2019

Co-registered photoacoustic and ultrasound imaging of human colorectal cancer

Guang Yang

Eghbal Amidi

William C. Chapman Jr.

Sreyankar Nandy

Atahar Mostafa

See next page for additional authors

Authors

Guang Yang, Eghbal Amidi, William C. Chapman Jr., Sreyankar Nandy, Atahar Mostafa, Heba Abdelal, Zahra Alipour, Deyali Chatterjee, Matthew Mutch, and Quing Zhu

Co-registered photoacoustic and ultrasound imaging of human colorectal cancer

Guang Yang
Eghbal Amidi
William C. Chapman, Jr.
Sreyankar Nandy
Atahar Mostafa
Heba Abdelal
Zahra Alipour
Deyali Chatterjee
Matthew Mutch
Quing Zhu

Co-registered photoacoustic and ultrasound imaging of human colorectal cancer

Guang Yang,[a,†] Eghbal Amidi,[a,†] William C. Chapman Jr.,[b,†] Sreyankar Nandy,[a] Atahar Mostafa,[a] Heba Abdelal,[c] Zahra Alipour,[c] Deyali Chatterjee,[c] Matthew Mutch,[b] and Quing Zhu[a,d,*]
[a]Washington University in St. Louis, Department of Biomedical Engineering, St. Louis, Missouri, United States
[b]Washington University School of Medicine, Department of Surgery, St. Louis, Missouri, United States
[c]Washington University School of Medicine, Department of Pathology and Immunology, St. Louis, Missouri, United States
[d]Washington University School of Medicine, Department of Radiology, St. Louis, Missouri, United States

Abstract. Colorectal cancer is the second most common malignancy diagnosed globally. Critical gaps exist in diagnostic and surveillance imaging modalities for colorectal neoplasia. Although prior studies have demonstrated the capability of photoacoustic imaging techniques to differentiate normal from neoplastic tissue in the gastrointestinal tract, evaluation of deep tissue with a fast speed and a large field of view remains limited. To investigate the ability of photoacoustic technology to image deeper tissue, we conducted a pilot study using a real-time co-registered photoacoustic tomography (PAT) and ultrasound (US) system. A total of 23 *ex vivo* human colorectal tissue samples were imaged immediately after surgical resection. Co-registered photoacoustic images of malignancies showed significantly increased PAT signal compared to normal regions of the same sample. The quantitative relative total hemoglobin (rHbT) concentration computed from four optical wavelengths, the spectral features, such as the mean spectral slope, and 0.5-MHz intercept extracted from PAT and US spectral data, and image features, such as the first- and second-order statistics along with the standard deviation of the mean radon transform of PAT images, have shown statistical significance between untreated colorectal tumors and the normal tissue. Using either a logistic regression model or a support vector machine, the best set of parameters of rHbT and PAT intercept has achieved area-under-the-curve (AUC) values of 0.97 and 0.95 for both training and testing data sets, respectively, for prediction of histologically confirmed invasive carcinoma.

1 Introduction

Photoacoustic imaging (PAI) is an emerging technique that can provide high optical absorption contrast images at reasonable microscale resolution and clinically relevant depths.[1] Several studies have established that optical absorption parameters are important biomarkers directly related to the tissue microvasculature, tumor angiogenesis, or tumor hypoxia.[2–4] In general, PAI is classified into photoacoustic microscopy (PAM) and photoacoustic tomography (PAT).[1] Previously, PAM and photoacoustic endoscopy have demonstrated the capability of detecting human colorectal cancer.[5,6] However, the low imaging speed (limited by the laser repetition rate and scanning scheme), small imaging area, and moderate penetration depth created obstacles for clinical applications.

Compared with PAM, PAT is able to penetrate deeper with a faster data acquisition speed and a larger field of view due to the use of ultrasonic arrays and a wide optical beam. Several studies have demonstrated that a PAT/US dual-modality imaging system can provide anatomical and functional information in tumors,[7–13] but no prior applications in the human distal GI tract have been reported using PAT/US dual-modality imaging.

Adenocarcinoma of the colon and rectum is the second most common malignancy diagnosed globally and the fourth leading cause of cancer mortality, with more than 100,000 new cases diagnosed annually in the U.S.[14,15] Accurate staging and post-treatment surveillance of this prevalent disease are critical because treatment strategies are predicated upon the stage at presentation and response to therapy—in some instances, detailed imaging allows certain patients to avoid surgery altogether. Although colonoscopy and biopsy are the gold-standard diagnostic tests for colorectal cancers,[16] multiple imaging modalities including optical imaging,[17,18] endoscopic ultrasound (EUS), pelvic magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET) are also utilized.

Unfortunately, each of these modalities has critical weaknesses when evaluating colorectal tumors. White light endoscopy only detects macroscopic morphology and provides no functional assessment of the imaged tissue. MRI has limited between-slice resolution and is often unable to differentiate early tumors from benign neoplasia, committing patients to potentially more invasive treatment regimens than needed.[19,20] Monitoring of tumors after chemotherapy and radiation with MRI is often confounded by fibrotic reaction and edema, which can appear similar to residual tumor.[21] CT has poorer resolution of the bowel wall layers in comparison to MRI, subsequently limiting its ability to describe circumferential resection margin status or serosal invasion in locally advanced cases. Additionally, CT also cannot distinguish induration or peritumoral fibrosis from frank malignant disease with a high degree of specificity, further limiting its application in local tumor staging.[19] PET imaging is also plagued by poor resolution, and

*Address all correspondence to Quing Zhu, Email: zhu.q@wustl.edu

†These authors contributed equally to this study.

EUS remains highly user-dependent and unable to resolve small islands of the tumor.[19] Therefore, a critical need exists for precise imaging modalities of colorectal tumors for both staging and therapeutic response evaluations.

PAT, in contrast, uniquely provides functional imaging at high resolution using hemoglobin as an endogenous contrast agent. By detecting the abnormal vasculature that accompanies colorectal malignancies, we hypothesized that this modality might be able to identify malignant or residual tumors, otherwise, undetectable by current clinical imaging. We, therefore, performed the following pilot study to test a real-time co-registered PAT/US system prototype and assess its ability to delineate differences between benign and malignant tissue. To the best of our knowledge, this study is the first utilizing co-registered PAT/US to evaluate human colon samples.

2 Methods

2.1 Human Sample Preparation

Freshly resected colon and rectum samples obtained from patients undergoing surgery at Washington University School of Medicine were imaged immediately after surgery. Patients with known benign neoplasia (polyps) as well as malignancies (adenocarcinoma) were eligible for imaging. Cancer patients who had received preoperative treatment with chemotherapy and /or radiation were also included. The study was approved by the Institutional Review Board at Washington University (#201707066). Informed consent was obtained from all patients. Specimens were obtained from the operating room as previously described.[6]

A total of 23 tissue samples were imaged in the pilot study using the PAT/US system. This included untreated colorectal adenocarcinomas ($n = 12$), precancerous polyps ($n = 6$), colorectal cancer following chemotherapy or radiation and chemotherapy ($n = 4$), and postpolypectomy ($n = 1$). Two treated patients have achieved complete pathological response and two partial response. The majority of patients underwent hemicolectomy for cancer and were found to have malignancy on histologic analysis (Table 1).

2.2 Co-Registered Ultrasound-Guided Photoacoustic Tomography System

Details of the real-time, co-registered PAT/US system used in this study were discussed previously.[11,22] The system consists of three main parts: a Ti:sapphire laser (Symphotics TII, LS-2134, Symphotics, Camarillo, California) optically pumped with a Q-switched Nd: YAG laser (Symphotics TII, LS-2122), an optimized optical fiber-based light delivery system,[23] and a commercial US system (EC-12R, Alpinion Medical System, Republic of Korea) used for acquiring the corresponding US and PAT data. With this system, pulsed laser light (pulse duration: 10 ns, pulse repetition rate: 15 Hz, 20 mJ/pulse at 750 nm wavelength) was delivered to tissue placed on a two-dimensional motorized scanning stage. B-scan images were acquired for four wavelengths (730, 780, 800, and 830 nm) at each area of imaged tissue among the 23 specimens. The overall scanning area varied from 1 to 3 cm while image acquisition time for a region of interest (ROI) was 12 to 15 s with a frame rate of 15 frames per second. For most of the samples, care was taken to acquire images from abnormal tissues as well as corresponding normal areas, as marked by a trained pathologist.

