

Article

Insights into Biochemical Sources and Diffuse Reflectance Spectral Features for Colorectal Cancer Detection and Localization

Marcelo Saito Nogueira ^{1,2,*} , Siddra Maryam ^{1,2}, Michael Amissah ^{1,2}, Andrew McGuire ³, Chloe Spillane ³, Shane Killeen ³, Stefan Andersson-Engels ^{1,2}  and Micheal O'Riordain ³

¹ Tyndall National Institute, Lee Maltings, Dyke Parade, T12 R5CP Cork, Ireland

² Department of Physics, University College Cork, College Road, T12 K8AF Cork, Ireland

³ Department of Surgery, Mercy University Hospital, T12 WE28 Cork, Ireland

* Correspondence: marcelo.nogueira@tyndall.ie or marcelosaitonogueira@gmail.com

Simple Summary: Colorectal cancer (CRC) is the third most common and second most deadly type of cancer worldwide. The early detection and accurate characterization of colorectal cancer are associated with improved outcomes. With increasing emphasis on early cancer detection and the use of minimally invasive microsurgical techniques, the development of accurate diagnostic technologies to identify tumors and define their boundaries in real time becomes of paramount importance. The current research identifies potential cancer biomarkers and associated light-based instrument specifications to manufacture next-generation medical devices for CRC detection. These specifications have been chosen so that miniaturized instruments can be integrated into colonoscopes. Cancer biomarkers are listed to enable the use of complementary biochemical methods to analyze biological tissues. The impact of using next-generation colonoscopes in reducing cancer deaths can be assessed once medical devices are manufactured.

Abstract: Colorectal cancer (CRC) is the third most common and second most deadly type of cancer worldwide. Early detection not only reduces mortality but also improves patient prognosis by allowing the use of minimally invasive techniques to remove cancer while avoiding major surgery. Expanding the use of microsurgical techniques requires accurate diagnosis and delineation of the tumor margins in order to allow complete excision of cancer. We have used diffuse reflectance spectroscopy (DRS) to identify the main optical CRC biomarkers and to optimize parameters for the integration of such technologies into medical devices. A total number of 2889 diffuse reflectance spectra were collected in *ex vivo* specimens from 47 patients. Short source-detector distance (SDD) and long-SDD fiber-optic probes were employed to measure tissue layers from 0.5 to 1 mm and from 0.5 to 1.9 mm deep, respectively. The most important biomolecules contributing to differentiating DRS between tissue types were oxy- and deoxy-hemoglobin (Hb and HbO₂), followed by water and lipid. Accurate tissue classification and potential DRS device miniaturization using Hb, HbO₂, lipid and water data were achieved particularly well within the wavelength ranges 350–590 nm and 600–1230 nm for the short-SDD probe, and 380–400 nm, 420–610 nm, and 650–950 nm for the long-SDD probe.

Keywords: diffuse reflectance spectroscopy; colorectal cancer; optical spectroscopy; optical diagnostics; colonoscopy; surgical guidance; metabolomics; biomedical optics; biophotonics; multivariate analysis; chemometrics



Citation: Saito Nogueira, M.; Maryam, S.; Amissah, M.; McGuire, A.; Spillane, C.; Killeen, S.; Andersson-Engels, S.; O'Riordain, M. Insights into Biochemical Sources and Diffuse Reflectance Spectral Features for Colorectal Cancer Detection and Localization. *Cancers* **2022**, *14*, 5715. <https://doi.org/10.3390/cancers14225715>

Academic Editor: Fiona Lyng

Received: 9 October 2022

Accepted: 9 November 2022

Published: 21 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Colorectal cancer (CRC) is the third most common type of cancer worldwide, representing 11.3% (1.85 million) of diagnosed cancer cases and resulting in 10.2% (0.88 million) of cancer-related deaths in 2020 [1,2]. The large mortality caused by CRC can be attributed

to late-stage detection leading to poor patient prognosis. This, along with the significant morbidity and cost associated with standard surgery and associated adjuvant treatment modalities has led to the widespread adoption of screening methods to detect cancer at an early stage [3]. This is paralleled by a proliferation of minimally invasive methods such as Endoscopic Mucosal Resection (EMR), Endoscopic Submucosal Dissection (ESD), Transanal endoscopic microsurgery (TEM) and Transanal Minimally Invasive Surgery (TAMIS) to deal with these early lesions [4]. Currently, standard CRC screening tests such as the detection of blood in the stool have high false positive rates and accurate CRC detection is achieved by colonoscopy followed by relevant biopsies. However, screening and early diagnostic are limited by suboptimal compliance and lack of accessibility. With this in mind, global initiatives to develop early CRC detection methods have focused on CRC biomarkers and monitoring biochemical changes in tissues and biofluids. The high specificity of such biomarker detection methods is associated with their ability to detect low concentrations of specific molecules [5]. In tissue, structural changes are related to molecular concentrations in the millimolar range, whereas metabolic, immunologic and genetic features are associated with micromolar, nanomolar, and picomolar ranges, respectively [5]. Metabolic markers (e.g., enzymes, pO_2 , pH, and minerals) and immunologic markers (e.g., growth factors, hormones and cytokines) can be directly expressed in tissue at sufficient levels for real-time and non-invasive detection by optical techniques. Therefore, *in vivo* optical detection of tissue metabolic biomarkers is attractive to identify and localize precisely the tumor area, for example at colonoscopy, with no requirements for sample preparation for analysis of tissue sections, biofluids, and other samples used for cancer screening purposes. However, optical techniques can only be integrated into medical devices (e.g., colonoscopes) if technology allows for miniaturization and cost-effectiveness. One of the main cost-effective and miniaturizable techniques capable of probing tissue endogenous biomolecules is diffuse reflectance spectroscopy (DRS), which includes elastic scattering spectroscopy and near-infrared (NIR) spectroscopy for point measurements and can be translated to imaging by using hyperspectral imaging techniques.

DRS is an optical technique capable of tissue identification based on its biochemical composition and microstructure [6–10]. DRS works by sending light to the interrogated biological tissue and detecting the diffusively reflected light (i.e., the light that emerges from the tissue surface after being scattered inside it). Since the detected light traveled inside the tissue, it contains information about the tissue's optical properties such as scattering and absorption [11–18]. Light scattering is dependent on the refractive index mismatches of tissue including sets of molecules (e.g., collagen fibers and fibrils), organelles (e.g., mitochondria), cell membranes, and inhomogeneity of the intracellular and extracellular environment [10,19–21]. Therefore, scattering is associated with the tissue microstructure. On the other hand, absorption is associated with the tissue biomolecular composition, as its absorption bands are dependent on the molecular energy levels (including electronic [10,22–45], vibrational [46–56] and rotational levels [57]). Tissue absorption is related to biological chromophores (or absorbing biomolecules), which were previously investigated in a number of clinical and preclinical studies [58]. These chromophores include β -carotene, bile, bilirubin, ceroid, collagen, deoxyhemoglobin (Hb), oxyhemoglobin (HbO_2), methemoglobin (MetHb), water, lipid, and melanin [58]. The chromophores probed with DRS are typically present in biomolecule concentrations from micromolar to nanomolar [5]. These concentrations describe metabolic and immunologic features of contrast between normal and cancer tissues [5].

Previous DRS studies have investigated the average concentration of specific tissue biomolecules (i.e., chromophores) based on spectral fitting models of diffuse reflectance, diffuse transmittance and fluorescence [24,26,32,58–67]. These concentrations were analyzed in clinical breast cancer studies about tumor vascularity [68], blood oxygenation [69], quantitative chemical information of oxy- and deoxyhemoglobin (HbO_2 and Hb, respectively), water, and lipids [70–72]. Bile, blood, water and lipid concentrations were also reported for clinical studies on liver cancer [61,73]. In addition, previous NIR spectroscopy studies

on brain, muscle, mammary, lung and prostate cancers of rats and mice reported altered vasculature, oxygen dynamics and HbO₂, Hb, water, lipid, protein and DNA content in tumor tissues [74–82]. In terms of CRC detection, DRS studies reported that hemoglobin concentration and blood oxygenation can be used to differentiate between normal and tumor tissues [83–87]. Indices of statistical differences in biomolecule concentrations have been reported for application in colorectal cancer surgery [88]. In addition, DRS was used together with machine learning methods for direct tissue classification. These studies include those distinguishing tumor tissue from healthy surrounding tissues in the oral cavity (head and neck cancer) [89], breast [90,91], lung [92], and liver [93,94]. Recent studies have investigated the detection of colorectal cancer during surgery [88,95–101]. In terms of primary colorectal cancer, DRS and related optical spectroscopic techniques have mainly been used to guide colonoscopy by distinguishing the normal mucosa and malignant tissue inside the colon (luminal side). However, DRS studies analyzing multiple chromophores in an extended NIR wavelength range into the short wavelength infrared (SWIR) region are rare. To the best of our knowledge, this is the first CRC study optimizing wavelength ranges of reflectance spectra based on the combination between statistical tests and machine learning methods. In addition, the extended wavelength range between 350 and 1920 nm has been probed and our analysis objectively determined the biomolecules most important for tissue classification based on their depth-dependent absorption and scattering. As a final novel point of our study, our objective analysis does not need any homogeneous-medium assumptions such as those required by reflectance spectral fitting models to work.

In this study, we analyzed the spectral regions of best diagnostic potential and the biochemical sources of the classification between normal mucosa and tumor tissues based on the DRS spectral machine-learning features associated with chromophore absorption. The importance of these sources for such classification was determined by (1) assessing the amplitude of partial least-squares components (PLSCs) at relevant wavelength ranges of tissue scattering and absorption, (2) comparing the shape of the PLSC loading curves and chromophore absorption spectra, and (3) evaluating the classification performance (sensitivity, specificity, accuracy and area under the receiver operating characteristic curve) for wavelength ranges of statistically significant difference at a significance level of $p = 0.001$ of t -test after ensuring normality by using Anderson-Darling and Lilliefors tests. Our study evaluated the tissue classification in both superficial and deeper layers by using probes with short and long source-to-detector distances (SDD) of 630 μm and 2500 μm , respectively. Biomolecules were probed in a wide wavelength range from 350 nm to 1920 nm, which allowed for the collection of tissue structural and biomolecular signals at different depths. Our DRS spectral dataset contained 2889 spectra of freshly excised *ex vivo* tissues of 47 patients, which was used for robust and reliable analysis of the tissue biochemistry.

2. Materials and Methods

2.1. Diffuse Reflectance Spectroscopy (DRS) Instrumentation

The DRS equipment used in this study, illustrated in Figure 1, consisted of a broadband light source (HL-2000-HP, Ocean Optics, Edinburgh, United Kingdom) with emission ranging from 350 nm to 2400 nm, a quadrifurcated fiber optic probe with source-to-detector distance (SDD) of 630 μm (BF46LS01 1-to-4 Fan-Out Bundle, Thorlabs, Munich, Germany), a trifurcated fiber optic probe with SDD of 2500 μm (Fibertech Optica, Anjou, Canada), a visible/near-infrared (NIR) wavelength spectrometer (QE-Pro, Ocean Optics, Edinburgh, UK) and a NIR/SWIR spectrometer (NIR-Quest, Ocean Optics, Edinburgh, UK). The fiber optic probes were made of low-OH silica in order to allow better transmission at the SWIR range. These probes were used for both illumination and collection of the reflected light to be detected by the spectrometers. The visible/NIR spectrometer collected light in the wavelength range between 350 nm and 1140 nm, while the NIR/SWIR spectrometer detects light from 1090 nm to 1920 nm. The overlapping region was used to merge the spectra into one broadband spectrum from 350 nm to 1920 nm. Once reflected light was detected by the

spectrometers, the intensity readings were preprocessed in order to obtain the tissue DRS spectra according to Section 2.5.

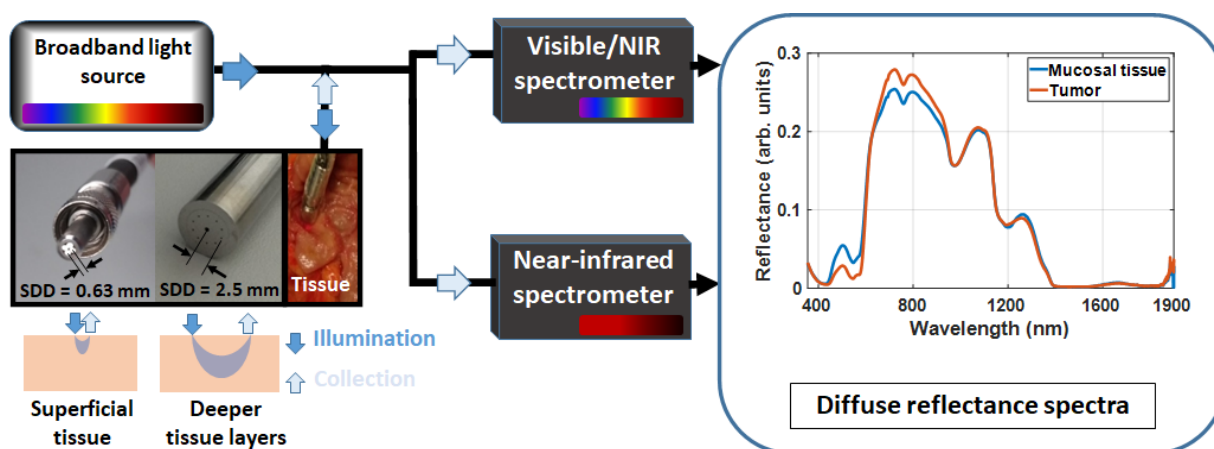


Figure 1. Schematic drawing of our DRS system. The obtained broadband reflectance contains information about a larger variety of tissue biomolecules compared to studies probing shorter wavelength ranges [61,86,87,102–111]. Broadband reflectance spectra were obtained by merging the visible and NIR spectra based on the overlapping spectral region between the two spectrometers (from 1090 nm to 1140 nm). The spectral merging procedure is described in detail in [11,12] and in Section 2.5. Briefly, in the overlapping spectral region, the two spectra are added with a smoothing weighting.

