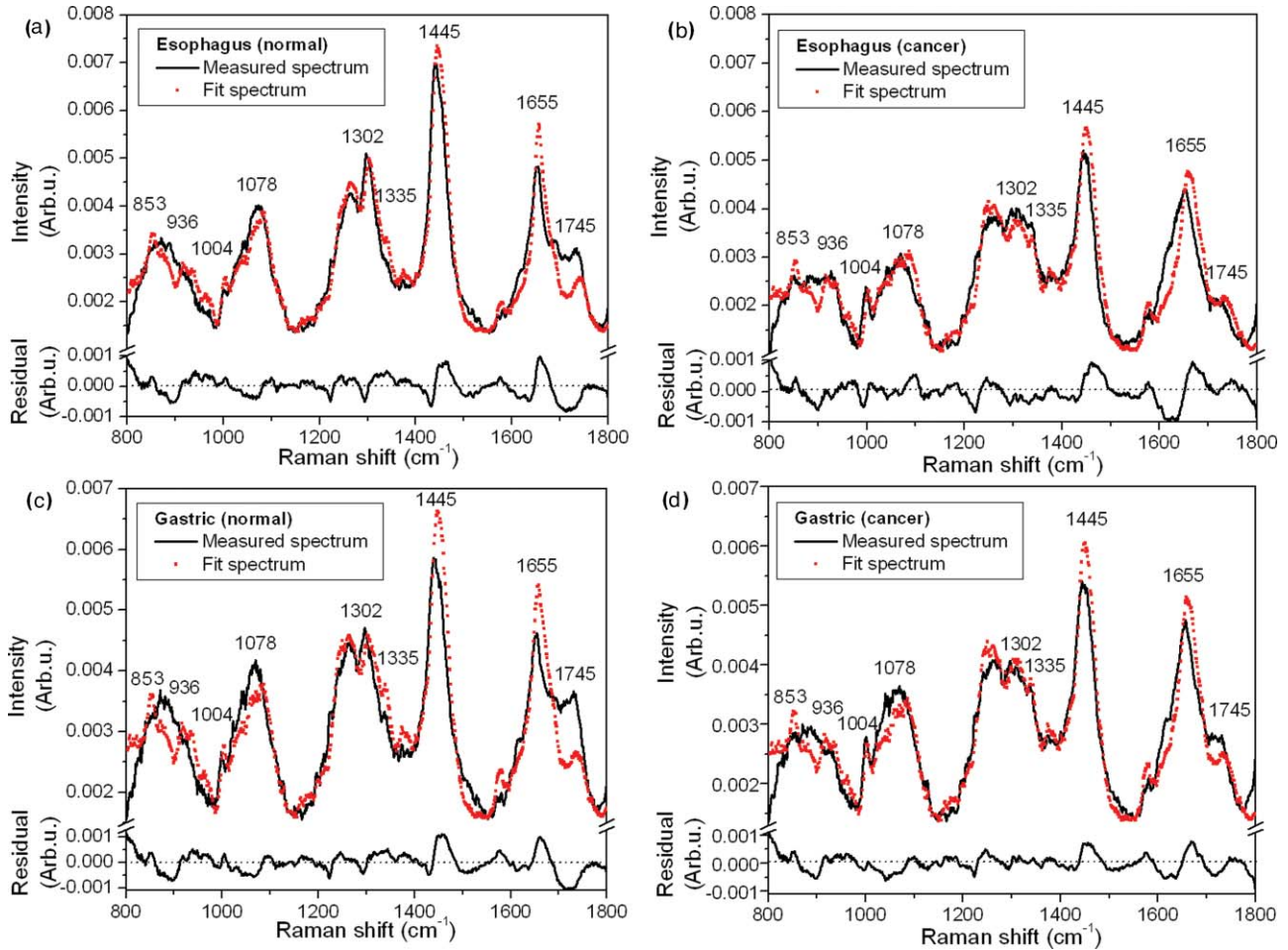


Fig. 4 Comparison of the *in vivo* Raman spectra measured with the reconstructed tissue Raman spectra through the employment of the five basis reference Raman spectra: (a) normal esophagus, (b) esophageal cancer, (c) normal gastric, (d) gastric cancer. Residuals (measured spectrum minus fit spectrum) are also shown in each plot.

which early neoplastic tissue transformation (e.g., dysplasia, carcinoma *in situ*) is most likely to originate,^{15,16,22,23} our fiber-optic Raman endoscopic probe still ensures a relatively small probing volume (of less than 1 mm³ in tissue volume) covering the epithelium tissue layer,²²⁻²⁵ which permits the Raman evaluation of precancer and early cancer in the upper GI.