2.3 Extraction of Functional, Spectral, and Textural Features

Several functional, spectral, and textural features were extracted from the PAT and US data and images as given in Table 2.

2.3.1 Functional features

The relative oxy-hemoglobin ($rHbO_2$) and deoxy-hemoglobin (rHb) at each pixel can be calculated using the following equations:

$$rHbO_2(r, \theta) = \overline{C(r, \Delta)} HbO_2(r, \theta), \qquad (1)$$

$$rHb(r, \theta) = \overline{C(r, \theta)} Hb(r, \theta), \qquad (2)$$

where $\overline{C(r, \theta)} = \Gamma C_0(r, \theta) \varnothing(r, \theta)$, Γ is the tissue's Grüneisen parameter, $C_0(r, \theta)$ is the system acoustic operator, and $\varnothing(r, \theta)$ is the local fluence, which can be approximated as wavelength independent at the narrow wavelength window we have used. Based on these equations, deriving $rHbO_2$ and rHb requires a known tissue fluence distribution. Since this distribution is difficult to determine in human tissue due to wide variation in composition, we computed relative $rHbO_2$ and rHb values instead. By summing the $rHbO_2$ and rHb at each pixel, the relative total hemoglobin ($rHbT$) for each pixel is computed; the average $rHbT$ for an ROI was then calculated by averaging the $rHbTs$ of all pixels in that ROI with a value at least half of the maximum $rHbT$. All PAT images in the co-registered US and PAT images were $rHbT$ without any normalization.

2.3.2 Spectral features

Ultrasound images were employed to select a proper ROI corresponding to the lesion for PAT spectral feature calculations.[9,24] First, PAT beam lines with a maximum value close to the background noise level of our co-registered US/PAT system (60 mV) were ignored. The rest of the beam lines were gated by a hamming window, and then their FFT in −10-dB frequency range were calculated. Moreover, to cancel the frequency response of the transducer and electrical receiving system,[24] the spectra of PAT beams were normalized to the spectra of an approximate point-like target (a 250-μm black thread orthogonal to PAT imaging place with a varied distance to transducer from 0.5 to 7 cm and a step of 0.25 cm).

After calibrating our data, each of the calibrated PAT spectra was fitted linearly. The mean spectral slope (SS), midband fit (MBF), and 0.5-MHz spectral intercept [0.5-MHz SI (PAT)] were then calculated (Fig. 1). We chose 0.5-MHz spectral intercept as a feature instead of 0 MHz because the lower bound of our transducer in PAT mode is ∼0.5 MHz.

US spectral features were also calculated. To do this, similar method as PAT spectral features extraction was followed. The only differences were: first, the analysis was performed in the frequency range of 3.5 to 7 MHz, which is the −10-dB frequency range of the transducer in US mode. Second, the calibration was performed using a reference gelatin-based phantom constructed in our lab.[25]

2.3.3 PAT image features

After visual inspection of PAT frames of malignant and normal colon samples, we noticed that the textures of these images

Table 1 Summary of specimens.

Patient ID	Surgery	Pathology
1	Total colectomy	Moderately differentiated adenocarcinoma (T3)
2	Right hemicolectomy	Tubular adenoma (precancerous polyp)
3	Right hemicolectomy	Moderately differentiated adenocarcinoma (T2)
4	Sigmoid colectomy	Treated moderately differentiated adenocarcinoma (T3); postchemotherapy
5	Right hemicolectomy	Tubular adenoma (precancerous polyp)
6	Low anterior resection	Complete pathologic response-no residual tumor after radiation and chemotherapy
7	Left colectomy	Moderately differentiated adenocarcinoma (T2)
8	Low anterior resection	Complete pathologic response-no residual tumor following radiation and chemotherapy
9	Low anterior resection	Tubulovillous adenoma (precancerous polyp)
10	Right hemicolectomy	Moderately differentiated adenocarcinoma (T3)
11	Right hemicolectomy	Moderately differentiated adenocarcinoma (T2)
12	Right hemicolectomy	Moderately differentiated adenocarcinoma (T4)
13	Right hemicolectomy	Tubular adenoma (precancerous polyp)
14	Left colectomy	Moderately differentiated adenocarcinoma (T4)
15	Left hemicolectomy	Tubulovillous Adenoma (precancerous polyp)
16	Low anterior resection	No residual tumor following prior polypectomy[a]
17	Sigmoidectomy	Moderately differentiated adenocarcinoma (T3)
18	Transverse colectomy	Moderately differentiated adenocarcinoma (T2)
19	Low anterior resection	Treated moderately differentiated adenocarcinoma (T3); postradiation and chemotherapy
20	Low anterior resection	Moderately differentiated adenocarcinoma (T3)
21	Total colectomy	Moderately to poorly differentiated adenocarcinoma (T3)
22	Right hemicolectomy	Moderately differentiated adenocarcinoma (T3)
22	Sigmoid colectomy	Tubular adenoma with a small invasive moderately
23	Right hemicolectomy	Differentiated adenocarcinoma (mixed polyp and adenocarcinoma)

Note: T is the primary tumor depth of invasion per TNM guidelines.
[a]No residual tumor was found after polypectomy. We have grouped this case with the complete responders.

looked different between the two types of samples. To confirm our observation, we extracted PAT image features from available image frames. To do so, a proper ROI was first chosen. To find the center of this ROI, the region surrounding the lesion was determined based on the US image, and the Radon transforms at the two angles of 0 deg and 90 deg of the PAT image in this region were calculated. Each of these Radon transforms was then normalized to its own maximum values and a Gaussian curve was fitted to each of them. The center of the square ROI where the image analysis was performed was determined by the means of these two Gaussian curves, and its size was assumed to be 1 cm for all cases (Fig. 2).

Textural features of the normalized PAT images were calculated in the specified ROI.[24] The first step in calculating these features is to construct a gray-level co-occurrence matrix (GLCM).[26] GLCM quantifies how the pixels are connected in the image. The size of this matrix was chosen as 16×16 pixels. The value of pixel (i, j) of this matrix was chosen to be the number of times that gray levels i and j are adjacent to each other in the PAT image. Note that we assumed that the two gray levels g1 and g2 are adjacent if g1 is positioned at the immediate left of g2. After constructing the GLCM matrix, four textural features were calculated for each PAT image frame using the following equations:

$$\text{contrast} = \sum_{|i-j|=0}^{N-1} |i-j|^2 \sum_{i=1}^{N} \sum_{j=1}^{N} c(i,j), \tag{3}$$

Table 2 Abbreviations.

Abbreviation	Description
rHbT	Relative total hemoglobin
SS (PAT)	Mean PAT spectral slope
0.5-MHz SI (PAT)	0.5-MHz spectral intercept from PAT spectra
MBF (PAT)	Mid-band fit from PAT spectra
SS (US)	Mean US spectral slope
0.5-MHz SI (US)	0.5-MHz spectral intercept from US spectra
MBF (US)	Mid-band fit from US spectra
Sig_rad	Standard deviation of the mean radon transform
Homogeneity	The homogeneity of image textures
Energy	The grayscale distribution homogeneity of images and texture crudeness
Contrast	The sharpness of images and the depth of texture grooves
Correlation	The consistency of image texture

$$\text{correlation} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}(i-\mu_i)(j-\mu_j)c(i,j)}{\sigma_i \sigma_j}, \qquad (4)$$

$$\text{energy} = \sum_{i=1}^{N}\sum_{j=1}^{N} c(i,j)^2, \qquad (5)$$

$$\text{homogeneity} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{c(i,j)^2}{1+|i-j|}, \qquad (6)$$

where $c(i,j)$ is the value of the (i,j) pixel of the GLCM, N is the dimension of this matrix, and σ and μ are the standard deviation and mean for row i or column j of the GLCM.

The standard deviation of the mean radon transform (Sig_rad) was the last image feature that was calculated in this study. To calculate this feature, the radon transform of the non-normalized PAT image at angles 0 deg to 90 deg with a step of 1 deg were calculated and an average was taken over these transforms. Then a Gaussian curve was fitted to this mean radon transform and the standard deviation of this curve was measured.

2.4 Feature Selection and Classification

A two-step approach was used to select features most likely to differentiate normal from untreated malignant tissue. In the first step, all previously discussed PAT/US features were tested in univariate analysis between untreated cancer and normal regions, and p values were generated by two-sample two-sided Student's t tests. The features where $p > 0.05$—which we concluded *a priori* not to be significantly associated with malignancy—were excluded from the classification model (Table 3).