2.2. Probing Superficial or Deeper Tissue Layers

By using the 630 μm SDD probe (630 μm fiber center-to-center distance, Figure 1), our DRS system collected reflectance signals from 0.5 to 1 mm deep into tissue (between 450 and 1590 nm). In this case, the probe contained 600 μm core diameter fibers for both illumination and collection and will be referred to as a short-SDD probe throughout this article. In order to collect light from deeper tissue layers (between 0.5 and 1.9 mm deep between 450 and 1590 nm), we used a 2500 μm SDD probe (long-SDD probe) containing one source fiber in the center and 10 collection fibers surrounding it. Each 5 collection fibers were positioned linearly in the proximal end of the fiber to match the slit of the visible/NIR and NIR/SWIR spectrometers to optimize the light detection configuration. The source and collection fibers of the long-SDD probe had 600 μm and 200 μm core diameters, respectively.

In order to estimate the chromophores and depth interrogated by each probe, we used a spectral fitting algorithm to extract the optical properties from our DRS measurements. The fitting was based on a look-up table of Monte Carlo simulations of the light propagation into tissues and iteratively modifying chromophore concentrations and scattering properties. As a probe depth estimate for each wavelength, we used the depth of the maximum fluence value at the mean position between the source and detector. Then, the minimum and maximum probed depth were reported in this study with the purpose of illustrating the independence of the datasets acquired with each probe and its impact on the evaluation of the biochemical composition of superficial or deeper tissue layers.

2.3. Optical Data Collection

Our data collection started with the background and reference measurements. Reference measurements were taken by positioning each probe on a specialized holder able to keep a fixed distance between the fiber optic probes and our reflectance standard (FWS-99-01c, Avian Technologies LLC, New London, CT, USA). Since the holder was closed to avoid interference from ambient light, it was used to take both reference and background measurements. Probe contamination was avoided by covering our probes with transparent polyvinyl chloride (PVC) film during measurements. After each set of measurements, the

plastic film was removed and the probes were cleaned with ethanol 70%. The same probes were used for every clinical measurement throughout this study. Every measurement was performed by positioning the probe as close to a perpendicular angle to the tissue surface as possible. Measurements of both probes were performed at similar locations within millimeters of each other. Briefly, we positioned the sample on a board with a coordinate system, which allowed us to come back to a similar position for both probes. the experimental procedure is described in detail in [11,12]. After completing the data collection, the data were safely stored for subsequent analysis. In this study, we collected a total of 1363 spectra for the short SDD probe (630 μm SDD) and 1526 for the long SDD probe (2500 μm SDD). Each spectrum is an average of a triplicate measurement in the same tissue location.

2.4. Clinical Protocol and Research Ethics

Our study included 47 patients undergoing bowel resection at the Mercy University Hospital (Cork, Ireland). Patient demographics and tumor characteristics are shown in Table 1. The study was approved by the Clinical Research Ethics Committee of the University College Cork. Our procedure consisted of collecting around 15 measurements of *ex vivo* mucosal tissues and 15 tumor tissues on the specimen after surgical resection. Measurements were taken from a typical area of 100 cm^2 . After a specimen was resected, the colonic lumen was exposed. The specimen was rinsed with water and cleaned afterward in order to remove the excess blood and any remaining feces from the mucosal surface. Then, the mucosa and tumor regions were identified by experienced surgeons. The time between the specimen removal and the start of the data collection was on average 40 min. All data collection was performed within an average time of 1 h after surgical resection. Physiological conditions were kept as much as possible throughout the data collection by keeping the tissue moist with a wet wipe. In order to correlate spectral readings with the tissues measured, the coordinate of every reading was registered by using a picture of the specimen over a grid. The boundary of each tissue type was determined by experienced surgeons. After the acquisition of all-optical DRS data, the specimen was returned to the Pathology Department for processing and analysis according to standard protocols. The ground truth of cancer tissue types was obtained by histopathology analysis.

Table 1. Patient demographics, cancer types and tumor staging classification.

Patient and Cancer Characteristics		Number of Patients/Tumors
Total		47
Gender	Male	32
	Female	15
Age (years)	Median	69
	Minimum	40
	Maximum	89
	Interquartile range	13.5
Cancer types	Adenocarcinoma	47
T (tumor) stage	pT1	5
	pT2	7
	pT3	26
	pT4	9
N (lymph node) stage	N0	19
	N1a	9
	N1b	12
	N1c	1
	N2	1
	N2a	4
	N2b	1

2.5. Data Preprocessing and Feature Selection

First, both visible and NIR tissue intensity spectra had their background subtracted and the resulting signal was divided by the reflected intensity of the reference (reflectance standard) according to the expression:

$$\text{Reflectance}(\lambda) = \frac{1}{\text{Reference reflectivity}} \times \frac{\text{Tissue reflected intensity} - \text{Background intensity}}{\text{Reference reflected intensity} - \text{Background intensity}} \quad (1)$$

Next, the broadband reflectance spectra were obtained by merging the visible and NIR spectra based on the overlapping spectral region between the two spectrometers (from 1090 nm to 1140 nm). The merging was performed by interpolating the overlapping region of the two spectrometers and performing the following weighted sum:

$$\text{Reflectance}(\lambda) = \sum_{i=0}^{100} \frac{((100-i) \times \text{reflectance of VIS spectrometer} + i \times \text{reflectance of NIR spectrometer})}{100} \quad (2)$$

The result is a smooth reflectance curve where the reflectance measured by each spectrometer has a higher contribution at their respective wavelength regions of higher sensitivity. As preparation for classification tests based on k-nearest neighbors (which may be prone to overfitting when using numerous variables in comparison to the sample size), feature extraction was performed by using partial-least squares (PLS).

PLS is a supervised method of orthogonal transformation used for linear dimension reduction in a given dataset. PLS is used to create a new set of linearly independent variables (partial least squares components or PLSCs) which maximize the covariance between the predictors (reflectance values for each wavelength) and responses (tissue types) [112]. As PLSCs are weighted sums of the original variables (or predictors), the combination of weights of the first PLSCs shows the wavelength regions which are responsible for most of the discrimination between two classes (tissue types). More details about the calculation of PLSCs and the importance of predictors for better tissue discrimination can be accessed from the publications of Brereton et al. [113] and Gromski et al. [114], respectively. We also emphasize that our PLSCs are the same as the principal components (PCs) of PLS.

In this study, the bias due to the incomparable scales of observations (reflectance values) at specific variables (reflectance at particular wavelengths) was avoided by scaling/normalizing observations between -1 and $+1$ for each wavelength. By compensating for the difference in scale on reflectance values, this scaling ensures the feature extraction equally takes into account the contributions of each wavelength on the new set of variables based on the PLS maximization of the discrimination between the two tissue types. This contribution is translated into the weights (loadings) of each wavelength on the variables selected to develop a tissue classification model. Finally, data preprocessing and analysis were performed using home-made MATLAB routines (MathWorks Inc., Natick, MA, USA).

2.6. Extraction of Spectral Features

The spectral features for differentiation between normal mucosa and cancer tissues were extracted by using the first four PLSCs, which were selected to avoid overfitting by stopping to include new PLSCs when accuracy increments were lower than approximately 2%. These features were interpreted based on the amplitude and spectral shape of PLSC loadings. The amplitude was used to determine the contribution of specific wavelength ranges to tissue classification, which were associated with the main absorbing ranges of typical tissue biomolecules (or chromophores). The spectral shape was associated with characteristics of the chromophore absorption spectra. More details of the interpretation of the spectral features are described in Section 4.2.

In order to provide information about which wavelength ranges were related to most of the tissue biomarkers associated with cancer detection, delineation and potential carcinogenesis, we evaluated the tissue classification performance parameters (sensitivity,

specificity, accuracy and area under the receiver operating characteristic curve; AUC) achieved by using wavelength ranges selected via statistically significant difference verified by a student *t*-test. First, normal distributions of DRS readings at each wavelength and tissue type were verified through Anderson-Darling and Lilliefors normality tests. Next, a student *t*-test was applied to the same distributions. Wavelength ranges leading to $p < 0.001$ were selected for building our tissue classification model. This model was built by applying PLS to the data on the relevant wavelength ranges and selecting the four highest-order PLSCs to be used with a weighted KNN classification algorithm. The weighted KNN algorithm used 10 neighbors and squared inverse distance between observations.

Based on the results of the KNN classification, its performance was evaluated by using two-fold cross-validation for 20 iterations. Each iteration consists of randomly dividing the dataset into training and test sets of equal size. The classification model was generated by using the training set and, then, applied for tissue classification on the test set. The process is repeated with the first test set used as the training set and vice versa. At the end of this process, the output was the mean of each classification performance parameter. The process is repeated 20 times. Then, the mean and standard deviation of the output of the 20 iterations were calculated. The reproducibility of these parameters was evaluated by the obtained standard deviations.

A flowchart summarizing the spectral analysis is shown in Figure 2.

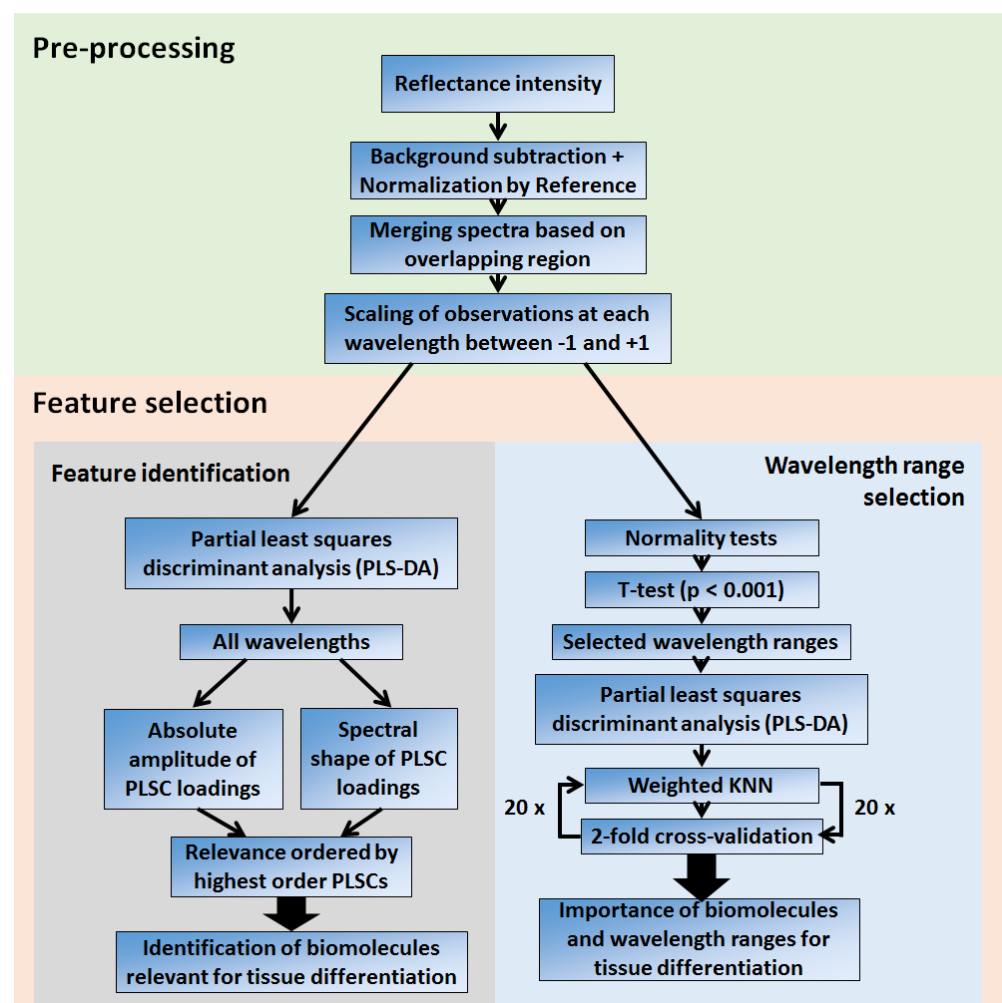


Figure 2. Flowchart of the steps of our spectral analysis.