On the other hand, the difference spectrum [Fig. 2(c)] of normal distal esophagus and gastric cardia within the GE junction exhibits significant Raman peaks shifting, bandwidths broadening/narrowing, and relative Raman peak intensity changes with the most prominent features at 1078, 1150, 1197, 1302, 1409, 1455, and 1715 cm⁻¹, while the NNCLSM

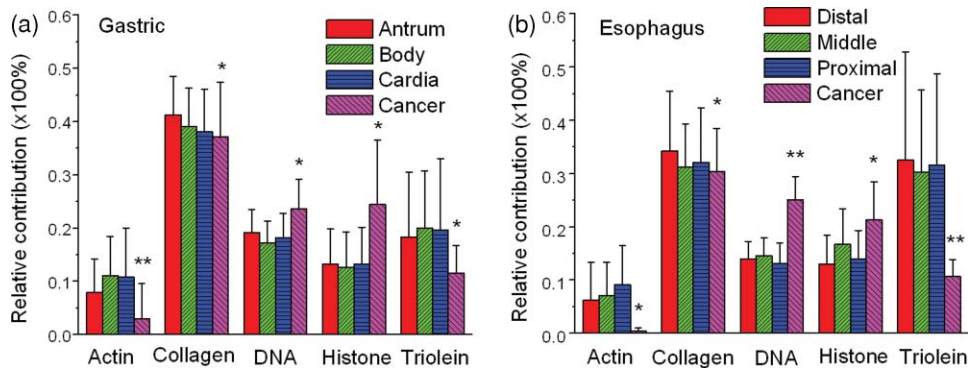


Fig. 5 Histogram displaying the average biochemical spectral contributions to the different tissue types in the (a) gastric and (b) esophagus tissue. The one SD confidence intervals are shown for each model component. Note: * indicates a significant differences ($p < 1.0 \times 10^{-3}$); whereas ** indicates a significant differences ($p < 1.0 \times 10^{-8}$).

modeling suggests a complex variation in Raman spectral contributions of proteins DNA and lipids of different organs (Fig. 5). These observations relate to the characteristic morphological architectures of the distinctive epithelial types (i.e., columnar lined versus nonkeratinized squamous epithelia) and glandular architectures etc. belonging to different organs^{26,27,38,39} as they were also visible under wide-field imaging observations. The biochemical modeling indicates a considerably elevated content in lipids of the distal esophagus as compared to the gastric cardia ($p = 3.2 \times 10^{-6}$), which is in agreement with the reports on accumulations of lipids in the dense cell layers of the esophageal epithelium, as a part of the mucosal barrier.^{40,41} Further, the relative less collagen of the esophagus compared to gastric is also linked with the thick squamous epithelium obscuring the Raman signal from the extensive connective tissues in the lamina propria. This optical filtering effect has also been observed in autofluorescence spectroscopy of gastric and esophageal neoplastic tissues as a result of the increased epithelial proliferation.^{23,24,42} It is particularly interesting that we observed inverse linear relationship (correlation coefficient $R > 0.92$) between collagen and triolein in the esophagus, suggesting that the large interpatient variation in the lipids signals could be connected with epithelial properties (e.g., thickness, lipid content)^{40,41} within the esophagus. To further evaluate the clinical implications of interorgan variability, we also employed a powerful classification algorithm [e.g., partial least squares-discriminant analysis (PLS-DA)]²⁵ to exploit subtle spectral variations of the entire Raman spectra that are not resolved by biomolecular modeling. The classification accuracy of 87.1% (142/163) could be achieved based on the leave-one tissue site-out, cross-validation method,⁴³ proving that the Raman spectra reflect surface and subsurface tissue structures and biomolecular constituents.