Next, logistic a general logistic mode (GLM) and a support vector machine (SVM) were used to evaluate the strength of association of each feature with the ultimate tissue diagnosis, and a prediction model was then constructed with significant covariates. In total, 18 areas selected from 18 specimens and 12 malignant areas from 12 untreated cancer specimens were used to construct and evaluate the prediction models. Out of these, 12 normal and 8 malignant areas were used for prediction
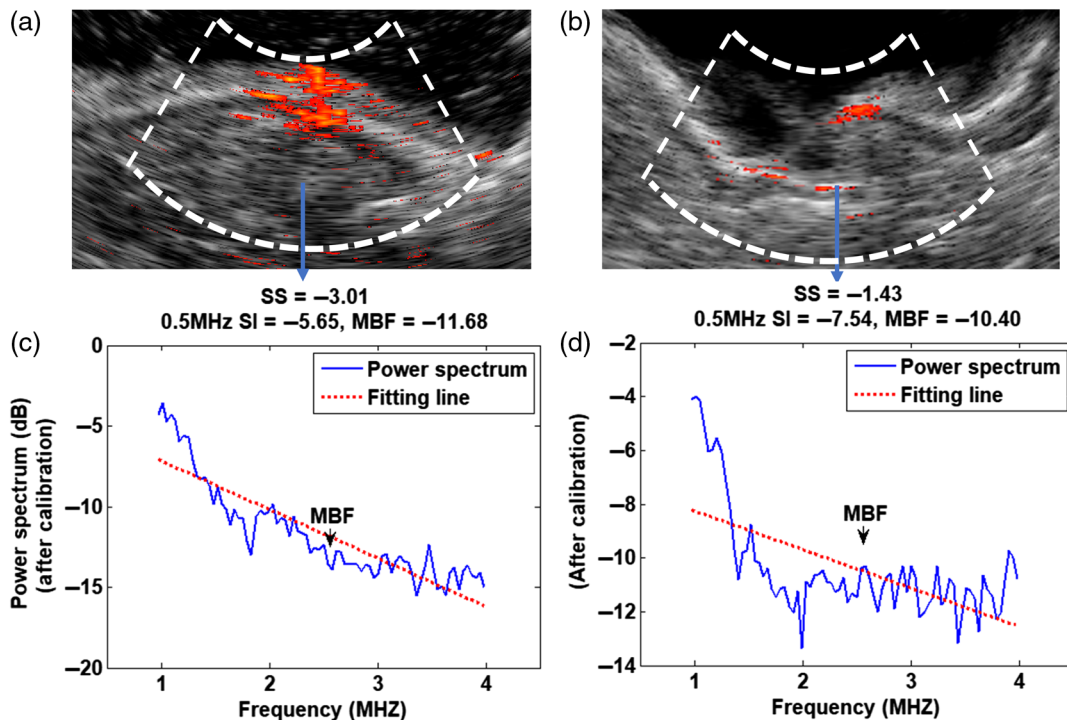
Fig. 1 (a), (b) Co-registered rHbT and US images of (a) a cancerous and (b) a normal colon sample. (c), (d) Calibrated PAT power spectra along with their fitted lines in the regions marked with the angular dashed lines.
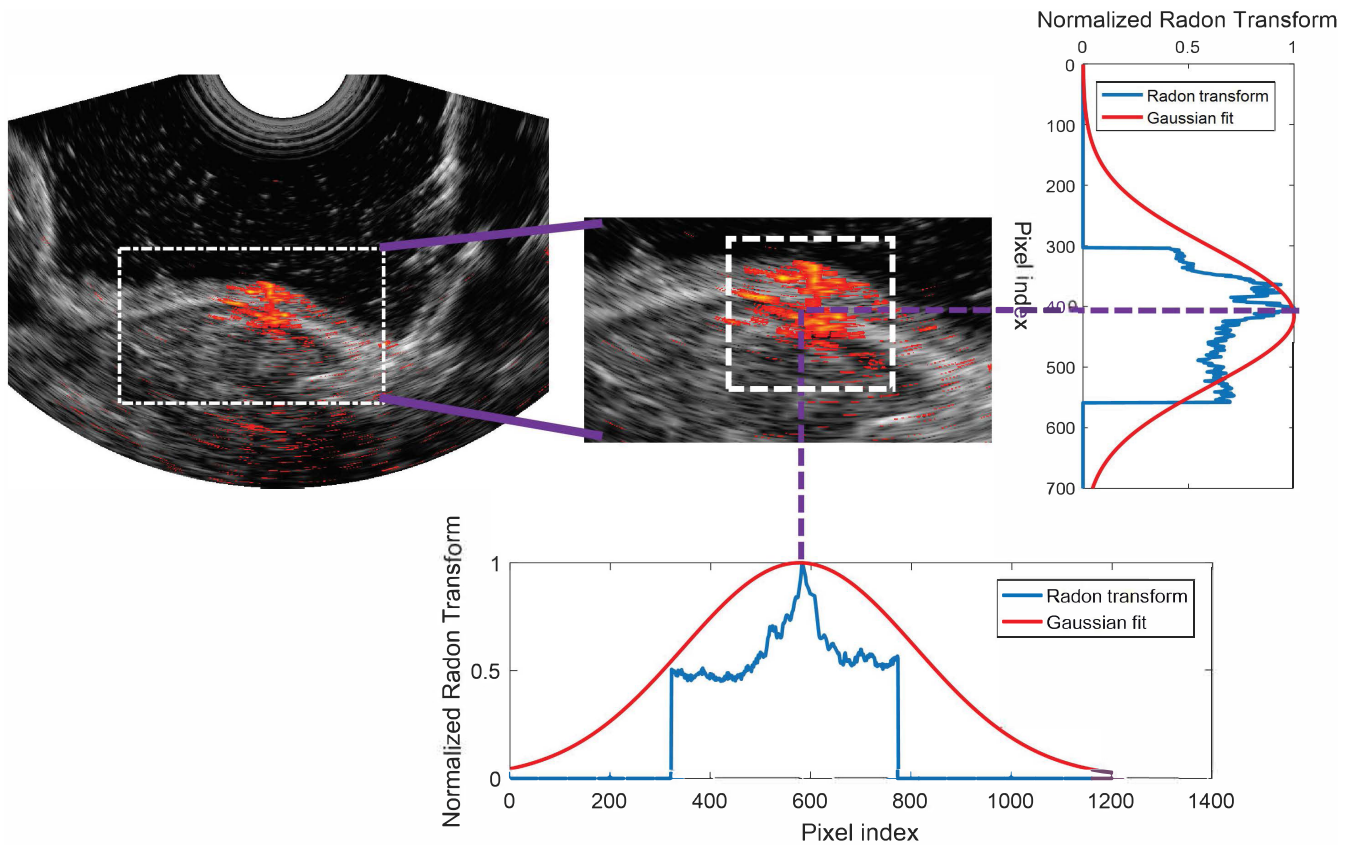
Fig. 2 ROI selection for image analysis. The region covering the lesion is determined (left), and the normalized Radon transforms at 0 and 90 deg are calculated. A Gaussian curve is then fitted to each Radon transform. The center of the square ROI where the image analysis is performed is determined by the means of these two Gaussian curves and its size is 1 cm.

model derivation and the rest (6 normal and 4 malignant areas) for internal model validation. The receiver operating curve (ROC) and the area under the curve (AUC) were used to evaluate the accuracy of the model. Finally, a second prediction model was constructed without rHbT to determine how limiting the PAT/US device to a single wavelength would affect identification of malignancies.

Table 3 Significance testing of individual covariates as related to tissue diagnosis.

| Feature	p-value
rHbT	<0.001
SS (PAT)	<0.001
0.5-MHz SI (PAT)	0.002
0.5-MHz SI (US)	0.01
Homogeneity	0.01
Energy	0.02
Sig_rad	0.03
MBF (PAT)	0.12
MBF (US)	0.23
SS (US)	0.55

3 Results

3.1 Qualitative Analysis: Baseline Characteristics of US and PAT Images

The colorectal tissues are composed primarily of fluid, lipid, collagen, and muscle. The general architecture (from superficial to deep) in a normal specimen is mucosa (fluid-filled cells surrounded by lipid bilayers), submucosa (largely composed of extracellular collagen matrix and some muscle fibers), muscularis propria (muscle), and adipose tissue (lipid). In malignancy, the individual cell types are similar but the architecture is distorted as cancerous cells of mucosal origin penetrate into the deeper layers of the organ. As these cells invade, the organized structure of the tissue is lost.