3. Results

3.1. Tissue Classification Features Based on PLSC Amplitudes

Our first analysis for the selection of the features relevant to tissue classification was based on the PLSC loadings. The interpretation of the PLSC loadings is described in Section 4.2. Figure 3 illustrates that the PLSC1 loadings of the short-SDD probe exhibit the highest and second-highest absolute amplitudes in the range from 600 nm to 1350 nm and from 350 nm to 600 nm, respectively. A similar behavior is observed on the amplitudes of the PLSC1 of the long-SDD probe, whose first, second and third highest absolute amplitudes occur between 450 nm and 600 nm, between 650 nm and 1350 nm, and between 1350 nm and 1900 nm, respectively. These amplitudes suggest that most of the differentiation between mucosal and cancerous tissues may originate from the absorption and scattering processes below 1350 nm. Processes at the visible wavelength range 350–600 nm are predominantly related to blood absorption, and information regarding the near-infrared range 700–1350 nm could be associated with relatively lower absorption of oxyhemoglobin (HbO₂) and deoxyhemoglobin (Hb), lipid and water at slightly deeper tissue layers.

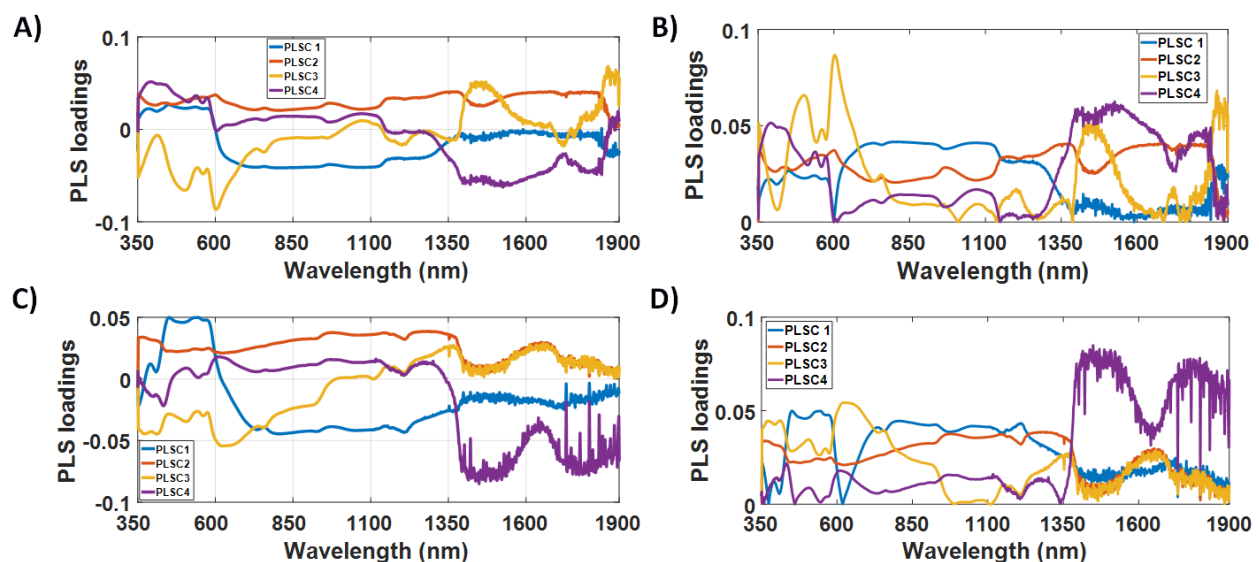


Figure 3. PLS components (PLSCs) for the classification between cancerous and mucosal/submucosal tissues. (A) Raw and (B) absolute values of PLSC loadings for the short-SDD probe and (C) Raw and (D) absolute values of PLSC loadings for the long-SDD probe.

On the PLSC2, PLSC3 and PLSC4 of the short-SDD probe, overall higher absolute amplitudes can be found for wavelengths below 700 nm and above 1350 nm, whereas the same components for the long-SDD probe have amplitude loadings more uniformly distributed over the full spectral window between 350 nm and 1900 nm. Particularly for PLSC4 of the long-SDD probe, the amplitude is higher for wavelengths above 1350 nm. While the high amplitudes below 700 nm for the short-SDD probe reinforce potentially relevant features of tumor detection at the UV-visible region (dominated by blood absorption), tissue classification using other wavelengths is unclear from only absolute PLS-loading amplitudes. With this in mind, we analyzed the shape of the PLSCs based on the wavelength regions and spectral shape of the chromophore absorption spectra, as shown in Section 3.3. On the other hand, this analysis is subjective and does not allow us to define specific wavelength ranges, as these would depend on subjective thresholds of the amplitude of PLSC loadings and the choice of how many PLSCs to be considered for thresholding. In addition, although PLSC1 is more important for tissue classification, the weight of each PLSC on the importance of wavelength ranges is unclear.

3.2. Wavelength Selection and Tissue Classification

Since analyzing the PLSC loadings may lead to subjective observations, we performed a more objective analysis by selecting wavelengths through statistically significant differences in the *t*-test ($p < 0.001$). In order to ensure that *t*-test could be applied to our dataset, Anderson-Darling and Lilliefors normality tests were applied to the data of each individual wavelength. Normal distributions were identified for all the wavelengths of both probes, except for those between 536 nm and 545 nm for the short-SDD probe. However, we confirmed statistically significant difference ($p < 0.001$) is obtained at those wavelengths upon application of the Wilcoxon rank-sum test (data not shown) and obtained no difference in tissue classification by including or excluding this wavelength range (Table 2).

Table 2. Tissue classification performance * of PLS-KNN for the short-SDD probe using wavelengths selected by *t*-test ($p < 0.001$). Blue fields represent the performance using visible/NIR light detection available in Si-detector-based spectrometers, orange fields cover the performance using NIR wavelengths detected by InGaAs-based spectrometers and green fields show the performance of both types of wavelength range combined. Means and standard deviations were taken from the outcomes of 20 iterations of 2-fold cross-validation.

Wavelengths	Sensitivity	Specificity	Accuracy	AUC
350–540 nm, 540–590 nm	(78.2 ± 0.9)%	(75.8 ± 1.6)%	(77.1 ± 1.0)%	(0.854 ± 0.007)
350–590 nm	(78.4 ± 1.1)%	(74.9 ± 1.4)%	(76.7 ± 0.8)%	(0.85 ± 0.005)
600–1230 nm	(79.8 ± 0.9)%	(84.4 ± 1.4)%	(81.9 ± 0.8)%	(0.894 ± 0.005)
350–590 nm, 600–1230 nm	(78.6 ± 0.7)%	(75.4 ± 1.0)%	(77.1 ± 0.7)%	(0.854 ± 0.006)
1530–1700 nm	(70.9 ± 1.1)%	(67.0 ± 1.6)%	(69.1 ± 1.0)%	(0.771 ± 0.009)
1730–1850 nm	(69.1 ± 1.0)%	(69.5 ± 1.3)%	(69.3 ± 0.9)%	(0.765 ± 0.007)
1530–1700 nm, 1730–1850 nm	(76.4 ± 0.9)%	(77.7 ± 1.1)%	(77.0 ± 0.7)%	(0.845 ± 0.006)
350–590 nm, 600–1230 nm, 1530–1700 nm, 1730–1850 nm	(85.5 ± 0.8)%	(84.0 ± 1.0)%	(84.8 ± 0.7)%	(0.919 ± 0.004)
350–1920 nm	(85.6 ± 0.9)%	(80.4 ± 1.1)%	(83.2 ± 0.8)%	(0.905 ± 0.005)

* Results corresponding to 2-fold cross-validation from 20 iterations of random sampling of training and test sets. Our validation provides a sufficiently robust estimation of classification performance metrics since the standard deviation of measurements within each individual patient is less than twice the standard deviation of all measurements of all patients. In addition, no clear trend among specific patients has been identified in our dataset of 47 patients. A typical measurement means and standard deviation for one patient can be found in Figure S1 of the Supplementary Material.

After confirmation of normality for most of the wavelengths, we applied a *t*-test for each wavelength. We used the wavelength ranges where a statistically significant difference ($p < 0.001$) was obtained (Figure 4) to select bands relevant for tissue classification.

By using the wavelength ranges of Figure 4 as well as combinations of those ranges for UV-visible or NIR, tissue PLS-KNN classification models were built and compared with the model using all the wavelengths. The classification performance of each model can be found in Tables 2 and 3.

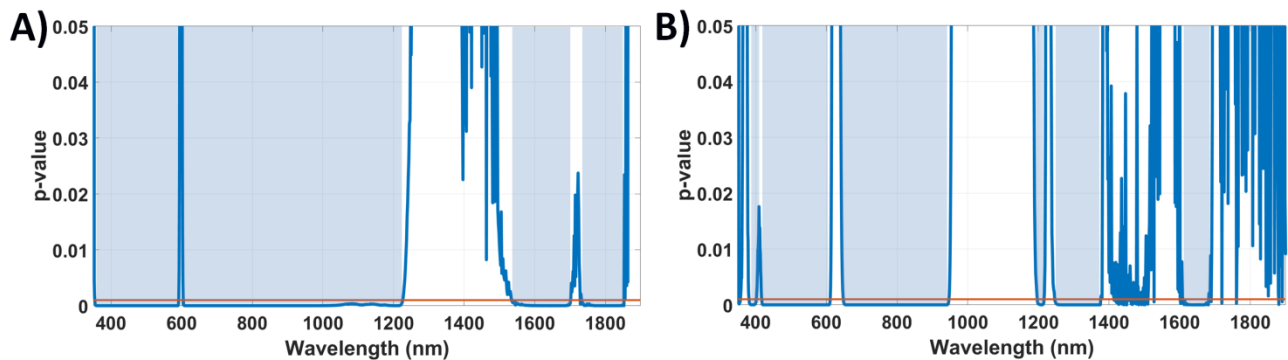


Figure 4. Selected spectral regions (blue) where statistically significant differences ($p < 0.001$) are found for (A) short-SDD probe and (B) long-SDD probe. The red line indicates the 0.001 cutoff for the p -value.

Table 3. Tissue classification performance * of PLS-KNN for the long-SDD probe using wavelengths selected by t -test ($p < 0.001$). Blue fields represent the performance using visible/NIR light detection available in Si-detector-based spectrometers, orange fields cover the performance using NIR wavelengths detected by InGaAs-based spectrometers and green fields show the performance of both types of wavelength range combined. Means and standard deviations were taken from the outcomes of 20 iterations of 2-fold cross-validation.

Wavelengths	Sensitivity	Specificity	Accuracy	AUC
380–400 nm	(86.0 ± 0.9)%	(85.0 ± 0.9)%	(85.6 ± 0.7)%	(0.925 ± 0.004)
420–610 nm	(85.6 ± 0.5)%	(87.2 ± 0.6)%	(86.3 ± 0.3)%	(0.93 ± 0.004)
650–950 nm	(89.6 ± 0.6)%	(89.7 ± 1.0)%	(89.7 ± 0.6)%	(0.96 ± 0.004)
380–400 nm, 420–610 nm, 650–950 nm	(87.0 ± 0.8)%	(85.5 ± 0.8)%	(86.3 ± 0.7)%	(0.931 ± 0.003)
1200–1220 nm	(67.8 ± 1.2)%	(63.3 ± 1.9)%	(65.7 ± 1.1)%	(0.707 ± 0.013)
1250–1380 nm	(77.1 ± 1.0)%	(80.7 ± 1.0)%	(78.8 ± 0.7)%	(0.87 ± 0.006)
1600–1690 nm	(62.4 ± 1.1)%	(58.3 ± 1.6)%	(60.4 ± 0.8)%	(0.654 ± 0.008)
1200–1220 nm, 1250–1380 nm, 1600–1690 nm	(77.6 ± 1.0)%	(84.7 ± 1.1)%	(81.0 ± 0.8)%	(0.883 ± 0.006)
380–400 nm, 420–610 nm, 650–950 nm, 1200–1220 nm, 1250–1380 nm, 1600–1690 nm	(89.1 ± 0.7)%	(90.2 ± 0.7)%	(89.6 ± 0.5)%	(0.957 ± 0.004)
350–1920 nm	(89.3 ± 0.6)%	(90.2 ± 0.7)%	(89.7 ± 0.5)%	(0.959 ± 0.003)

* Results corresponding to 2-fold cross-validation from 20 iterations of random sampling of training and test sets. Our validation provides a sufficiently robust estimation of classification performance metrics since the standard deviation of measurements within each individual patient is less than twice the standard deviation of all measurements of all patients. In addition, no clear trend among specific patients has been identified in our dataset of 47 patients. A typical measurement means and standard deviation for one patient can be found in Figure S1 of the Supplementary Material.

Table 2 indicates that the spectral regions 600–1230 nm and 350–590 nm are the first and second most important for tumor detection by using the short-SDD probe, as accuracy is higher for these wavelengths. When tissue classification is performed only with wavelengths of the first spectral region, the achieved sensitivity (79.8 ± 0.9)% and specificity (84.4 ± 1.4)% are comparable to that obtained by using all the selected wavelength regions combined. This result suggests that the difference between mucosa and tumors is generated from the combination of absorption of Hb, HbO₂, lipid and water as well as an optical

scattering of tissue layers slightly below the tissue surface. In particular, Hb and HbO₂ may play a significant role in classification across the superficial tissue layers, as the spectral regions leading to the highest discrimination cover all their absorption wavelengths.