The differences in *in vivo* Raman spectra between neoplastic and nonneoplastic gastric tissues have also been reported in our previous gastric Raman studies.^{22–25} The NNCLSM modeling demonstrates that the neoplastic tissues show different contributions of distinctive biomolecules across interorgan and interanatomical regions in the upper GI [Figs. 5(a) and 5(b)], substantiating the important biochemical/biomolecular changes of tissues associated with carcinogenesis processes *in situ*. The significance of cancerous tissue transformation as compared to interanatomical variability in the gastric was further verified by ANOVA and post-hoc Fisher's LSD analysis [e.g., cancer versus normal gastric tissues (antrum, body, and cardia) with significant differences in the Raman bands: 936 ($p < 2.1 \times 10^{-2}$), 1245 ($p < 5.5 \times 10^{-4}$), 1335 ($p < 3.7 \times 10^{-28}$), 1618 ($p < 3.1 \times 10^{-31}$), 1665 ($p < 1.7 \times 10^{-25}$), and 1745 cm^{-1} ($p < 2.1 \times 10^{-30}$)]. Particularly, the Raman peaks at 1575 and 1335 cm^{-1} are highly associated with increased DNA contents^{22–25} that were also confirmed by biomolecular modeling [Figs. 5(a) and 5(b)]; whereas the increased Raman band intensity and the bandwidth broadening of the 1655 cm^{-1} amide I $\nu(\text{C}=\text{O})$ vibration pertained to a higher content of α -helical proteins (e.g., histones, the main protein component that makes up the chromatin). Overall, these biomolecular changes suggest the increased nuclear activity (e.g., hyperchromatic state) in neoplastic tissue, which is among the main characteristics of tissue carcinogenesis and is of considerable pathological value for tissue diagnosis and

characterization.²⁷ The biomolecular modeling also shows a small decrease in collagen, which is in agreement with the reports that cancerous cells proliferate, invade into underlying stromal layer and express a class of metalloproteases, resulting in an overall reduction of collagen content in cancer tissue.^{22–25} The relative increase in Raman signals from proteins (e.g., 1004, 1265, and 1655 cm^{-1}) could also be related to increased mitotic activity,^{22–25,44} while the reduction in Raman signals related to lipids (at 1078, 1302, 1440, and especially at 1745 cm^{-1}) for precancer and cancer gastric tissues has also been found in our previous Raman studies.^{22–25} On top of these, the Raman peak increase at 1575–1618 cm^{-1} is highly associated with angiogenesis process of neoplastic tissue,^{23–25} which in fact, is reflected by the poor fitting ability of the NNCLSM model in this spectral region due to the omission of Raman-active blood constituents (e.g., hemoglobin) in the basis reference set. Figures 5(a) and 5(b) uncover the diagnostic significant fit coefficients for detection of neoplasms in the gastric and esophagus (e.g., actin, collagen, DNA, histones, and triolein); and the consistency in identifying similar diagnostic significant biochemicals across interorgans and interanatomical regions reconfirms the versatility of the Raman endoscopy technique for early detection and diagnosis of cancers based on Raman signals derived from proteins, DNA, and lipids.^{18,23} To further evaluate the clinical implications of interorgan variability, we also employed PLS-DA for cancer diagnosis²⁵ whereby a diagnostic accuracy of 89.3% [sensitivity: 92.6% (162/175); specificity 88.6% (665/751)] and of 94.7% [sensitivity: 90.9% (30/33); specificity 93.9% (216/230)] in the gastric and esophagus, respectively, can be achieved based on the leave-one tissue site-out, cross validation method. The Raman spectral changes associated with cancerous tissue transformation [Figs. 2(a) and 2(b)] were several orders of magnitude larger than interanatomical variability of normal tissue, illustrating the efficacy of *in vivo* Raman endoscopy technique for early cancer diagnosis and detection in the upper GI.