Figure 3 shows specimen photographs, US images, co-registered PAT/US rHbT maps as well as histologic images from two representative regions of normal colon samples [Figs. 3(a)–3(d) and 3(e)–3(h)] and two colorectal malignancies [Figs. 3(i)–3(l) and 3(m)–3(p)]. The white arrows indicate the scanning direction along which B-scans were recorded at four different wavelengths (the imaging plane is perpendicular to the scanning direction). In the standalone ultrasound images, the normal layered structure of the colorectal wall is clearly delineated

Normal areas
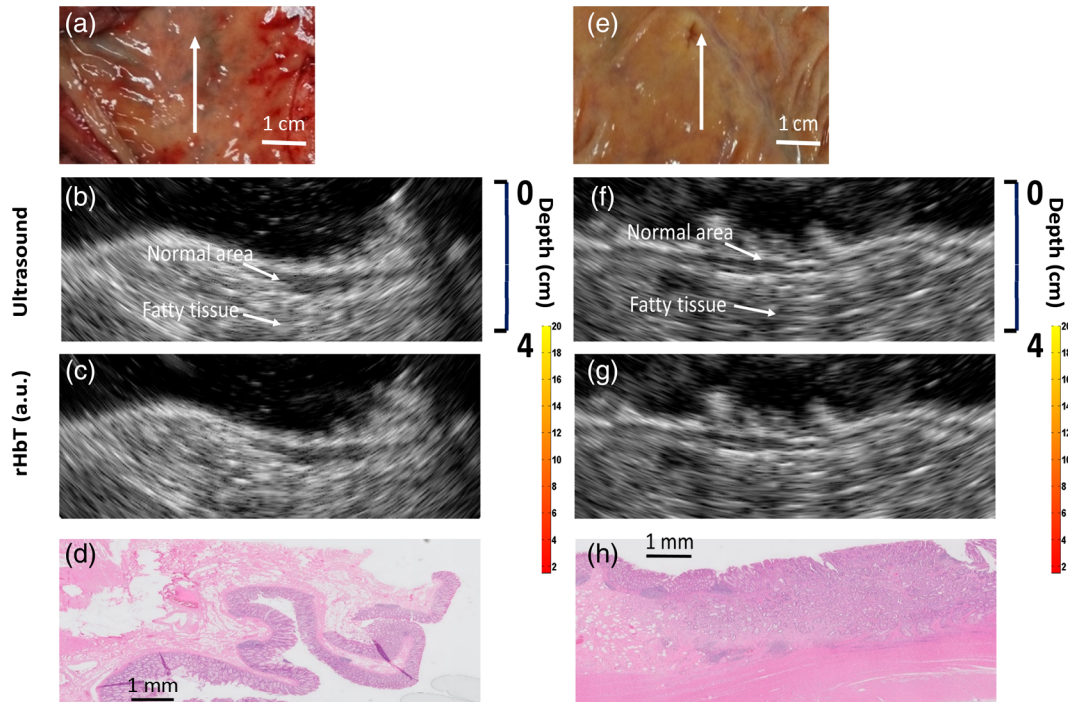
Malignancies
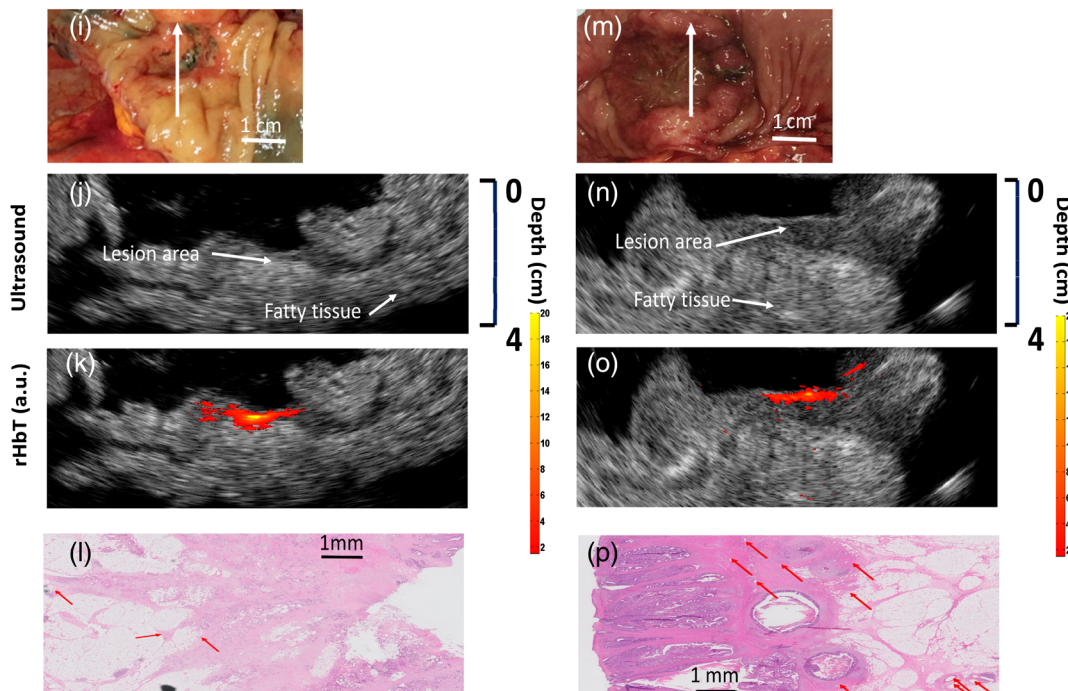
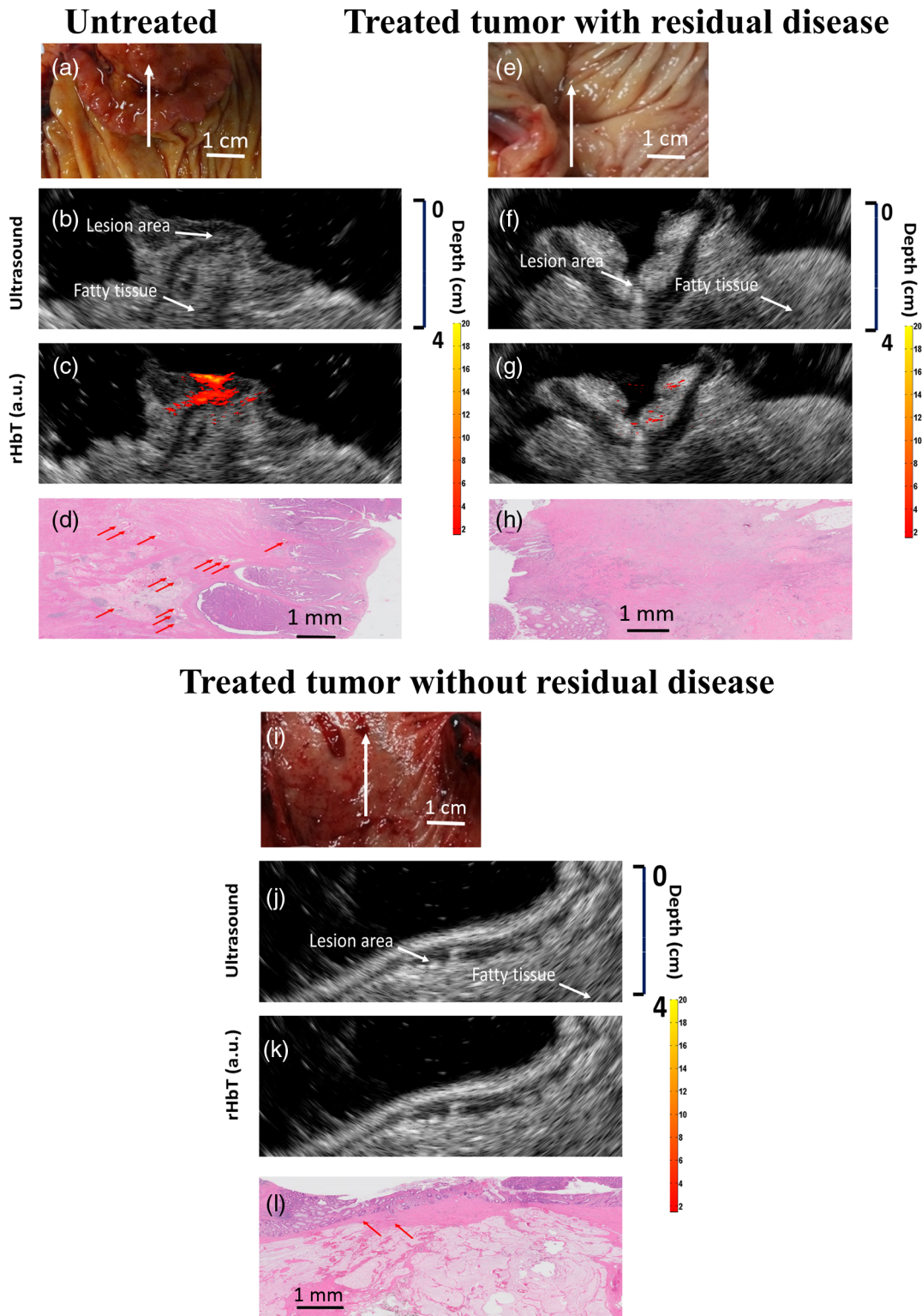**Fig. 3** Color photograph, US image, rHbT map, and H&E image from representative areas of (a)–(h) two normal regions and (i)–(p) two malignant regions of pretreatment colorectal cancer tissue. Red arrows identify blood vessels within the histologic images.