In terms of the combination of wavelength ranges, the specificity achieved by combining the 350–590 nm and 600–1230 nm wavelength ranges was lower than using 600–1230 nm alone. This lower classification performance may be associated with higher absorption and scattering properties between 350 and 590 nm, which may lead to higher separation between the centers of the mucosa and tumor distributions recognized by PLS, while less contrast between the two distributions was achieved due to higher variation in reflectance values. On the other hand, higher classification performance over all parameters was obtained by combining the ranges 1530–1700 nm and 1730–1850 nm. Then, the information provided by Hb and HbO₂ in both wavelength ranges may be redundant, whereas signals associated with water and lipid absorption may be complementary. The information from all the selected wavelength ranges is also complementary, as their combination leads to higher classification performance compared to all ranges tested in this study, including a 3.6% higher specificity than that obtained by using all wavelengths (350–1920 nm).

Table 3 shows that wavelength ranges below 950 nm led to higher sensitivity than those above 1200 nm for the long-SDD probe. This indicates that Hb and HbO₂ are the chromophores that most contribute to accurate tumor detection in deeper tissue layers. Similarly to the short-SDD probe, the highest classification performance was achieved by using wavelengths on the optical window (in this case, 650–950 nm). Additionally, since the performance obtained for the long-SDD probe is higher than that obtained by using the short-SDD probe, signals of deep tissue layers may contain more relevant information for tissue discrimination, and probes could be designed in future studies to obtain information from relevant tissue depths. The importance of deeper tissue layer information is further reinforced by the higher classification performance achieved by using the ultraviolet (380–400 nm) and visible (420–610 nm) wavelengths alone compared to the best performance obtained by using the short-SDD probe.

Although the higher performance achieved by using the range 650–950 nm compared to shorter wavelengths may be attributed to the contribution of water and lipid absorption in the optical window, NIR wavelength ranges led to relatively low classification performance. In this case, the most informative NIR range was 1250–1380 nm, where features of higher variations in lipid and water absorption can be simultaneously observed (Figure 5).

Similar behavior was observed between the classification performance achieved by using the long-SDD probe and the short-SDD probe. By combining wavelength ranges containing ultraviolet and visible wavelengths, lower classification performance was obtained compared to the 650–950 nm range alone. In addition, the combination of NIR wavelength ranges improved tissue classification. Those results suggest that information from water and lipid absorption at deeper tissue layers probed by using long SDDs may be complementary and useful for cancer detection, while signals from Hb and HbO₂ absorption may not bring useful information for tissue classification. Finally, the combination of selected wavelength ranges led to the achievement of as good performance as using the entire wavelength range (350–1920 nm) investigated in this study.

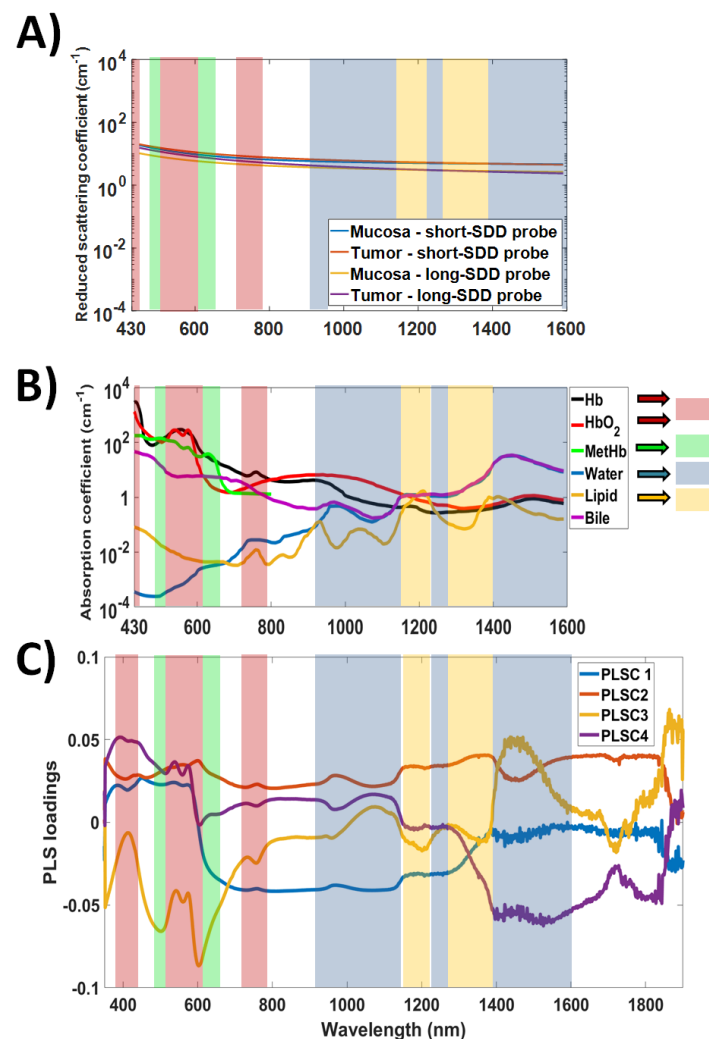


Figure 5. Wavelength regions with spectral features of mucosa and tumor scattering coefficients and each tissue chromophore shown at the (A) Reduced scattering spectra of mucosa and tumor tissues of both short-SDD and long-SDD probes, (B) chromophore absorption spectra and (C) PLSs of the short-SDD probe.

3.3. Relationship between Tissue Classification Features and Tissue Biochemistry/Microstructure

Once the importance of wavelength bands and their combinations were assessed, we used the shape of the PLSC loadings to understand the biochemical and microstructural sources of tissue classification. This section covers the results of our analysis, whereas the interpretation of spectral features of tissue classification is discussed in Section 4.2. Our analysis comprised of the determination of contributions of tissue scattering and chromophore absorption based on the shape of the PLSC loadings at wavelength bands of highest absorption of tissue chromophores as well as the “flatness” of PLSC loadings (indicative of the predominant contribution of tissue scattering). The selection of wavelength bands for subsequent interpretation is illustrated in Figure 5.

Figure 5 shows the spectral ranges where particular features of specific chromophore absorption can be observed. As an example, the red region shows the features of oxy-hemoglobin (HbO₂) and deoxyhemoglobin (Hb), including bands between 380 nm and 450 nm (peaks of HbO₂ at 414 nm, 542 nm and 576 nm [21], and peaks of Hb at 433 nm, 556 nm and 757 nm). In the case of the latter peak of Hb (at 757 nm), we assume the feature is from Hb instead of lipid (peak at 761 nm) due to the higher Hb absorption and potentially higher Hb concentration in biological tissues (resulting in overall higher tissue absorption due to Hb). In addition, the used bile spectrum (from Nachabe et al. [61]) includes the

water contained in the sample, as is obvious from the spectral range above 950 nm. When the shape of the PLSC loadings is relatively flat and monotonic as a function of wavelength, we attributed scattering as the most important factor contributing to tissue classification using such PLSC. It is important to remember that, even though the chromophore spectra are not shown for wavelengths longer than 1600 nm in Figure 5, double absorption peaks between 1700 nm and 1800 nm are associated with lipid, whereas the increase in absorption close to 1900 nm is related to water [115,116]. By analyzing the spectral shape of the PLSCs, we obtained Table 4 below.

Table 4. Scattering and absorption features of the superficial tissue PLSCs.

Scattering and Absorption Features								
		VIS Scat	NIR Scat	Hb	HbO ₂	MetHb	Water	Lipid
PLS (Short-SDD probe)	PLSC1			X	X		X	
	PLSC2	X	X	X	X		X	
	PLSC3			X	X		X	X
	PLSC4			X	X		X	X
PLS (Long-SDD probe)	PLSC1			X	X		X	X
	PLSC2	X	X	X	X		X	X
	PLSC3			X	X		X	X
	PLSC4			X	X		X	X

Based on the loadings of PLSC1, Table 4 suggests that the features contributing to tissue classification between mucosa and tumor are mostly related to Hb, HbO₂ and water for the short-SDD probe and the Hb, HbO₂, water and lipid for the long-SDD probe. For both probes, absorption features from the same chromophores of PLSC1 appear in PLSC2, which also contain features of scattering in visible and near-infrared wavelength ranges. In addition, absorption features of Hb, HbO₂, water and lipid can be observed on the loadings of PLSC3 and PLSC4 of both probes. On the other hand, features of lipid absorption appear only at the loadings of PLSC3 and PLSC4 of the short-SDD probe, whereas characteristics of the same chromophore are exhibited in the loadings of PLSC1, PLSC2, PLSC3, and PLSC4 of the long-SDD probe.

4. Discussion

4.1. Impact of Depth-Resolved Determination of Wavelength Ranges and Biomarkers for Tissue Classification

Our study investigated the most important wavelength regions for discriminating colorectal mucosa and cancer tissues and how these regions can be related to the biomarkers contributing to this discrimination (discussed in Section 4.2). Furthermore, we provide information for two tissue probed depths, as these depths are dependent on the source-to-detector distance of our fiber optic probes (Section 4.2). Our optimization of wavelength ranges of reflectance spectra and evaluation of main biomarkers contributing to the classification of normal mucosa and tumor tissues is an extension of our previous work [11] showing the successful classification of these tissues by using support vector machines, as well as our study [12] estimating biomolecule concentrations by using a spectral fitting algorithm based on Monte Carlo simulations of light propagation in tissues (assuming homogeneous tissue models).

To the best of our knowledge, previous DRS studies did not perform an objective analysis of the importance of wavelength ranges for discrimination between mucosa and cancer tissues. These studies have been reviewed in our previous publication [11] and have only investigated specific wavelength ranges from 300 to 800 nm or from 900 to 2500 nm. No comparison between wavelength ranges objectively selected by statistical

methods has been performed by using DRS for CRC detection. With that in mind, our study provides the objective analysis of the importance of ultraviolet, visible and near-infrared wavelength ranges for classification between colorectal mucosa and tumor tissues, which is especially useful (1) to design new optical spectroscopy and imaging systems restricted to specific wavelength ranges (for cost-effectiveness, performance maintenance upon miniaturization and integration into medical devices, higher accuracy at specific wavelengths, and higher spatial resolution), (2) to determine the range of probed tissue depths where improved tissue classification can be achieved and where optical tissue biomarkers tend to significantly influence tissue classification, (3) to identify which tissue biomarkers are the most important for discrimination between mucosa and cancer tissues based on DRS signals.

Our study improves the subjectivity of analysis of amplitude and shape of PLSC loadings (Figure 3) by adding an objective analysis based on the selection wavelength ranges for tissue classification based on $p < 0.001$ (statistically significant difference) for the t -test and comparison of PLS-KNN classification performance metrics by using each combination of selected wavelength ranges. Higher classification performance at wavelength ranges of specific tissue chromophores (biomolecules) indicates the most important wavelength ranges and respective biomarkers to classify normal mucosa and cancer (Section 3.2). This classification performance adds to biomarker identification based on features of tissue scattering and chromophore absorption based on the amplitude (Section 3.1) and shape (Section 3.3) of PLSC loadings since statistical wavelength selection and subsequent classification cannot identify scattering contributions spread out all wavelengths. The scattering contribution to tissue classification is only observed by analysis of PLSC loading shapes as a function of wavelength.

In this study, the analysis of PLSC loadings suggests that tissue scattering is secondary but still important for tissue classification since only PLSC2 loadings resemble scattering coefficient curves as a function of wavelength (Section 3.3). However, it is worth noting that different combinations of optical properties (scattering and absorption coefficients) lead to probing a different depth in tissue. Probing different depths means that each wavelength of each DRS spectrum extracts biomarkers (chromophore concentrations and scattering properties) at a different tissue depth. Additionally, the higher source-detector distance increases the chances of collecting light which traveled longer and deeper into tissue. Therefore, short-SDD and long-SDD probes capture information on biomarkers at different depths, as these depths depend on both probe geometry and tissue optical properties. Based on Monte Carlo simulations of light propagation in tissues performed in our previous study [11], the probed depth of the short-SDD probe was mostly within 0.5–1 mm for wavelengths between 450 and 1590 nm, whereas that of the long-SDD probe varied between 0.5 and 1.9 mm for the same wavelength range. Statistical methods presented in this paper enable objective depth-resolved biomarker identification and selection of wavelength ranges. This depth-resolved analysis is not achievable by spectral fitting models assuming that tissue is homogeneous because tissue heterogeneity is neglected to keep the number of fitted parameters to a minimum. At the cost of a tissue homogeneity assumption, such spectral fitting models can retrieve average tissue scattering properties and chromophore concentrations. Although average concentrations are easy to interpret, they contain neither information about tissue depths nor wavelength ranges to best differentiate normal and cancerous tissues.

Previous studies have calculated sensitivity and specificity for tumor detection based on biomolecular concentrations obtained from assumptions of homogeneous media, especially using spectral fitting models of diffuse reflectance, transmittance and fluorescence [24,26,32,58–67]. In imaging and tomography applications, these concentrations are typically extracted from the absorption at a few wavelengths and used as diagnostic indexes in different applications [117], whereas point spectroscopy evaluates a larger number and often a wider range of wavelengths. However, the potential of point spectroscopy is not fully exploited if homogeneous media assumptions are made, as useful information can

potentially be extracted from such a larger number of wavelengths and depths probed. Our previous work [12] has shown that biomolecular concentrations are different depending on the probed depth by varying the probe SDD and by using a spectral fitting model for homogeneous tissue. In this study, the importance of measurements at each wavelength is considered separately. The probed depth of reflectance at each wavelength is different and incorporated into our analysis using statistical tests and machine learning methods. Analyzing DRS measurements of each wavelength separately means that information is extracted from tissue biochemistry and microstructure at multiple depths. Therefore, the importance of chromophores obtained in this study is based on more complete information compared with previous studies and exploits the full potential of DRS point spectroscopy measurements. By using our depth-resolved analysis, tissue biomarker information was interpreted by using PLSC loadings while the objective selection of optimized tissue classification parameters was determined by evaluating classification performance directly.