One notes that biochemical modeling by means of NNCLSM is a linear semiquantitative estimation of biochemical compositions (e.g., proteins, DNA, and lipids, etc.) of tissue in the upper GI. Only the most essential biochemical constituents of tissue (i.e., biochemicals largely representing cell nucleus, cell cytoplasm, and intra-/extra-cellular matrices) that are known to undergo alterations during the carcinogenesis process are included in the NNCLSM modeling. The slight discrepancy (of $\sim 10\%$) between the reconstructed Raman spectra (by the NNCLSM modeling) and the measured Raman spectra may be due to the following reasons: 1. the *in vitro* biomolecular compositions and conformations used for NNCLSM modeling may not truly reflect *in vivo* tissue conditions; 2. the basis spectra are not orthogonal, and only a very limited number of biochemicals are included; 3. the NNCLSM biochemical modeling does not take the effect of tissue optical properties (e.g., absorption and scattering coefficients, and anisotropy) of different tissues into consideration; 4. the laser penetration depth may also affect the estimated distributions of the biochemicals in different tissue types.³⁷ Nevertheless, within these limitations, the results of NNCLSM modeling indicate that the clinical Raman endoscopic technique can provide effective biomolecular and morphological information about different anatomical tissue sites of the upper GI tract.

In summary, high-quality *in vivo* Raman spectra can be acquired from different sites of the upper GI (i.e., gastric and esophagus) in real-time by Raman endoscopy technique during clinical endoscopic examinations. Difference spectra together with biochemical modeling suggest that interorgan variability is significant whereas the interanatomical variability of the esophagus and gastric is subtle compared to neoplastic tissue transformation. Cancerous tissues can be identified across interanatomical regions with accuracies of 89.3% and 94.7% in the gastric and esophagus, respectively, demonstrating that the Raman endoscopy technique has promising clinical potential for real-time *in vivo* diagnosis and detection of early cancer in the upper GI at the molecular level.

Acknowledgments

This research was supported by the National Medical Research Council and the Biomedical Research Council, Singapore.