[Figs. 3(b) and 3(f)]. In the presence of malignancy, however, this organized structure is distorted by the tumor and loses the clear delineation of mucosal, submucosal, and muscular layers. These findings mirror the differences in histology among the specimens; in contrast to the ordered layering of the normal colonic wall [images 3(d) and 3(h)], the tumors appear disorganized with destruction of the underlying colonic architecture [images 3(l) and 3(p)].

Additionally, the rHbT maps computed from coregistered PAT/US images of benign regions show significantly lower rHbT signal [Figs. 3(c) and 3(g)] compared to the malignant lesions [3(k) and 3(o)]. As demonstrated in these representative

Fig. 4 Color photograph, US image, rHbT map, and H&E image from representative areas of (a)–(d) a pretreatment colorectal cancer, (e)–(h) a post-treatment colorectal cancer tissue with residual disease, and (i)–(l) a post-treatment colorectal cancer tissue without residual disease.

images, normal tissue was found to have almost no detectable rHbT signal. In contrast, malignant tissue showed much higher concentrations of hemoglobin around the tumor bed. Again, these findings appear corroborated by histologic examination. In comparison to the relative paucity of large blood vessels in normal tissue, the malignancies were more vascular and contained large blood vessels [red arrows in images 3(l) and 3(p)].

It is interesting to note that fatty tissues have limited PAT signals in the outer portions of the specimens. This is not surprising since we are specifically targeting hemoglobin—which is not concentrated in fatty tissue—as our chromophore of
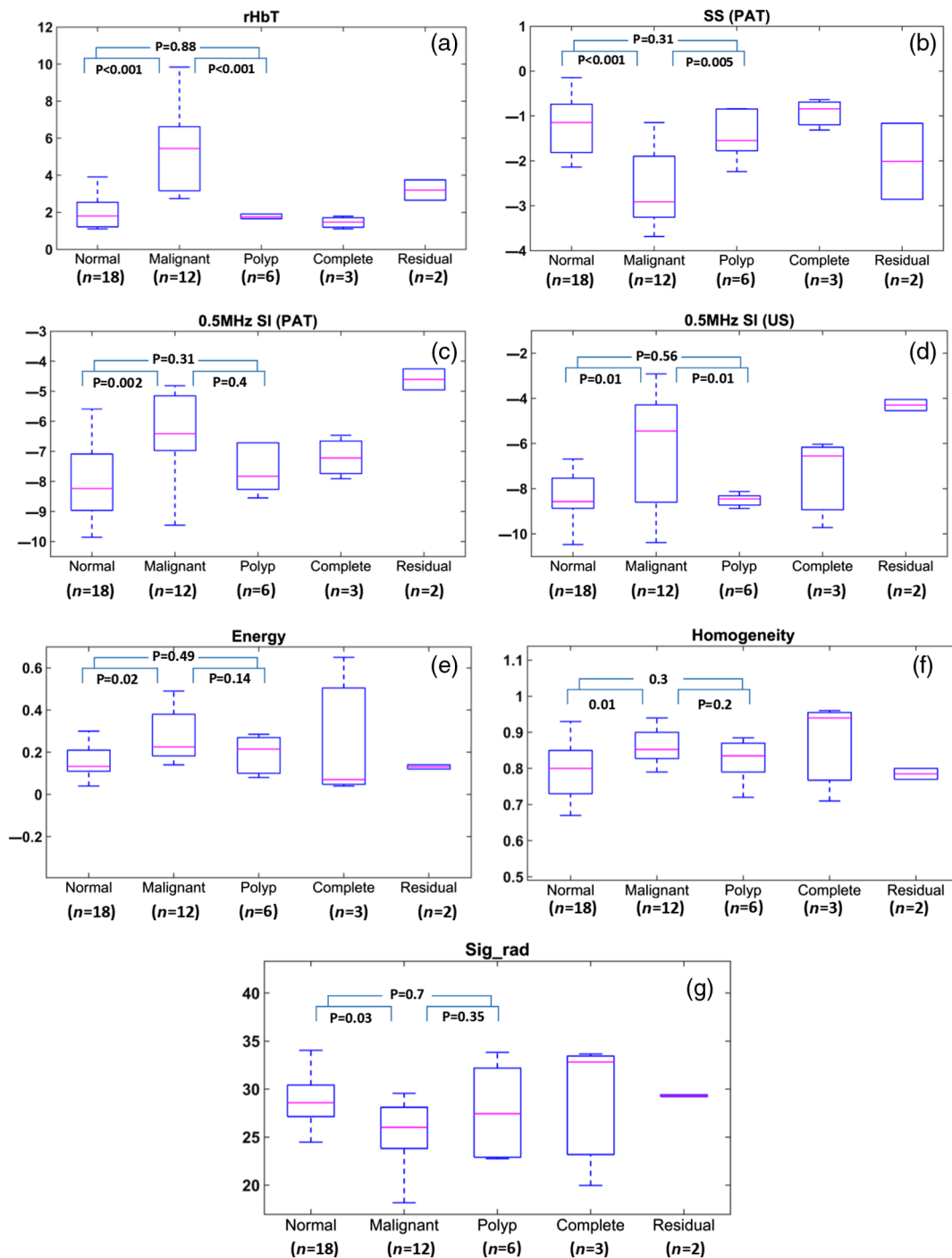
Fig. 5 Boxplots of (a) total hemoglobin, (b) the mean SS from PAT spectra, (c) 0.5-MHz spectral intercept from PAT spectra, (d) 0.5-MHz spectral intercept from US spectra, (e) energy from the second-order statistics of PAT images, (f) homogeneity from the second-order statistics of PAT images, and (g) standard deviation of the mean radon transform.

interest and therefore image within the 730- to 830-nm wavelength range. Additionally, all PAT images are displayed with the same dynamic range of -10 dB, so anything below this level is not displayed. The fatty tissue, due to its lack of vascular structures, falls below this range.

3.2 Evaluation of Treated Tumors

Figure 4 shows corresponding images from a representative untreated colon cancer sample [Figs. 4(a)–4(d)], a colon tumor treated with preoperative chemotherapy [Figs. 4(e)–4(h)], and a rectal malignancy that received radiation and chemotherapy prior to surgical resection [Figs. 4(i)–4(l)]. The untreated tumor displayed findings consistent with other untreated colorectal cancers: loss of layered wall structure, increase in rHbT signal, and increased vascularity throughout the tumor bed. However, treatment appears to reverse these changes. For example, PAT signal appears to diminish with chemotherapy [image 4(g)] and even disappear altogether with complete destruction of the tumor [image 4(k)]. Additionally, ultrasound imaging

demonstrated a return to the normal wall structure with complete tumor destruction [image 4(j)]. Histologic comparison among specimens also correlated with these findings; reduction in vasculature along with return to a semiorganized mucosal structure was noted throughout the treated specimens.