4.2. Spectral Features for Colorectal Cancer (CRC) Detection

The present study investigated the spectral regions and biomolecules associated with cancer development by identifying the most important spectral features for the differentiation between normal and cancerous tissues. This identification was performed by using partial least-squares (PLS). One of the advantages of using PLS methods is that they can provide insight into the variables most likely to be responsible for the differentiation between two groups via the interpretation of weights and loadings. With this in mind, PLS is typically used in exploratory studies focusing on which variables are best discriminators [113]. In molecular biology applications (e.g., metabolomics, proteomics, lipidomics, glycomics and others), these variables can be related to biomolecules via features of their generated physical processes present in the measured signal (e.g., the fluorescence emission of specific molecules in a tissue fluorescence measurement) [114]. In terms of optical techniques, this approach was used by Wang et. al., who found that NADH, collagen, and porphyrin were related to oral cancer detection by using fluorescence spectroscopy [118]. In this study, biomolecular contribution to the DRS signal is related to absorption features of tissue chromophores such as oxyhemoglobin (HbO₂) and deoxyhemoglobin (Hb), lipid and water, as well as scattering tissue features related to size and refractive index of the main scatterers. As a result, absorption and scattering features related to discrimination can be observed on the PLS loadings, since these loadings are based on the differentiation between mucosa and tumor tissues present on the DRS signal.

PLSC loadings can be interpreted based on their amplitude and spectral shape. Higher absolute amplitudes of PLSC loadings on spectral regions of absorption of specific molecules mean these molecules are associated with the PLSC(s). The same association is reinforced if the spectral shape of the PLSC loadings is similar to that of the biomolecular absorption spectra within the pertinent spectral regions. The first PLSCs are more relevant for tissue discrimination. Therefore, analyzing the absolute amplitudes and spectral shape of loadings of the first PLSCs allows us to determine which biomolecules are potentially more relevant for differentiating mucosal and cancerous tissues.

By analyzing parameters such as loadings of principal component analysis (PCA) or PLS, it is possible to determine which biomolecules are involved with the classification between groups of tissues. The loadings of PLS components are typically analyzed in metabolomics [114], where the focus is determining which variables (related to chemicals) are best discriminators for certain tissue groups (e.g., cancerous and healthy tissue) rather than understanding whether tissue classification is possible [113]. A similar rationale is used to analyze spectral features associated with other physical phenomena related to biomolecules. For optical techniques such as fluorescence spectroscopy, biomolecules include NADH, FAD, collagen, elastin, porphyrins and lipopigments [24,26,27,32,60,67,119–121], whereas Fourier-transform infrared (FTIR) spectroscopy and Raman spectroscopy probe chemicals such as amino acids, proteins, lipids, carbohydrates, nucleic acids, porphyrins, and water [47,49,122,123]. A more extensive list of altered biochemical composition

in CRC was reported by previous studies on lipidomics, proteomics, metabolomics, genomics, glycomics and other molecular biology sciences [124–133]. Out of all biomolecules in the extensive list of CRC tissue discriminators, our DRS study focuses on biomolecules with detectable absorption (chromophores) in the wavelength range between 350 and 1920 nm. These biomolecules include β -carotene, bile, bilirubin, ceroid, collagen, deoxyhemoglobin (Hb), oxyhemoglobin (HbO_2), methemoglobin (MetHb), water, lipid, and melanin.

Previous studies using PLS to investigate biomolecular features in optical measurements are scarce. By using fluorescence spectroscopy, Wang et al. [118] used PLS loadings to determine that the biomolecules involved with oral cancer detection were collagen, NADH and porphyrins. This determination was based on the peaks of the loadings as a function of wavelength, which occurred at 390 nm and 470 nm for the 320 nm excitation, and at 460 nm and 640 nm for the 360 nm excitation. Apart from the study of Wang et. al., our study used the amplitudes and shapes of PLSC loadings to identify CRC biomarkers and corresponding wavelength ranges based on the loadings of the four first PLSCs. These loadings suggest that only Hb, HbO_2 , lipid, water and light scattering in tissue contribute to PLS features to be used for tissue classification (Section 3.3). In a qualitative analysis based on the amplitude of PLSC1 loadings, Hb and HbO_2 were the most important biomolecules for tissue classification since the highest and second highest absolute amplitudes (Figure 3) occurred between 350 and 1350 nm for both probes used (Section 3.1). Light scattering, lipid and water were of secondary importance, since only PLSC2 appears to contribute to scattering, and the wavelength range between 1350 and 1900 nm has the third highest absolute amplitude only for PLSC1 of the long-SDD probe (Section 3.1). It is important to note that all PLSCs used for tissue classification exhibit spectral features of water, and all PLSCs except PLSC1 and PLSC2 of the short-SDD probe have lipid features. The absence of PLSC1 and PLSC2 features for lipids suggests that no superficial lipid signal up to 1 mm deep contributes to CRC detection, as the maximum probed depth for the short-SDD probe is ~ 1 mm. Additionally, scattering may have contributed to PLSCs other than PLSC2 even though its contribution may not show any strong trend on the PLSC loadings (Figure 3). Finally, one reason the scattering properties may not be more relevant for tissue classification is due to the relatively short source-detector distance providing a low sensitivity to alterations in scattering properties between tissue types. This means there might be a scattering-related contrast between tissues that our instrument is not sensitive to because of our future endeavors of translating our findings to an endoscopic *in vivo* study requiring that the short-SDD probe and all supporting equipment fit within the endoscope. Non-endoscopic applications targeting the classification of normal mucosa and cancer tissues relying on longer SDD and/or quantities associated with tissue scattering may still be useful to find scattering differences between these tissues.

4.3. Considerations on Biomolecular Concentrations and Probed Depth for CRC Detection

Based on the current literature evidence that Hb, HbO_2 , lipid and water concentrations differ in normal tissues and tumors, our results agree that these biomolecules can be used for tumor detection. In our study, the features of each biomolecule were illustrated in the shape of the PLSC loadings, while the molecular relevance for differentiation between colorectal mucosa and tumor could be determined with the amplitude of the loadings and higher classification performance of selected wavelength ranges. In general, there is an agreement between the wavelength ranges leading to higher classification performance and higher absolute amplitudes of the PLSC1 loadings. Since taking these ranges solely based on the loadings is subjective, we used a more objective analysis by selecting ranges of $p < 0.001$ in the *t*-test. This analysis suggested that Hb and HbO_2 information relevant to tumor detection can be collected from specific wavelength ranges within the UV-visible region and the optical window. A combination of these ranges may not lead to more accurate tissue classification. On the other hand, classification can be improved by probing water and lipid at several near-infrared (NIR) ranges which provide complementary information for

tissue discrimination. Our results regarding the importance of Hb, HbO₂, lipid, water and scattering for CRC detection agree with results based on our objective statistical analysis and add to the results of our previous work [11], which suggested that wavelengths in the optical window contribute to higher accuracies of CRC detection. When neglecting local blood oxygen saturation (StO₂) for reflectance spectral fitting under the homogeneous medium assumption, it is worth noting that lipid and scattering contributions to tissue classification become more important than contributions of total hemoglobin content [12]. However, when considering StO₂ and distribution of chromophores scattering properties over all probed tissue layers in DRS, near-infrared spectroscopy, elastic scattering spectroscopy and hyperspectral imaging, previous *in vivo* and *ex vivo* studies have shown that contributions of Hb and HbO₂ are comparable to those of lipid and water [11]. These contributions are evidenced by similar accuracies obtained when using wavelength ranges between 400 and 1000 nm (91.2 ± 0.9 accuracy) and between 1000 and 1920 nm (92.2 ± 1.3 accuracy) for the short-SDD probe [11].

In terms of wavelength ranges for differentiation between normal and cancer tissues, those in the optical window resulted in the highest classification performance for both short-SDD and long-SDD probes. This performance could have been achieved by considering the information on a number of biomolecules at variable depths and targeting wavelengths of higher light penetration. Our study suggests the latter is one of the main discriminatory factors, as probing deeper tissue layers with the long-SDD probe also led to higher performance compared with the short-SDD probe. Since more accurate tissue classification is achieved by using wavelengths at the optical window and using the long-SDD probe, the most relevant biomolecular changes associated with CRC detection, and potential carcinogenesis may occur at deeper tissue layers, which can be exploited by future studies.

Similar to the short-SDD probe, the highest classification performance was achieved by using wavelengths on the optical window (in this case, 650–950 nm). Additionally, since the performance obtained for the long-SDD probe is higher than that obtained by using the short-SDD probe, signals of deep tissue layers may contain more relevant information for tissue discrimination. This evidence can be further reinforced by the higher performance achieved by using the ultraviolet (380–400 nm) and visible (420–610 nm) wavelengths alone compared to the best performance obtained by using the short-SDD probe.

Our analysis of wavelength selection indicated that probing specific wavelength ranges for Hb, HbO₂, water and lipid absorption lead to similar classification accuracies as those achieved by using the wavelength range 350–1920 nm (Tables 2 and 3). These ranges may be used for future instrument design by keeping the geometrical configuration of the probe used in this study. In addition, extending the wavelength range towards longer NIR wavelengths may add information from other biomolecules as well as the contribution of the complementary data of water and lipid. With this in mind, future studies may investigate the complementarity of water and lipid information at wavelengths longer than 1920 nm and include features of other biomolecules absorbing at NIR and mid-infrared wavelengths (e.g., proteins, carbohydrates and nucleic acids).

4.4. Strength of the Cross-Validation of Our Model

In this study, we have used our PLS model to show the importance of objectively chosen wavelength ranges in tissue classification. Our tissue classification model was not used to assess the maximum classification performance to be obtained by testing several machine learning algorithms with our dataset. In fact, we have previously shown that higher classification performance can be obtained [11]. For correct interpretation of the influence of each wavelength range in the tissue classification, it is extremely important that our PLS model is general enough so that high classification performance parameters (sensitivity, specificity, accuracy and AUC) are not obtained due to overfitting by using large subsets of the dataset to build our classification model.

Although 5-fold cross-validation is frequently used for validation, this study used two-fold cross-validation in order to show a robust tissue classification in our dataset by using 50% of the data as a training set (two-fold) instead of 80% (5-fold). Two-fold cross-validation also allows the model to be tested in a larger dataset compared to 5-fold cross-validation, while the model is trained in a smaller subset of the total dataset. If the model is not general enough, training the model with smaller random subsets of the dataset while testing it in larger random subsets can lead to lower classification performance parameters. Therefore, robust classification through two-fold cross-validation means a stronger potential of generalization of the model upon an increase in sample size, especially compared to validation using more than 50% of the dataset for training and less than 50% for testing.

4.5. Limitations of Our Study

In order to evaluate oxygenation changes that could affect the results of the present study, we conducted a pilot observation of the DRS signal in 3 patients. We observed no significant variations in average Hb and HbO₂ of 7 mucosa sites and 7 tumor sites 15 min from the beginning of our measurements (data not shown). In this case, the DRS signal was monitored every 5 min during the collection time period. In terms of expected biochemical differences when translating our findings to *in vivo* studies, we expect a similar behavior showed by Baltussen et al. [88] in fat, tumor, and healthy colorectal wall tissues. The authors exhibited a statistically significant difference in blood concentration (%) and StO₂ (%) when measured *ex vivo* (within 1 h after resection) compared to *in vivo*. According to their study, both blood content and StO₂ increased, presumably due to an increase in blood volume in the capillaries after excision [134], exposure to air and decreased oxygen consumption by the cells in the specimen. Their study also suggests the water content decreases due to dehydration (vaporization and leakage) of the resected tissue, which was minimized in this study by keeping the tissue moist with wet wipes. However, it is important to remember that measurements by Baltussen et al. were taken by probing a different tissue depth compared to our study (probe with SDD 1.29 mm center-to-center distance) and the analyzed tissues are different from colorectal mucosa. Additionally, the behavior of StO₂ (%) is still not well understood, as results of Baltussen et al. disagrees with those of Salomatina et al. [135], who evaluated mouse ear tissues 5 to 10 min after excision (*ex vivo*) of tissue and after 24 and 72 h of storage.