References

- D. M. Parkin, F. Bray, J. Ferlay, and P. Pisani, "Global cancer statistics, 2002," *Ca-Cancer J. Clin.* **55**, 74–108 (2005).
- R. Wong and R. Malthaner, "Esophageal cancer: a systemic review," *Curr. Probl Cancer* **24**, 297–373 (2000).
- C. J. Clark, R. C. Thirlby, V. Picozzi Jr, D. B. Schembre, F. P. Cummings, and E. Lin, "Current problems in surgery: gastric cancer," *Curr. Probl Surg.* **43**, 566–570 (2006).
- A. J. Overhiser and P. Sharma, "Advances in endoscopic imaging: narrow band imaging," *Rev. Gastroenterol. Disord.* **8**, 186–193 (2008).
- M. A. Kara, F. P. Peters, P. Fockens, F. J. ten Kate, and J. J. Bergman, "Endoscopic video-autofluorescence imaging followed by narrow band imaging for detecting early neoplasia in Barrett's esophagus," *Gastrointest. Endosc.* **64**, 176–185 (2006).
- C. H. Liu, B. B. Das, W. L. Sha Glassman, G. C. Tang, K. M. Yoo, H. R. Zhu, D. L. Akins, S. S. Lubicz, J. Cleary, and R. Prudente, "Raman, fluorescence, and time-resolved light scattering as optical diagnostic techniques to separate diseased and normal biomedical media," *J. Photochem. Photobiol. B* **16**, 187–209 (1992).
- C. H. Liu, W. L. Sha Glassman, H. R. Zhu, D. L. Akins, L. I. Deckelbaum, M. L. Stetz, K. O'Brien, J. Scott, and R. R. Alfano, "Near-IR Fourier transform Raman spectroscopy of normal and atherosclerotic human aorta," *Lasers Life. Sci.* **4**, 257 (1992).
- Z. Huang, S. K. Teh, W. Zheng, J. Mo, K. Lin, X. Shao, K. Y. Ho, M. Teh, and K. G. Yeoh, "Integrated Raman spectroscopy and trimodal wide-field imaging techniques for real-time *in vivo* tissue Raman measurements at endoscopy," *Opt. Lett.* **34**, 758–760 (2009).
- A. Mahadevan-Jansen and R. Richards-Kortum, "Raman spectroscopy for the detection of cancers and precancers," *J. Biomed. Opt.* **1**, 31–70 (1996).
- N. Stone, C. Kendall, J. Smith, P. Crow, and H. Barr, "Raman spectroscopy for identification of epithelial cancers," *Faraday Discuss.* **126**, 141–157 (2004).
- Z. Huang, H. Lui, D. I. McLean, M. Korbelik, and H. Zheng, "Raman spectroscopy in combination with background near-infrared autofluorescence enhances the *in vivo* assessment of malignant tissues," *Photochem. Photobiol.* **81**, 1219–1226 (2005).
- S. K. Teh, W. Zheng, D. P. Lau, and Z. Huang, "Spectroscopic diagnosis of laryngeal carcinoma using near-infrared Raman spectroscopy and random recursive partitioning ensemble techniques," *Analyst* **134**, 1232–1239 (2009).
- E. Widjaja, W. Zheng, and Z. Huang, "Classification of colonic tissues using near-infrared Raman spectroscopy and support vector machines," *Int. J. Oncol.* **32**, 653–662 (2008).
- S. K. Teh, W. Zheng, K. Y. Ho, M. Teh, K. G. Yeoh, Z. Huang, "Diagnosis of gastric cancer using near-infrared Raman spectroscopy and classification and regression tree techniques," *J. Biomed. Opt.* **13**, 034013 (2008).
- S. K. Teh, W. Zheng, K. Y. Ho, M. Teh, K. G. Yeoh, and Z. Huang, "Near-infrared Raman spectroscopy for early diagnosis and typing of adenocarcinoma in the stomach," *Br. J. Surg.* **97**, 550–557 (2010).
- S. K. Teh, W. Zheng, K. Y. Ho, M. Teh, K. G. Yeoh, and Z. Huang, "Diagnostic potential of near-infrared Raman spectroscopy in the stomach: differentiating dysplasia from normal tissue," *Br. J. Cancer* **98**, 457–465 (2008).
- S. K. Teh, W. Zheng, K. Y. Ho, M. Teh, K. G. Yeoh, and Z. Huang, "Near-infrared Raman spectroscopy for optical diagnosis in the stomach: Identification of Helicobacter-pylori infection and intestinal metaplasia," *Int. J. Cancer* **126**, 1920–1927 (2010).
- G. Shetty, C. Kendall, N. Shepherd, N. Stone, and H. Barr, "Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus," *Br. J. Cancer* **94**, 1460–1464 (2006).
- C. Kendall, N. Stone, N. Shepherd, K. Geboes, B. Warren, R. Bennett, and H. Barr, "Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus," *J. Pathol.* **200**, 602–609 (2003).
- L. M. Wong Kee Song, A. Molckovsky, K. K. Wang, L. J. Burgart, B. Dolenko, R. J. Somorjai, and B. C. Wilson, "Diagnostic potential of Raman spectroscopy in Barrett's esophagus," *Proc. SPIE* **5692**, 140–146 (2005).
- Z. Huang, H. Zeng, I. Hamzavi, D. I. McLean, and H. Lui, "Rapid near-infrared Raman spectroscopy system for real-time *in vivo* skin measurements," *Opt. Lett.* **26**:1782–1784 (2001).
- Z. Huang, M. S. Bergholt, W. Zheng, K. Lin, K. Y. Ho, M. Teh, and K. G. Yeoh, "In vivo early diagnosis of gastric dysplasia using narrow-band image-guided Raman endoscopy," *J. Biomed. Opt.* **15**, 037017 (2010).
- Z. Huang, S. K. Teh, W. Zheng, K. Lin, K. Y. Ho, M. Teh, and K. G. Yeoh, "In vivo detection of epithelial neoplasia in the stomach using image-guided Raman endoscopy," *Biosens. Bioelectron.* **26**, 383–389 (2010).
- M. S. Bergholt, W. Zheng, K. Lin, K. Y. Ho, M. Teh, K. G. Yeoh, J. B. Y. So, and Z. Huang, "In vivo diagnosis of gastric cancer using Raman endoscopy and ant colony optimization techniques," *Int. J. Cancer* (2010).
- M. S. Bergholt, W. Zheng, K. Lin, K. Y. Ho, M. Teh, K. G. Yeoh, J. B. Y. So, and Z. Huang, "Raman endoscopy for *in vivo* differentiation between benign and malignant ulcers in the stomach," *Analyst* **135**, 3162–3168 (2010).
- D. A. Owen, "Normal histology of the stomach," *Am. J. Surg. Pathol.* **10**, 48–61 (1986).
- K. Takubo, *Pathology of the Esophagus: An Atlas and Textbook*, Springer-Verlag, New York (2009).
- American National Standard for the Safe Use of Lasers: ANSI Standard Z136.1–1986*, American National Standards Institute, Washington D.C. (1986).
- F. P. Incropera and D. P. D. Witt, *Fundamentals of Heat and Mass Transfer*, John Wiley and Sons, New York (1990).
- J. H. Torres, M. Motamedi, J. A. Pearce, and A. J. Welch, "Experimental evaluation of mathematical models for predicting the thermal response of tissue to laser irradiation," *Appl. Opt.* **32**, 597–606 (1993).
- S. Thomsen, "Pathologic analysis of photothermal and photomechanical effects of laser-tissue interactions," *Photochem. Photobiol.* **53**, 825–835 (1991).
- K. G. Yeoh, "How do we improve outcomes for gastric cancer?," *J. Gastroenterol. Hepatol* **22**, 970–972 (2007).
- R. Bender and S. Lange, "Adjusting for multiple testing: when and how?" *J. Clin. Epidemiol.* **54**, 343–349 (2001).
- J. R. Levin, R. C. Serlin, and M. A. Seaman, "A controlled, powerful multiple-comparison strategy for several situations," *Psychol. Bull.* **115**, 153–159 (1994).
- T. D. Wang, G. Triadafilopoulos, J. M. Crawford, L. R. Dixon, T. Bhandari, P. Sahbaie, S. Friedland, R. Soetikno, and C. H. Contag, "Detection of endogenous biomolecules in Barrett's esophagus by Fourier transform infrared spectroscopy," *Proc. Natl. Acad. Sci. U.S.A* **104**, 15864–15869 (2007).
- M. Diem, S. Boydston-White, and L. Chiriboga, "Infrared spectroscopy of cells and tissues: Shining light onto a novel subject," *Appl. Spectrosc.* **53**, 148A–161A (1999).
- P. Thueller, I. Charvet, F. Bevilacqua, M. S. Ghislain, G. Ory, P. Marquet, P. Meda, B. Vermeulen, and C. Depeursinge, "In vivo