3.3 Quantitative Analysis

In addition to the above qualitative comparisons, a total of 23 areas obtained from 12 untreated malignant tumors, 6 polyps (one has a small invasive component), 2 post-treatment complete responders, 1 no residual tumor cell following prior polypectomy, and 2 post-treatment nonresponders, as well as 18 normal areas from specimens of normal regions were used for quantitative feature extraction. Thus a total of 41 areas were used in Fig. 5. Note that one tumor area was selected from each specimen. Five specimens did not have normal regions or normal regions were too close to the tumor for imaging, therefore, 18 normal areas were selected from 18 specimens. All tumor and normal regions were identified by the attending pathologist.

Figures 5(a)–5(g) show the boxplots of the seven features calculated from the functional, spectral, and image differences between the different types of colorectal tissue. The n number given in the plots corresponds to the total number of areas. The malignant regions demonstrated elevated rHbT, 0.5-MHz SI (PAT), and 0.5-MHz SI (US) score smaller (less negative) compared to normal and precancerous regions. For SS (PAT), the malignant regions score below normal and precancerous polyps. Treated tumors with complete response were found to have similar scores to normal tissue, whereas treated regions with residual cancer have scores similar to untreated cancers. Due to the limited number of treated cancers, statistics were not performed for these two treated categories.

To distinguish untreated malignant from normal colon tissues, GLM and SVM classifiers were established. These classifiers were developed using the independent features with a p-value <0.05 between malignant and normal colon tissues. To determine if two features are independent, a spearman's correlation was calculated between each pair of features (Table 4). To train each classifier, we first used the feature with the lowest p-value and then added other features to the feature set one by one. We continued inclusion of the features to the feature set until no increase in the AUC value for the testing data set was observed. We found that when rHbT is included in the feature set, the best performance of both GLM and SVM classifiers (the highest AUC value for the testing data set) is achieved when rHbT and 0.5-MHz SI (PAT) are employed to train the classifier although SS (PAT) has a lower p-value than 0.5-MHz SI (PAT). Adding other features did not improve the AUC for the testing data set.

Figure 6 shows the ROC curves and AUC values of the training (left) and testing (right) data sets using GLM (top) and SVM (bottom) classifiers. As shown in this figure, when the features set include just rHbT, the AUC value for the training and testing data sets are 0.95 and 0.93 for both classifiers, respectively. Adding 0.5-MHz SI (PAT) to the features set, results in a significant improvement in the AUC values for both training and testing data sets (0.97 and 0.95 for the training and testing data sets for both classifiers, respectively). The three image features (Sig_rad, homogeneity, and energy) did not improve the AUC values for both training and testing data sets.

Finally, the performance of GLM (top) and SVM (bottom) classifiers without rHbT (the single-wavelength model) are presented in Fig. 7. Note that although the difference between some of the PAT image features in malignant and normal samples is statistically significant, none of these features improve the AUC for the testing data sets. The best performance of GLM classifier is achieved when the only spectral feature of SS(PAT) is included in the feature set. The best performance of SVM classifier is achieved when spectral features of SS(PAT), 0.5-MHz SI(PAT), and 0.5-MHz SI(US) are included in the feature set. The testing AUC in this case is 0.89 for the GLM classifier and 0.91 for the SVM classifier.

4 Discussion and Summary

In this pilot study of co-registered ultrasound and PAT, we found significant qualitative and quantitative differences between malignant tumors and normal tissue within human colorectal specimens. Specifically, the parameters rHbT, 0.5-MHz SI (PAT), 0.5-MHz SI (US), and SS (PAT) differ between the two tissue types imaged, suggesting that PAT may be able to differentiate malignant from normal tissue in the colon and rectum. Combined with the PAT system's tissue penetration depth of over 4 to 5 cm (depending on the background tissue optical properties), these findings suggest that PAT may be able to augment extant radiographic technology in the diagnosis, management, and surveillance of colorectal cancer.

As demonstrated by Xu et al.[27] and Kumon et al.,[28] PAT spectral features are related to the size and concentration of the optical absorbers. The slope decreases (more negative) as PA absorber sizes increase and the intercept increases (less negative) as the sizes and concentrations of the absorbers increase. We believe that malignant lesions have larger absorber sizes and higher concentrations compared with normal colorectal tissues

Table 4 The correlation between significant features used in this study.

	SS (PAT)	0.5-MHz SI (PAT)	0.5-MHz SI (US)	Homogeneity	Energy	Sig_rad
rHbT	0.65	0.45	0.27	0.34	0.41	0.37
SS (PAT)		0.67	0.23	0.42	0.46	0.29
0.5-MHz SI (PAT)			0.21	0.49	0.41	0.4
0.5-MHz SI (US)				0.41	0.37	0.31
Homogeneity					0.79	0.91
Energy						0.82
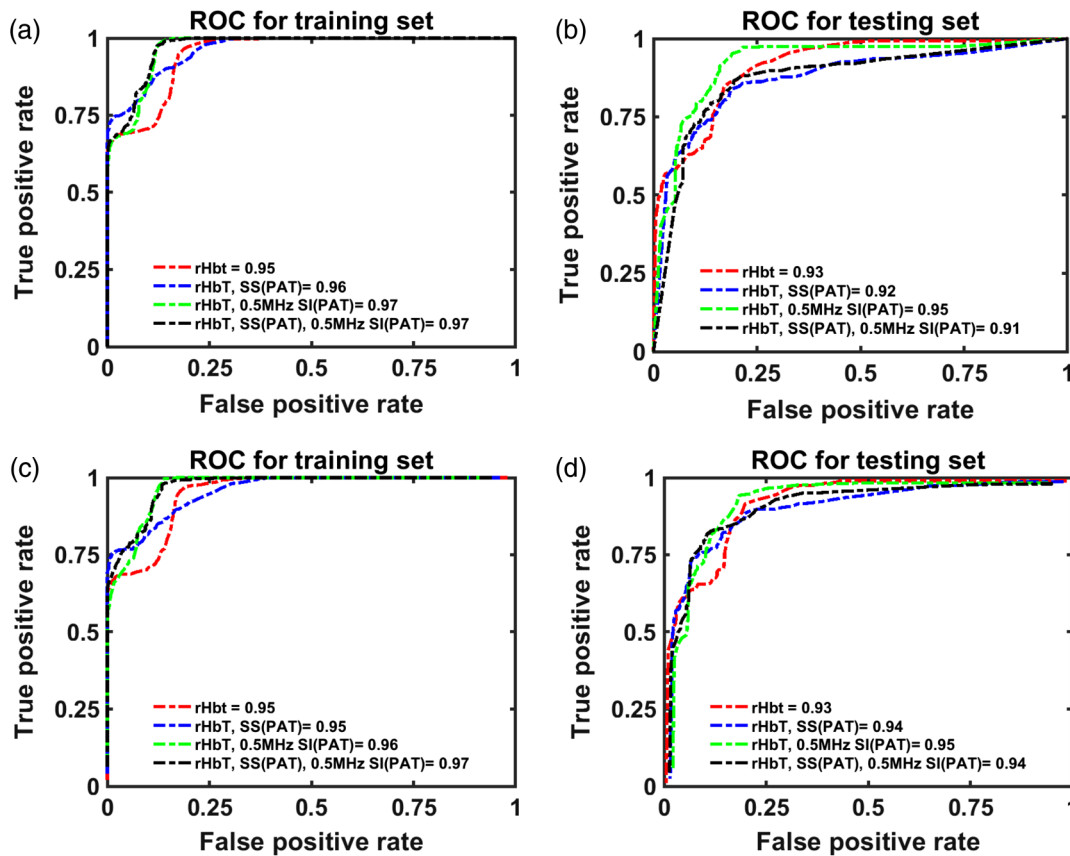
Fig. 6 The ROC curve and their associated AUC values for the training and testing data sets in the presence of rHbT in the feature set. (a), (b) GLM classifier performance. (c), (d) SVM classifier performance.

due to their increased microvessel networks. As introduced by Lizzi et al.,[29] US SS depends on acoustic scatter size, whereas spectral intercept depends on scatter sizes, concentrations, and acoustic impedances of tissue scatter matrix. These parameters have been found valuable to characterize liver, eye,[29] prostate,[30] and breast lesions.[31] We believe that the distorted tissue architecture and abundance of cancerous cells are the source of the US spectral contrast between malignant and normal colorectal tissues. However, the findings of the PAT and US spectral features of colorectal diseases may or may not be applicable to diseases of other organs.