Parts of our methodology are dependent on results achieved by using our dataset of almost 3000 measurements on 47 patients, which is assumed to be sufficiently robust for all analyses presented in this study. To ensure the wavelength selection was robust, the statistical test was based on all collected data. One should be aware that the wavelength selection was independent of the machine learning model. Therefore, classification performance metrics were affected by which wavelength ranges were selected, but not by the process to select wavelengths. It is worth noting that wavelengths were not used directly as descriptors/features of our machine learning (ML) models/classifiers. The four selected PLS components (PLSC) have been used as spectral features for tissue classification. A number of PLSCs have been selected as the minimum number of PLSCs before only increments of ~2% accuracy were gained upon inclusion of a new PLSC (data not shown). This threshold has been chosen to provide the best possible description of the data while avoiding overfitting. The four PLSCs were calculated on the entire dataset before two-fold cross-validation was performed for KNN models. Since PLSC loadings for tissue classification have been determined based on all the collected data (including the test set for the 20 iterations of two-fold cross-validation), it is important to note that the classification performance metrics were not calculated for optimized ML models in which spectral bands and PLSCs are selected at specific subsets of our dataset (discovery set) and subsequently training classifiers. Our classification performance metrics were used for the comparison of results obtained for each objectively/statistically selected wavelength range with the aim of placing wavelength ranges in the order of importance for tissue

classification, and subsequently, the main cancer biomarkers upon the association between these ranges and tissue chromophores and scattering properties were identified. For the estimation of classification performance metrics obtained with wider wavelength ranges, one should check our previous study [11].

Regarding the validation of our tissue classification model, the two-fold cross-validation with random sampling was performed spectrum-wise and not patient-wise. Hence, the data from different patients can be present in both training and test sets, but the same spectrum will only belong to either the training set or test set for each fold of cross-validation. However, this does not mean that the results of this study are invalid or that classification performance metrics have been overestimated. It is important to consider that the intra-patient is comparable to the inter-patient variation in our dataset (the standard deviation of measurements within each individual patient is less than twice the standard deviation of all measurements of all patients). A typical mean and standard deviation across all measured locations of one patient for both tumor and normal mucosa measured with both probes is given in Figure S1 in Supplementary Material. No clear trend has been identified in the data of specific patients, possibly due to the limited number of patients. If such a trend is identified in future studies, a patient-wise validation of our model would estimate the classification performance metrics of Tables 2 and 3 most likely as robustly. Since this trend has not been identified, a spectrum-wise validation of our tissue classification model could potentially be stronger because this validation includes the data of more patients compared with a patient-wise validation. This is especially important given that our study was limited to 47 patients.

Furthermore, the ultimate aim of our study is to classify normal mucosa and cancer tissues at each location within the same patient. Therefore, robust validation of our tissue classification model should consider both measurements of different patients as well as measurements from the same patient, but at different tissue locations, as we did in our study.

Our study focused on biochemical and DRS spectral differences between mucosa and cancer tissues, which, from a clinical perspective, would be especially useful for guidance on cancer margins and tumor delineation. Future studies exploring tumor detection will include data from tissues with non-cancer pathology. This would include a variety of non-neoplastic processes commonly seen in CRC patients including inflammatory bowel disease, radiation-induced fibrosis, and scarring following local excision of cancers.

From a research perspective, the results of our study can be used to design optimized optical instruments targeting the specific wavelength ranges and tissue probed depth, as well as to evaluate the feasibility of employing biochemical analysis methods targeting Hb, HbO₂, water and lipid for tissue classification. These methods include other optical sensors and/or combinations with existing instruments for CT, MRI, electrical measurements, mass spectrometry, and others. Simulations of light transport in tissues and DRS measurements may benefit from estimating accurate data of the chromophores most relevant for tissue classification. The same applies to the creation of 3D cell models (e.g., spheroids) mimicking mucosal and cancerous tissues in tissue engineering studies, where the right tissue types and thicknesses should accurately reproduce the tissue biochemistry.

5. Conclusions

In the present study, we evaluated the most important tissue biomarkers and wavelength ranges for CRC detection by using diffuse reflectance spectroscopy in the visible/near-infrared wavelength range from 350 nm to 1920 nm. In this range, the biomolecules most relevant for classification between normal and cancer tissues were Hb, HbO₂, lipid and water. Tissue classification using Hb and HbO₂ data can be achieved with wavelength ranges 350–590 nm or 600–1230 nm for superficial tissue (short-SDD probe), and 380–400 nm, 420–610 nm, and 650–950 nm for deeper tissue layers (long-SDD probe). Information collected in those wavelength ranges is redundant and the combination between them did not enhance the classification. On the other hand, water and lipid information are

complementary and may improve cancer detection and investigation of carcinogenesis upon extension of the wavelength range. Our results suggested that information from deeper tissue layers either accessed by probing the optical window or through the long-SDD probe can differentiate normal and cancer tissues more accurately. Wavelength ranges and probe geometrical configuration used in this study may be used for more specific future instrument design. From a practical perspective, these wavelength ranges and geometrical configurations will be used to develop an optical system for CRC detection during colonoscopy and intra-operatively. By optimizing the features for discrimination between normal and cancerous tissues, tissue identification is performed in real-time with a single reading of about 2–3 s. By means of the integration of real-time tissue identification into a flexible fiberoptic probe which could be passed down the working channel of an endoscope, optical spectroscopy may provide a powerful tool that can be used to detect cancer cells and direct management in real time. In addition, this spectroscopic technique can detect more subtle mucosal abnormalities such as sessile serrated polyps, which may be difficult to identify during colonoscopy. Once instruments with wavelength ranges and probe geometrical configuration found in this study have been miniaturized and integrated into colonoscopies, next-generation instruments can be manufactured and their impact in reducing CRC morbidity and morbidity can be assessed in future *in vivo* studies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers14225715/s1>, Figure S1: Mean and standard deviation across broadband reflectance spectra of all measured locations of both tumor and normal mucosa measured with (A) the short-SDD and (B) the long-SDD probe for one patient.

Author Contributions: M.S.N. Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing—Original Draft, Writing—Review and Editing, Visualization, Supervision, Software, Validation, Project administration, S.M. Investigation, Data curation, M.A. Investigation, Data Curation, A.M. Methodology, Investigation, Data Curation, Writing—Review and Editing, C.S. Methodology, Investigation, Data Curation, Writing—Review and Editing, S.K. Investigation, Resources, Data Curation, S.A.-E. Conceptualization, Methodology, Resources, Writing—Review and Editing, Supervision, Project administration, Funding Acquisition, M.O. Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing—Review and Editing, Supervision, Project administration, Funding Acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This study received financial support from Science Foundation Ireland (SFI): grant ID SFI/15/RP/2828. Shauni Fitzgerald and Edmund Manning were funded by the Mercy Foundation.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Clinical Research Ethics Committee of University College Cork (protocol approval numbers of ECM 4 (n) 07/03/18 and ECM 3 (hhh) 18/06/2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors acknowledge Eduardo Moriyama for the advice and logistics of the project associated with this manuscript, Shree Krishnamoorthy for the discussions on the same project, Haiyang Li for the design and assembly of the holder of the DRS system, Noel Lynch for specimen retrieval of the first set of patients of this study and all the logistics related to it, Vivienne Curran, Aoife Foyle, Una McAuliffe, Evelyn Flanagan, Shauni Fitzgerald, and Edmund Manning for all the logistics of clinical procedures, research ethics documents, and search for patient information.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [\[CrossRef\]](#)
- Ferlay, J.; Ervik, M.; Lam, F.; Colombet, M.; Mery, L.; Piñeros, M.; Znaor, A.; Soerjomataram, I.; Bray, F. *Global Cancer Observatory: Cancer Today*; International Agency for Research on Cancer: Lyon, France, 2018; Available online: <https://gco.iarc.fr/> (accessed on 8 October 2022).
- Shaukat, A.; Levin, T.R. Current and future colorectal cancer screening strategies. *Nat. Rev. Gastroenterol. Hepatol.* **2022**, *19*, 521–531. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ichkhanian, Y.; Zuchelli, T.; Watson, A.; Piraka, C. Evolving management of colorectal polyps. *Ther. Adv. Gastrointest. Endosc.* **2021**, *14*, 26317745211047010. [\[CrossRef\]](#) [\[PubMed\]](#)
- Pogue, B.W. Optics in the molecular imaging race. *Opt. Photonics News* **2015**, *26*, 24–31. [\[CrossRef\]](#)
- Nogueira, M.S.; Raju, M.; Gunther, J.; Maryam, S.; Amissah, M.; Lu, H.; Killeen, S.; O’Riordain, M.; Andersson-Engels, S. Accurate colorectal cancer detection and delineation by probing superficial and deeper tissue biochemistry and microstructure using diffuse reflectance spectroscopy. In Proceedings of the Molecular-Guided Surgery: Molecules, Devices, and Applications VIII, San Francisco, CA, USA, 22 January–28 February 2022; Volume 11943.
- Nogueira, M.S.; Raju, M.; Komolibus, K.; Grygoryev, K.; Andersson-Engels, S. Assessment of tissue biochemical and optical scattering changes due to hypothermic organ preservation: A preliminary study in mouse organs. *J. Phys. D Appl. Phys.* **2021**, *54*, 37.
- Nogueira, M.S.; Raju, M.; Gunther, J.; Maryam, S.; Amissah, M.; Killeen, S.; O’Riordain, M.; Andersson-Engels, S. Optical determination of superficial and deeper tissue biochemistry and microstructure for delineation and early detection of colorectal cancer. In Proceedings of the European Conference on Biomedical Optics, Munich, Germany, 20–24 June 2021; p. ETu3A-1.
- Nogueira, M.S.; Maryam, S.; Amissah, M.; Lynch, N.; Killeen, S.; O’Riordain, M.; Andersson-Engels, S. Benefit of extending near-infrared wavelength range of diffuse reflectance spectroscopy for colorectal cancer detection using machine learning. In Proceedings of the European Conference on Biomedical Optics, Munich, Germany, 20–24 June 2021; p. EW4A-16.
- Nogueira, M.S.; Raju, M.; Gunther, J.; Grygoryev, K.; Komolibus, K.; Lu, H.; Andersson-Engels, S. Diffuse reflectance spectroscopy for determination of optical properties and chromophore concentrations of mice internal organs in the range of 350 nm to 1860 nm. In Proceedings of the Biophotonics: Photonic Solutions for Better Health Care VI, Strasbourg, France, 22–26 April 2018; Volume 10685. [\[CrossRef\]](#)
- Nogueira, M.S.; Maryam, S.; Amissah, M.; Lu, H.; Lynch, N.; Killeen, S.; O’Riordain, M.; Andersson-Engels, S. Evaluation of wavelength ranges and tissue depth probed by diffuse reflectance spectroscopy for colorectal cancer detection. *Sci. Rep.* **2021**, *11*, 798. [\[CrossRef\]](#)
- Nogueira, M.S.; Raju, M.; Gunther, J.; Maryam, S.; Amissah, M.; Lu, H.; Killeen, S.; O’Riordain, M.; Andersson-Engels, S. Tissue biomolecular and microstructure profiles in optical colorectal cancer delineation. *J. Phys. D Appl. Phys.* **2021**, *54*, 454002. [\[CrossRef\]](#)
- Mousavi, M.; Moriyama, L.T.T.; Grecco, C.; Nogueira, M.S.; Svanberg, K.; Kurachi, C.; Andersson-Engels, S. Photodynamic therapy dosimetry using multiexcitation multiemission wavelength: Toward real-time prediction of treatment outcome. *J. Biomed. Opt.* **2020**, *25*, 63812. [\[CrossRef\]](#)
- Nogueira, M.S.; Maryam, S.; Amissah, M.; Lynch, N.; Killeen, S.; Lu, H.; O’Riordain, M.; Andersson-Engels, S. Improving colorectal cancer detection by extending the near-infrared wavelength range and tissue probed depth of diffuse reflectance spectroscopy: A support vector machine approach. In Proceedings of the Optical Biopsy XX: Toward Real-Time Spectroscopic Imaging and Diagnosis, San Francisco, CA, USA, 22 January–28 February 2022; Volume 11954.
- Nogueira, M.S.; Amissah, M.; Maryam, S.; Lynch, N.; Killeen, S.; O’Riordain, M.; Andersson-Engels, S. Optimization of tissue classification for colorectal cancer detection using support vector machines and diffuse reflectance spectroscopy. In Proceedings of the European Conference on Biomedical Optics, Munich, Germany, 20–24 June 2021; p. EW4A-17.
- Nogueira, M.S.; Jayet, B.; Matias, J.S.; Gunther, J.E.; Tyndall, C.; Andersson-Engels, S. Biophotonics web application for computer simulations in diffuse optics: Fostering multidisciplinary education and research. In Proceedings of the Optical Interactions with Tissue and Cells XXXIII; and Advanced Photonics in Urology, San Francisco, CA, USA, 22 January–28 February 2022; Volume 11958.
- Saito Nogueira, M.; Gunther, J.E.; Jayet, B.; Souza Matias, J.; Tyndall, C.; Andersson-Engels, S. Biophotonics computer app: Fostering multidisciplinary distance self-paced learning with a user-friendly interface. In Proceedings of the OSA Technical Digest, San Francisco, CA, USA, 1–6 February 2021; pp. 1–2.
- Nogueira, M.S.; Gunther, J.E.; Komolibus, K.; Andersson-Engels, S. User-friendly graphical user interface for simulating tissue optical properties and fluence rates: Improving students learning in tissue optics. In Proceedings of the Optical Interactions with Tissue and Cells XXXI, San Francisco, CA, USA, 1–6 February 2020; Volume 11238.
- Mourant, J.R.; Freyer, J.P.; Hielscher, A.H.; Eick, A.A.; Shen, D.; Johnson, T.M. Mechanisms of light scattering from biological cells relevant to noninvasive optical-tissue diagnostics. *Appl. Opt.* **1998**, *37*, 3586–3593. [\[CrossRef\]](#)
- Mourant, J.R.; Johnson, T.M.; Carpenter, S.; Guerra, A.; Aida, T.; Freyer, J.P. Polarized angular-dependent spectroscopy of epithelial cells and epithelial cell nuclei to determine the size scale of scattering structures. *J. Biomed. Opt.* **2002**, *7*, 378–387. [\[CrossRef\]](#)