- endoscopic tissue diagnostics based on spectroscopic absorption, scattering, and phase function properties," *J. Biomed. Opt.* **8**, 495–503 (2003).
38. R. D. Odze and J. R. Goldblum, *Surgical Pathology of the GI Tract, Liver, Biliary Tract and Pancreas*, Saunders, Philadelphia (2009).
 39. A. Bansal, O. Ulusaarac, S. Mathur, and P. Sharma, "Correlation between narrow band imaging and nonneoplastic gastric pathology: a pilot feasibility trial," *Gastrointest. Endosc.* **67**, 210–216 (2008).
 40. D. Hopwood, K. R. Logan, G. Coghill, and I. A. D. Bouchier, "Histochemical studies of mucous substances and lipids in normal human oesophageal epithelium," *Histochem. J.* **9**, 153–161 (1977).
 41. C. A. Squier and M. J. Kremer, "Biology of oral mucosa and esophagus," *J. Natl. Cancer Inst. Monogr.* **29**, 7–15 (2001).
 42. B. Mayinger, P. Horner, M. Jordan, C. Gerlach, T. Horbach, W. Hohenberger, and E. G. Hahn, "Light induced autofluorescence spectroscopy for the endoscopic detection of esophageal cancer," *Gastrointest. Endosc.* **54**, 195–201 (2001).
 43. R. Picard and D. Cook, "Cross-validation of regression models," *J. Am. Stat. Assoc.* **79**, 575–583 (1984).
 44. K. W. Short, S. Carpenter, J. P. Freyer, and J. R. Mourant, "Raman spectroscopy detects biochemical changes due to proliferation in mammalian cell cultures," *Biophys. J.* **88**, 4274–4288 (2005).