Several technical limitations must be considered with our data. First, we imaged colorectal specimens obtained from routine surgeries and these tissues were typically with large pathologic components that often appeared malignant by visual inspection after specimens were open. These lesions may or may not need advanced PAT and US features for diagnosis. However, these lesions are excellent examples for identifying PAT and US feature characteristics that differ between cancerous and normal tissue. With this information known, we can target less obvious lesions as we look to test the utility of the device in identifying cancer margins and residual tumors after chemoradiation treatment in patient.

The second limitation of this study is the low image resolution of our prototype. The image resolution is only ~250 μm due to the commercial endocavity ultrasound transducer array (6-MHz central frequency, 80% bandwidth). Because this resolution will impact future clinical applications of the device, we plan to upgrade the ultrasound system with a transducer

array of more than 15 MHz to address this problem in future studies. Third, hemoglobin oxygen saturation (sO2) was not calculated in this study since all specimens were imaged after resection, resulting in significantly altered oxygen saturation compared to normal living tissue. sO2 is a significant biomarker for characterization of cancer[11] and assessment of treatment response.

Third, the limited sample size could lead to overfitting of the classifiers if enough care is not taken to develop the classifiers. As a rule of thumb, overfitting is least possible to occur if the number of samples is 10 times or higher than the number of independent predictors.[32] Based on this rule, as we have a total of 30 samples (18 normal colorectal tissues and 12 untreated malignant colorectal tissues) for ROC analysis, the maximum number of the predictors that should be used to avoid overfitting would be three. Figure 6 shows that when rHbT is present in the feature set, the best performances of both GLM and SVM are achieved when rHbT and 0.5-MHz SI (PAT) are the only features used to train the classifiers. Adding SS (PAT) to these feature sets neither changes the value of the AUC for training data sets, nor increases the AUC for the testing data set. Moreover, when rHbT is not included in the feature set, employing the combination of SS (PAT), 0.5-MHz SI (PAT), and 0.5-MHz SI (US) features for developing the classifiers would result in the best performance of SVM classifier and SS (PAT) only would result in the best performance of GLM classifier (Fig. 7). Although adding Sig_rad increases the AUC value for the training data set, it decreases the AUC values for the testing data set in both classifiers. This would mean that our classifiers have
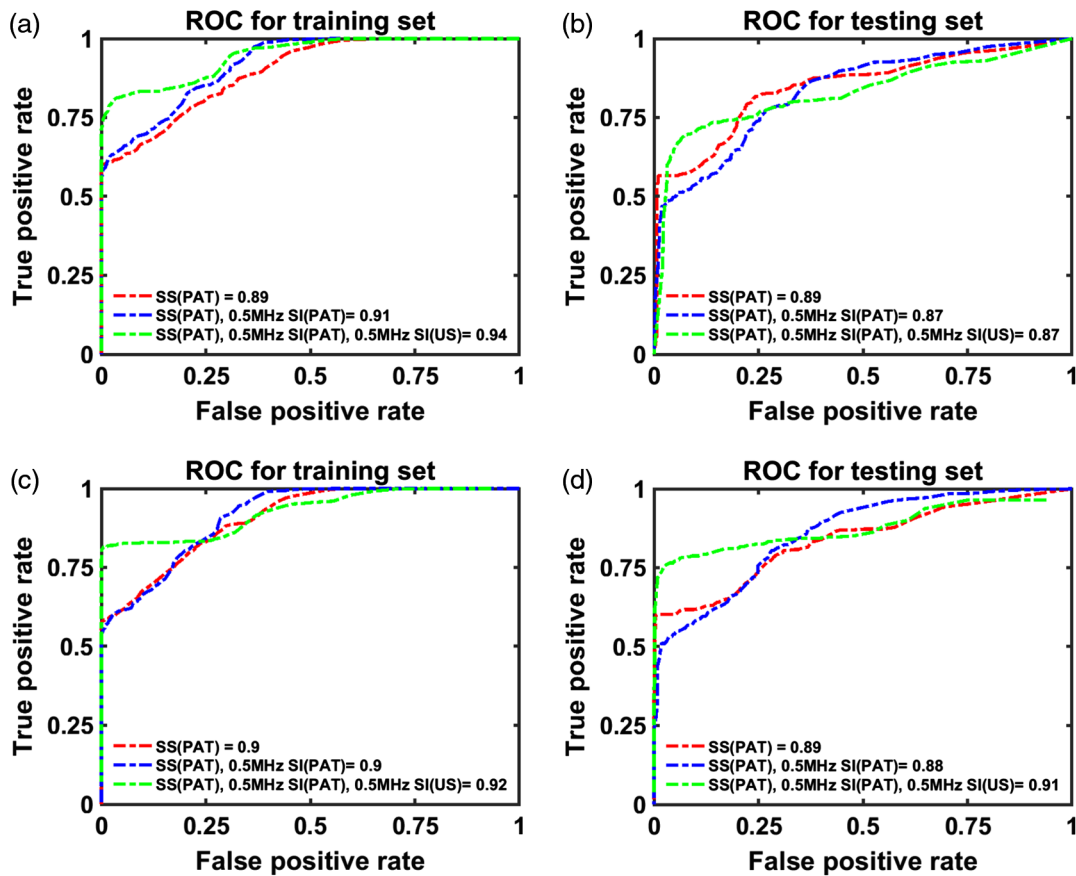
Fig. 7 The ROC curve and their associated AUC values for the training and testing data sets in the absence of rHbT in the feature set. (a), (b) GLM classifier performance. (c), (d) SVM classifier performance.

most probably been overfitted when four features have been used. In this study, to further protect our classifiers from overfitting, repeated rounds (100 times) of cross validation were applied by randomly selecting 2/3 of the samples for training and 1/3 of the samples for testing. The average ROC and AUC values were reported as the results.

In summary, a real-time co-registered PAT/US system was used to image and characterize colorectal masses *ex vivo* in this pilot study. Twenty-three colon and rectum samples (nineteen colon and four rectums) were imaged, rHbT was computed from four wavelength data, and seven quantitative features were extracted from PAT and US power spectra and images. In pre-treated malignant colorectal tumors, we found the cross-section structure to be highly disorganized with a significantly higher rHbT concentration compared to normal and precancerous regions. We performed classifications on the malignant and normal colon regions using GLM and SVM classifiers both with and without tHb in the feature set. When rHbT was employed to construct the classifiers with 0.5-MHz SI (PAT), GLM and SVM classifiers achieved optimal AUC values for the training and testing data sets (0.97 and 0.95, respectively). The small number of treated tumors included in this dataset limits the statistical power of the analysis, but the functional, spectral, and image parameters do appear more similar to normal colorectal tissue in tumors that have experienced complete responses compared to partial responders. These results indicate potential of using PAT/US for future cancer screening and post-treatment surveillance of the colon and rectum. Moving forward, we plan

to increase the resolution of our system using a high-frequency US array and then adapt the technology to an endorectal probe, which will allow us to test the functional and spectral feature differences in *in vivo* human tissue.

Disclosures

No potential conflicts of interest to disclose.

Acknowledgments

References

1. L. V. Wang and S. Hu, "Photoacoustic tomography: in vivo imaging from organelles to organs," *Science* **335**(6075), 1458–1462 (2012).
2. N. Weidner et al., "Tumor angiogenesis and metastasis—correlation in invasive breast carcinoma," *N. Engl. J. Med.* **324**(1), 1–8 (1991).
3. P. Vaupel, F. Kallinowski, and P. Okunieff, "Blood flow, oxygen and nutrient supply, and metabolic microenvironment of human tumors: a review," *Cancer Res.* **49**(23), 6449–6465 (1989).
4. M. Hockel and P. Vaupel, "Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects," *J. Natl. Cancer Inst.* **93**(4), 266–276 (2001).

5. Y. Yuan, S. Yang, and D. Xing, "Preclinical photoacoustic imaging endoscope based on acousto-optic coaxial system using ring transducer array," *Opt. Lett.* **35**(13), 2266–2268 (2010).

6. X. Leng et al., "Feasibility of co-registered ultrasound and acoustic-resolution photoacoustic imaging of human colorectal cancer," *Biomed. Opt. Express* **9**(11), 5159–5172 (2018).

7. H. S. Salehi et al., "Coregistered photoacoustic and ultrasound imaging and classification of ovarian cancer: *ex vivo* and *in vivo* studies," *J. Biomed. Opt.* **21**(4), 046006 (2016).

8. T. Wang et al., "Characterization of ovarian tissue based on quantitative analysis of photoacoustic microscopy images," *Biomed. Opt. Express* **4**(12), 2763–2768 (2013).