21. Jacques, S.L. Optical properties of biological tissues: A review. *Phys. Med. Biol.* **2013**, *58*, R37. [CrossRef] [PubMed]
22. Nogueira, M.S. *Fluorescence Lifetime Spectroscopy for Diagnosis of Clinically Similar Skin Lesions*; Universidade de São Paulo: São Paulo, Brazil, 2016.
23. Nogueira, M.S.; Kurachi, C. Assessing the photoaging process at sun exposed and non-exposed skin using fluorescence lifetime spectroscopy. In Proceedings of the Optical Biopsy XIV: Toward Real-Time Spectroscopic Imaging and Diagnosis, San Francisco, CA, USA, 13–18 February 2016; Volume 9703.
24. Cosci, A.; Nogueira, M.S.; Pratavieira, S.; Takahama, A.; de Souza Azevedo, R.; Kurachi, C.; Kurachi, C. Time-resolved fluorescence spectroscopy for clinical diagnosis of actinic cheilitis: Erratum. *Biomed. Opt. Express* **2016**, *7*, 4210–4219. [CrossRef] [PubMed]
25. Lacerenza, M.; Pacheco, A.; Sekar, S.K.V.; Nogueira, M.S.; Buttafava, M.; Tosi, A.; Pifferi, A.; Contini, D.; Andersson-Engels, S. Functional monitoring of lung tissue using a hybrid hyperspectral Time-Resolved GASMAS system: A systematic study on ex vivo sample. In Proceedings of the Optical Tomography and Spectroscopy, Washington, DC, USA, 20–23 April 2020; p. SW1D-2.
26. Pires, L.; Nogueira, M.S.; Pratavieira, S.; Moriyama, L.T.; Kurachi, C. Time-resolved fluorescence lifetime for cutaneous melanoma detection. *Biomed. Opt. Express* **2014**, *5*, 3080. [CrossRef]
27. Kurachi, C.; Pires, L.; Nogueira, M.S.; Pratavieira, S. Lifetime fluorescence for the detection of skin lesions. In Proceedings of the Biomedical Optics 2014, Miami, FL, USA, 26–30 April 2014. [CrossRef]
28. Nogueira, M.S.; Cosci, A.; Pratavieira, S.; Takahama, A., Jr.; Azevedo, R.S.; Kurachi, C. Evaluation of actinic cheilitis using fluorescence lifetime spectroscopy. In Proceedings of the Optical Biopsy XIV: Toward Real-Time Spectroscopic Imaging and Diagnosis, San Francisco, CA, USA, 13–18 February 2016; Volume 9703.
29. Salvio, A.G.; Ramirez, D.P.; Inada, N.M.; Stringasci, M.D.; Nogueira, M.S.; Bagnato, V.S. Fractionated Illumination in a Single Visit Photodynamic Therapy for Basal Cell Carcinoma; Book of Abstracts. 2017. Available online: <https://www.internationalphotodynamic.com/> (accessed on 8 October 2022).
30. Salvio, A.G.; Ramirez, D.P.; Nogueira, M.S.; Stringasci, M.D.; Oliveira, E.R.; Inada, N.M.; Bagnato, V.S. Evaluation of Pain and Treatment Effect during Large Area Photodynamic Therapy in 140 Patients with Widespread Actinic Keratosis of Upper Limbs; Book of Abstracts. 2017. Available online: <https://www.internationalphotodynamic.com/> (accessed on 8 October 2022).
31. Nogueira, M.S.; Cosci, A.; Rosa, R.G.T.; Salvio, A.G.; Pratavieira, S.; Kurachi, C. Portable fluorescence lifetime spectroscopy system for in-situ interrogation of biological tissues. *J. Biomed. Opt.* **2017**, *22*, 121608.
32. de Paula Campos, C.; de Paula D’Almeida, C.; Nogueira, M.S.; Moriyama, L.T.; Pratavieira, S.; Kurachi, C. Fluorescence spectroscopy in the visible range for the assessment of UVB radiation effects in hairless mice skin. *Photodiagnosis Photodyn. Ther.* **2017**, *20*, 21–27. [CrossRef]
33. Nogueira, M.S.; Guimarães, F.E.G. Photophysical processes on biological tissues and photodynamic therapy using steady-state and time-resolved fluorescence techniques: Diagnosis applications, dosimetry and photodegradation kinetics. In *Livro de Resumos; SIBiUSP-Integrated USP Libraries*: São Paulo, Brazil, 2016.
34. Ono, B.A.; Nogueira, M.; Pires, L.; Pratavieira, S.; Kurachi, C. Subcellular localization and photodynamic activity of Photodithazine (glucosamine salt of chlorin e6) in murine melanoma B16-F10: An in vitro and in vivo study. In Proceedings of the Optical Methods for Tumor Treatment and Detection: Mechanisms and Techniques in Photodynamic Therapy XXVII, San Francisco, CA, USA, 27 January–1 February 2018; Volume 1047616, p. 44. [CrossRef]
35. Nogueira, M.S.; Bagnato, V.S.; Panhoca, V.H. Effectiveness of whitening treatments employing violet illumination alone or combined with bleaching agents. In Proceedings of the Lasers in Dentistry XXVIII, San Francisco, CA, USA, 22 January–28 February 2022; Volume 11942.
36. Panhóca, V.H.; Nogueira, M.S.; Bagnato, V.S. Treatment of facial nerve palsies with laser and endermotherapy: A report of two cases. *Laser Phys. Lett.* **2020**, *18*, 15601. [CrossRef]
37. Panhóca, V.H.; Nogueira, M.S.; Bagnato, V.S. Laser and vacuum therapy for treatment of facial nerve palsies. In Proceedings of the Imaging, Therapeutics, and Advanced Technology in Head and Neck Surgery and Otolaryngology 2022, San Francisco, CA, USA, 22 January–28 February 2022; Volume 11935, pp. 39–48.
38. Bomfin, L.S.; Kitakawa, D.; Nogueira, M.S.; de Silva, L.F. Low Level Laser Therapy as adjuvant treatment for lower lip lesion. In Proceedings of the Latin America Optics and Photonics Conference, Recife, Brazil, 7–11 August 2022; p. M2B-6.
39. Nogueira, M.S.; Lacerenza, M.; Sekar, S.K.V.; Buttafava, M.; Pifferi, A.; Tosi, A.; Contini, D.; Andersson-Engels, S. Broadband extraction of tissue optical properties using a portable hybrid time-resolved continuous wave instrumentation: Characterization of ex vivo organs. In Proceedings of the Clinical and Translational Biophotonics, Washington, DC, USA, 20–23 April 2020; p. TM2B-3.
40. Nogueira, M.S.; Bagnato, V.S.; Panhoca, V.H. Characterization of teeth fluorescence properties due to coffee pigmentation: Towards optimization of quantitative light-induced fluorescence for tooth color assessment. In Proceedings of the Optical Interactions with Tissue and Cells XXXI, San Francisco, CA, USA, 1–6 February 2020; Volume 11238.
41. Nogueira, M.S.; Pinto Junior, F.F.; Caface, R.A.; de Oliveira, K.T.; Guimarães, F.E.G. Optimization of curcumin formulations for fluorescence-based applications. In Proceedings of the Optical Interactions with Tissue and Cells XXXI, San Francisco, CA, USA, 1–6 February 2020; Volume 11238.
42. Nogueira, M.S.; Junior, F.F.P.; Caface, R.A.; de Oliveira, K.T.; Bagnato, V.S.; Guimarães, F.E.G. Characterization of photophysical properties of curcumin for theranostics of neurodegenerative diseases. In Proceedings of the Optical Interactions with Tissue and Cells XXX, San Francisco, CA, USA, 2–7 February 2019; Volume 10876.

43. Nogueira, M.S.; Komolibus, K.; Grygoryev, K.; Gunther, J.E.; Andersson-Engels, S. Fluorescence spectroscopy of mouse organs using ultraviolet excitation: Towards assessment of organ viability for transplantation. In Proceedings of the Optical Interactions with Tissue and Cells XXX, San Francisco, CA, USA, 2–7 February 2019; Volume 10876.
44. Nogueira, M.S.; Panhóca, V.H.; Bagnato, V.S. Fluorescence spectroscopy analysis of light-induced tooth whitening. In Proceedings of the Optical Interactions with Tissue and Cells XXX, San Francisco, CA, USA, 2–7 February 2019; Volume 10876.
45. De Andrade, C.T.; Nogueira, M.S.; Kanick, S.C.; Marra, K.; Gunn, J.; Andreozzi, J.; Samkoe, K.S.; Kurachi, C.; Pogue, B.W. Optical spectroscopy of radiotherapy and photodynamic therapy responses in normal rat skin shows vascular breakdown products. In Proceedings of the Optical Methods for Tumor Treatment and Detection: Mechanisms and Techniques in Photodynamic Therapy XXV, San Francisco, CA, USA, 13–18 February 2016; Volume 9694.
46. Saito Nogueira, M.; Ribeiro, V.; Pires, M.; Peralta, F.; Carvalho, L.F.D.C.E.S.D. Biochemical Profiles of In Vivo Oral Mucosa by Using a Portable Raman Spectroscopy System. *Optics* **2021**, *2*, 134–147. [\[CrossRef\]](#)
47. Carvalho, L.F.C.S.; Nogueira, M.S.; Bhattacharjee, T.; Neto, L.P.M.; Daun, L.; Mendes, T.O.; Rajasekaran, R.; Chagas, M.; Martin, A.A.; Soares, L.E.S. In vivo Raman spectroscopic characteristics of different sites of the oral mucosa in healthy volunteers. *Clin. Oral Investig.* **2019**, *23*, 3021–3031. [\[CrossRef\]](#)
48. Leal, L.B.; Nogueira, M.S.; Mageski, J.G.A.; Martini, T.P.; Barauna, V.G.; dos Santos, L.; Carvalho, L.F.D.C.E.S.D. Diagnosis of Systemic Diseases Using Infrared Spectroscopy: Detection of Iron Overload in Plasma—Preliminary Study. *Biol. Trace Element Res.* **2021**, *199*, 3737–3751. [\[CrossRef\]](#)
49. Carvalho, L.F.C.S.; Nogueira, M.S. New insights of Raman spectroscopy for oral clinical applications. *Analyst* **2018**, *143*, 6037–6048. [\[CrossRef\]](#)
50. Carvalho, L.F.C.S.; Nogueira, M.S.; Neto, L.P.M.; Bhattacharjee, T.T.; Martin, A.A. Raman spectral post-processing for oral tissue discrimination—A step for an automatized diagnostic system. *Biomed. Opt. Express* **2017**, *8*, 5218, Erratum in *Biomed. Opt. Express* **2018**, *9*, 649. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Nogueira, M.S. Biophotonic telemedicine for disease diagnosis and monitoring during pandemics: Overcoming COVID-19 and shaping the future of healthcare. *Photodiagnosis Photodyn. Ther.* **2020**, *31*, 101836. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Maryam, S.; Nogueira, M.S.; Krishnamoorthy, S.; Sekar, S.K.V.; Lu, H.; Gautam, R.; Burke, R.; Andersson-Engels, S.; Riordain, R.N.; Sheahan, P. Multi-configuration Raman spectrometer for early stage diagnosis of oral cancer. In Proceedings of the Biomedical Vibrational Spectroscopy 2022: Advances in Research and Industry, San Francisco, CA, USA, 22 January–28 February 2022; Volume 11957, pp. 20–26.
53. Nogueira, M.S.; Barreto, A.L.; Furukawa, M.; Rovai, E.S.; Bastos, A.; Bertoncello, G.; e Silva de Carvalho, L.F.d.C. FTIR spectroscopy as a point of care diagnostic tool for diabetes and periodontitis: A saliva analysis approach. *Photodiagn. Photodyn. Ther.* **2022**, *40*, 103036. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Carnevali, A.C.; Leal, L.; Scherma, A.; Morais, C.; Martin, F.; Bonnier, F.; Baker, M.; Byrne, H.J.; Chagas e Silva Carvalho, L.F.; Nogueira, M.S. Identification of diabetic patients via urine analysis by FTIR: Preliminary study (Conference Presentation). In Proceedings of the Photonic Diagnosis and Treatment of Infections and Inflammatory Diseases II, San Francisco, CA, USA, 2–7 February 2019; Volume 10863.
55. Ferreira, M.C.C.; Monteiro, G.R.; Peralta, F.; Castro, P.A.A.; Zzell, D.; Nogueira, M.S.; Carvalho, L.F.C.S. Assessment of bound water of saliva samples by using FT-IR spectroscopy. In Proceedings of the Latin America Optics and Photonics Conference, Recife, Brazil, 7–11 August 2022; p. M4B-1.
56. Nogueira, M.S.; Leal, L.B.; Marcarini, W.D.; Pimentel, R.L.; Muller, M.; Vassallo, P.F.; Campos, L.C.G.; Dos Santos, L.; Luiz, W.B.; Mill, J.G.; et al. Rapid diagnosis of COVID-19 using FT-IR ATR spectroscopy and machine learning. *Sci. Rep.* **2021**, *11*, 15409. [\[CrossRef\]](#)
57. Chow, K.K.; Short, M.; Lam, S.; McWilliams, A.; Zeng, H. A Raman cell based on hollow core photonic crystal fiber for human breath analysis. *Med. Phys.* **2014**, *41*, 92701. [\[CrossRef\]](#)
58. Bydlon, T.M.; Nachabé, R.; Ramanujam, N.; Sterenborg, H.J.C.M.; Hendriks, B.H.W. Chromophore based analyses of steady-state diffuse reflectance spectroscopy: Current status and perspectives for clinical adoption. *J. Biophotonics* **2015**, *8*, 9–24. [\[CrossRef\]](#)
59. Nachabé, R.; Hendriks, B.H.W.; van der Voort, M.; Desjardins, A.E.; Sterenborg, H.J.C.M. Estimation of biological chromophores using diffuse optical spectroscopy: Benefit of extending the UV-VIS wavelength range to include 1000 to 1600 nm. *Biomed. Opt. Express* **2010**, *1*, 1432–1442. [\[CrossRef\]](#)
60. Nogueira, M.S.; Brugnera, J.A.; Bagnato, V.S.; Panhóca, V.H. Evaluation of the whitening effectiveness of violet illumination alone or combined with hydrogen peroxide gel. *Photobiomodulation Photomed. Laser Surg.* **2021**, *39*, 395–402. [\[CrossRef\]](#)
61. Nachabé, R.; Evers, D.J.; Hendriks, B.H.W.; Lucassen, G.W.; van der Voort, M.; Wesseling, J.; Ruers, T.J.M. Effect of bile absorption coefficients on the estimation of liver tissue optical properties and related implications in discriminating healthy and tumorous samples. *Biomed. Opt. Express* **2011**, *2*, 600–614. [\[CrossRef\]](#)
62. Nogueira, M.S.; Matthews, R.; Killeen, S.; O'Riordain, M.; Andersson-Engels, S. Colorectal cancer detection based on the extraction of scattering properties and biochemical concentrations from fluorescence spectroscopy measurements. In Proceedings of the Clinical and Translational Biophotonics, Fort Lauderdale, FL, USA, 24–27 April 2022; 24–27 April 2022; p. TS2B-5. [\[CrossRef\]](#)
63. Swartling, J.; Pifferi, A.; Enejder, A.M.K.; Andersson-Engels, S. Accelerated Monte Carlo models to simulate fluorescence spectra from layered tissues. *JOSA A* **2003**, *20*, 714–727. [\[CrossRef\]](#)