9. H. Li et al., "Utilizing spatial and spectral features of photoacoustic imaging for ovarian cancer detection and diagnosis," *J. Biomed. Opt.* **20**(1), 016002 (2015).

10. T. Wang et al., "A low-cost photoacoustic microscopy system with a laser diode excitation," *Biomed. Opt. Express* **5**(9), 3053–3058 (2014).

11. S. Nandy et al., "Evaluation of ovarian cancer: initial application of coregistered photoacoustic tomography and US," *Radiology* **289**(3), 740–747 (2018).

12. G. Yang et al., "Optimized light delivery probe using ball lenses for coregistered photoacoustic and ultrasound endo-cavity subsurface imaging," *Photoacoustics* **13**, 66–75 (2019).

13. G. Yang et al., "Co-registered photoacoustic and ultrasound real-time imaging of colorectal cancer: ex-vivo studies," *Proc. SPIE* **10878**, 108784K, (2019).

14. R. L. Siegel et al., "Colorectal cancer statistics, 2017," *CA Cancer J. Clin.* **67**(3), 177–193 (2017).

15. American Cancer Society, *Cancer Facts & Figures*, American Cancer Society, Atlanta, Georgia (2014–2016).

16. J. Pan et al., "Colonoscopy reduces colorectal cancer incidence and mortality in patients with non-malignant findings: a meta-analysis," *Am. J. Gastroenterol.* **111**(3), 355–365 (2016).

17. Y. Zeng et al., "The angular spectrum of the scattering coefficient map reveals subsurface colorectal Cancer," *Sci. Rep.* **9**(1), 2998 (2019).

18. T. Wang et al., "Label-free biomolecular imaging using scanning spectral interferometry," *Chin. Opt. Lett.* **11**(11), 111102–111105 (2013).

19. S. P. Raman, Y. Chen, and E. K. Fishman, "Evolution of imaging in rectal cancer: multimodality imaging with MDCT, MRI, and PET," *J. Gastrointest. Oncol.* **6**(2), 172 (2015).

20. E. Al-Sukhni et al., "Diagnostic accuracy of MRI for assessment of T category, lymph node metastases, and circumferential resection margin involvement in patients with rectal cancer: a systematic review and meta-analysis," *Ann. Surg. Oncol.* **19**(7), 2212–2223 (2012).

21. S. P. Hiotis et al., "Assessing the predictive value of clinical complete response to neoadjuvant therapy for rectal cancer: an analysis of 488 patients," *J. Am. Coll. Surg.* **194**(2), 131–135 (2002).

22. A. Mostafa et al., "Dual-mode photoacoustic and ultrasound system for real-time in-vivo ovarian cancer imaging," *Proc. SPIE* **10494**, 104944K, (2018).

23. H. S. Salehi et al., "Design of miniaturized illumination for transvaginal co-registered photoacoustic and ultrasound imaging," *Biomed. Opt. Express* **5**(9), 3074–3079 (2014).

24. E. Amidi et al., "Classification of human ovarian cancer using functional, spectral, and imaging features obtained from in vivo photoacoustic imaging," *Biomed. Opt. Express* **10**(5), 2303–2317 (2019).

25. E. Amidi et al., "Low-cost ultrasound and optical gelatin-based phantoms," *Proc. SPIE* **10878**, 108784A, (2019).

26. R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man Cybern.* **SMC-3**(6), 610–621 (1973).

27. G. Xu et al., "Photoacoustic spectrum analysis for microstructure characterization in biological tissue: a feasibility study," *Appl. Phys. Lett.* **101**(22), 221102 (2012).

28. R. Kumon, C. X. Deng, and X. Wang, "Frequency-domain analysis of photoacoustic imaging data from prostate adenocarcinoma tumors in a murine model," *Ultrasound Med. Biol.* **37**(5), 834–839 (2011).

29. F. L. Lizzi et al., "Relationship of ultrasonic spectral parameters to features of tissue microstructure," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **34**(3), 319–329 (1987).

30. E. Feleppa et al., "Ultrasonic spectral-parameter imaging of the prostate," *Imaging Syst. Technol.* **8**(1), 1098 (1997).

31. H. Tadayyon et al., "Quantitative ultrasound characterization of locally advanced breast cancer by estimation of its scatterer properties," *Med. Phys.* **41**(1), 012903 (2014).

32. J. Concato, A. R. Feinstein, and T. R. Holford, "The risk of determining risk with multivariable models," *Ann. Intern. Med.* **118**(3), 201–210 (1993).

Guang Yang received his BS degree in electrical engineering from Beijing University of Posts and Telecommunications, China, and his MS degree in biomedical engineering from Tsinghua University, China. Currently, he is pursuing his PhD in biomedical engineering at Washington University in St. Louis, USA. His research interest includes design and development of novel photoacoustic imaging systems for early-stage cancer detection.

Eghbal Amidi received his BS degree in electrical engineering from Urmia University, Iran, and his MS degree in the same major from Iran University of Science and Technology, Iran. Currently, he is pursuing his PhD in biomedical engineering at Washington University in St. Louis, USA. His research interest includes photoacoustic and ultrasound imaging.

William C. Chapman Jr. is a physician at Washington University School of Medicine in St. Louis. He received his Bachelor of Arts in military history from the University of North Carolina and his Doctor of Medicine from Georgetown University. His primary clinical and research focus is the surgical management of cancers of the colon and rectum.

Sreyankar Nandy received his BS degree in physics from University of Calcutta in 2007 and MS degree in applied physics from Indian Institute of Engineering Science and Technology (IIEST), Shibpur, in 2010. He subsequently obtained his MTech degree in applied optics from Indian Institute of Technology (IIT), Delhi, in 2013, and his PhD in biomedical engineering from Washington University in St. Louis in 2018, where his research was focused on developing novel optical and photoacoustic imaging modalities such as optical coherence tomography, photoacoustic tomography, and spatial frequency domain imaging for diagnosis and detection of ovarian and colorectal cancers. Since 2018, he has been a research fellow at Harvard Medical School and Massachusetts General Hospital, where his current research interest is development and translation of optical endomicroscopy systems for early *in vivo* detection of pulmonary diseases such as lung cancer and pulmonary fibrosis.

Atahar Mostafa received his BS degree in electrical and electronics engineering from the Bangladesh University of Engineering and Technology in Bangladesh and his MS degree in electrical engineering from the University of Saskatchewan in Canada. He received his PhD in biomedical engineering from Washington University in St. Louis in 2019, where his research was focused on developing novel diffused light and photoacoustic imaging modalities for diagnosis and detection of breast, ovarian, and colorectal cancers. Since January 2019, he is a senior engineer at Philips Boston, Massachusetts.

Heba Abdelal is a resident physician of anatomic and clinical pathology in Washington University in St. Louis and will complete her training in 2021. She received her MD degree from Ain Shams University in 2013. Her research interest is gastrointestinal/pancreas pathologies.

Zahra Alipour earned her MD degree from Azad University, Tehran Medical School in Iran. She is a resident physician of anatomic and clinical pathology in Washington University in St. Louis and will complete her training in 2021. Her research interest is gastrointestinal/ pancreas pathologies.

Deyali Chatterjee is an assistant professor of pathology and immunology at Washington University School of Medicine in St. Louis. She has a strong interest in clinicopathologic and collaborative research in the field of pancreatobiliary and gastrointestinal pathology. She studied various histopathologic aspects after neoadjuvant therapy in pancreatic ductal adenocarcinoma and has also studied pancreatic histology of mouse models, including xenografts for pancreatic cancer.

Matthew Mutch chief of the Section of Colon and Rectal Surgery at Washington University School of Medicine in St. Louis, has been named the Solon and Bettie Gershman Chair in Colon and Rectal Surgery at Barnes-Jewish Hospital (BJH). He received his MD from Washington University in St. His research and clinical interests include Colorectal cancer, colon cancer, cancer of the rectum, benign colorectal disorders, inflammatory bowel disease, etc.

Quing Zhu is a professor of the Department of Biomedical Engineering, Electrical and System Engineering and Radiology at Washington University in St. Louis. She is a pioneer of combining ultrasound and near infrared imaging modalities for diagnosis and treatment assessment of breast cancers. By collaborating with physicians/researchers, she and her team are the first group to explore coregistered photoacoustic and ultrasound techniques for ovarian cancer diagnosis, and for colorectal cancer diagnosis and treatment response assessment. She has been named fellow of the Optical Society of America, fellow of SPIE.