64. Muller, M.; Hendriks, B.H.W. Recovering intrinsic fluorescence by Monte Carlo modeling. *J. Biomed. Opt.* **2013**, *18*, 27009. [[CrossRef](#)] [[PubMed](#)]
65. Chang, S.K.; Mar\`in, N.; Follen, M.; Richards-Kortum, R.R. Model-based analysis of clinical fluorescence spectroscopy for in vivo detection of cervical intraepithelial dysplasia. *J. Biomed. Opt.* **2006**, *11*, 24008. [[CrossRef](#)] [[PubMed](#)]
66. Pfefer, T.J.; Wang, Q.; Drezek, R.A. Monte Carlo modeling of time-resolved fluorescence for depth-selective interrogation of layered tissue. *Comput. Methods Programs Biomed.* **2011**, *104*, 161–167. [[CrossRef](#)] [[PubMed](#)]
67. Marcu, L.; French, P.M.W.; Elson, D.S. *Fluorescence Lifetime Spectroscopy and Imaging: Principles and Applications in Biomedical Diagnostics*; CRC Press: Boca Raton, FL, USA, 2014.
68. Pogue, B.W.; Poplack, S.P.; McBride, T.O.; Wells, W.A.; Osterman, K.S.; Osterberg, U.L.; Paulsen, K.D. Quantitative hemoglobin tomography with diffuse near-infrared spectroscopy: Pilot results in the breast. *Radiology* **2001**, *218*, 261–266. [[CrossRef](#)] [[PubMed](#)]
69. Heffer, E.L.; Fantini, S. Quantitative oximetry of breast tumors: A near-infrared method that identifies two optimal wavelengths for each tumor. *Appl. Opt.* **2002**, *41*, 3827–3839. [[CrossRef](#)] [[PubMed](#)]
70. Cerussi, A.E.; Berger, A.J.; Bevilacqua, F.; Shah, N.; Jakubowski, D.; Butler, J.; Holcombe, R.F.; Tromberg, B.J. Sources of absorption and scattering contrast for near-infrared optical mammography. *Acad. Radiol.* **2001**, *8*, 211–218. [[CrossRef](#)]
71. Svensson, T.; Swartling, J.; Taroni, P.; Torricelli, A.; Lindblom, P.; Ingvar, C.; Andersson-Engels, S. Characterization of normal breast tissue heterogeneity using time-resolved near-infrared spectroscopy. *Phys. Med. Biol.* **2005**, *50*, 2559. [[CrossRef](#)]
72. Chance, B.; Nioka, S.; Zhang, J.; Conant, E.F.; Hwang, E.; Briest, S.; Orel, S.G.; Schnall, M.D.; Czerniecki, B.J. Breast cancer detection based on incremental biochemical and physiological properties of breast cancers: A six-year, two-site study1. *Acad. Radiol.* **2005**, *12*, 925–933. [[CrossRef](#)]
73. Nilsson, J.H.; Reistad, N.; Brange, H.; Öberg, C.-F.; Stureson, C. Diffuse reflectance spectroscopy for surface measurement of liver pathology. *Eur. Surg. Res.* **2017**, *58*, 40–50. [[CrossRef](#)]
74. Conover, D.L.; Fenton, B.M.; Foster, T.H.; Hull, E.L. An evaluation of near infrared spectroscopy and cryospectrophotometry estimates of haemoglobin oxygen saturation in a rodent mammary tumour model. *Phys. Med. Biol.* **2000**, *45*, 2685. [[CrossRef](#)]
75. Steen, R.G.; Kitagishi, K.; Morgan, K. In vivo measurement of tumor blood oxygenation by near-infrared spectroscopy: Immediate effects of pentobarbital overdose or carmustine treatment. *J. Neurooncol.* **1994**, *22*, 209–220. [[CrossRef](#)]
76. Kragh, M.; Quistorff, B.; Horsman, M.R.; Kristjansen, P.E.G. Acute effects of vascular modifying agents in solid tumors assessed by noninvasive laser Doppler flowmetry and near infrared spectroscopy. *Neoplasia* **2002**, *4*, 263–267. [[CrossRef](#)] [[PubMed](#)]
77. Kragh, M.; Quistorff, B.; Lund, E.L.; Kristjansen, P.E.G. Quantitative estimates of vascularity in solid tumors by non-invasive near-infrared spectroscopy. *Neoplasia* **2001**, *3*, 324–330. [[CrossRef](#)] [[PubMed](#)]
78. Gu, Y.; Chen, W.R.; Xia, M.; Jeong, S.W.; Liu, H. Effect of Photothermal Therapy on Breast Tumor Vascular Contents: Noninvasive Monitoring by Near-infrared Spectroscopy. *Photochem. Photobiol.* **2007**, *81*, 1002–1009. [[CrossRef](#)]
79. Howe, F.A.; Connelly, J.P.; Robinson, S.P.; Springett, R.; Griffiths, J.R. The effects of tumour blood flow and oxygenation modifiers on subcutaneous tumours as determined by NIRS. In *Oxygen Transport to Tissue XXVI*; Springer: Boston, MA, USA, 2005; pp. 75–81.
80. Hirosawa, N.; Sakamoto, Y.; Katayama, H.; Tonooka, S.; Yano, K. In vivo investigation of progressive alterations in rat mammary gland tumors by near-infrared spectroscopy. *Anal. Biochem.* **2002**, *305*, 156–165. [[CrossRef](#)]
81. Xia, M.; Kodibagkar, V.; Liu, H.; Mason, R.P. Tumour oxygen dynamics measured simultaneously by near-infrared spectroscopy and 19F magnetic resonance imaging in rats. *Phys. Med. Biol.* **2005**, *51*, 45. [[CrossRef](#)]
82. Yu, G.; Durduran, T.; Zhou, C.; Wang, H.-W.; Putt, M.E.; Saunders, H.M.; Sehgal, C.M.; Glatstein, E.; Yodh, A.G.; Busch, T.M. Noninvasive monitoring of murine tumor blood flow during and after photodynamic therapy provides early assessment of therapeutic efficacy. *Clin. Cancer Res.* **2005**, *11*, 3543–3552. [[CrossRef](#)]
83. Roy, H.K.; Gomes, A.; Turzhitsky, V.; Goldberg, M.J.; Rogers, J.; Ruderman, S.; Young, K.L.; Kromine, A.; Brand, R.E.; Jameel, M.; et al. Spectroscopic microvascular blood detection from the endoscopically normal colonic mucosa: Biomarker for neoplasia risk. *Gastroenterology* **2008**, *135*, 1069–1078. [[CrossRef](#)]
84. Wang, H.-W.; Jiang, J.-K.; Lin, C.-H.; Lin, J.-K.; Huang, G.-J.; Yu, J.-S. Diffuse reflectance spectroscopy detects increased hemoglobin concentration and decreased oxygenation during colon carcinogenesis from normal to malignant tumors. *Opt. Express* **2009**, *17*, 2805–2817. [[CrossRef](#)]
85. Zonios, G.; Perelman, L.T.; Backman, V.; Manoharan, R.; Fitzmaurice, M.; Van Dam, J.; Feld, M.S. Diffuse reflectance spectroscopy of human adenomatous colon polyps in vivo. *Appl. Opt.* **1999**, *38*, 6628–6637. [[CrossRef](#)]
86. Dhar, A.; Johnson, K.S.; Novelli, M.R.; Bown, S.G.; Bigio, I.J.; Lovat, L.B.; Bloom, S.L. Elastic scattering spectroscopy for the diagnosis of colonic lesions: Initial results of a novel optical biopsy technique. *Gastrointest. Endosc.* **2006**, *63*, 257–261. [[CrossRef](#)] [[PubMed](#)]
87. Mourant, J.R.; Bigio, I.J.; Boyer, J.D.; Johnson, T.M.; Lacey, J.; Bohorfoush, A.G.; Mellow, M.H. Elastic scattering spectroscopy as a diagnostic tool for differentiating pathologies in the gastrointestinal tract: Preliminary testing. *J. Biomed. Opt.* **1996**, *1*, 192–200. [[CrossRef](#)] [[PubMed](#)]
88. Baltussen, E.J.M.; Brouwer De Koning, S.G.; Hendriks, B.H.W.; Jóźwiak, K.; Sterenberg, H.J.C.M.; Ruers, T.J.M. Comparing in vivo and ex vivo fiberoptic diffuse reflectance spectroscopy in colorectal cancer. *Transl. Biophotonics* **2019**, *1*, e201900008. [[CrossRef](#)]

89. De Koning, S.G.B.; Baltussen, E.J.M.; Karakullukcu, M.B.; Dashtbozorg, B.; Smit, L.A.; Dirven, R.; Hendriks, B.H.W.; Sterenberg, H.J.C.M.; Ruers, T.J.M. Toward complete oral cavity cancer resection using a handheld diffuse reflectance spectroscopy probe. *J. Biomed. Opt.* **2018**, *23*, 121611.
90. De Boer, L.L.; Molenkamp, B.G.; Bydlon, T.M.; Hendriks, B.H.W.; Wesseling, J.; Sterenberg, H.; Ruers, T.J.M. Fat/water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries. *Breast Cancer Res. Treat.* **2015**, *152*, 509–518. [[CrossRef](#)] [[PubMed](#)]
91. Soares, J.S.; Barman, I.; Dingari, N.C.; Volynskaya, Z.; Liu, W.; Klein, N.; Plecha, D.; Dasari, R.R.; Fitzmaurice, M. Diagnostic power of diffuse reflectance spectroscopy for targeted detection of breast lesions with microcalcifications. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 471–476. [[CrossRef](#)]
92. Spliethoff, J.W.; Evers, D.J.; Klomp, H.M.; van Sandick, J.W.; Wouters, M.W.; Nachabe, R.; Lucassen, G.W.; Hendriks, B.H.W.; Wesseling, J.; Ruers, T.J.M. Improved identification of peripheral lung tumors by using diffuse reflectance and fluorescence spectroscopy. *Lung Cancer* **2013**, *80*, 165–171. [[CrossRef](#)]
93. Evers, D.J.; Nachabe, R.; Hompes, D.; Van Coevorden, F.; Lucassen, G.W.; Hendriks, B.H.W.; van Velthuysen, M.-L.; Wesseling, J.; Ruers, T.J.M. Optical sensing for tumor detection in the liver. *Eur. J. Surg. Oncol.* **2013**, *39*, 68–75. [[CrossRef](#)]
94. Tanis, E.; Evers, D.J.; Spliethoff, J.W.; Pully, V.V.; Kuhlmann, K.; van Coevorden, F.; Hendriks, B.H.W.; Sanders, J.; Prevoo, W.; Ruers, T.J.M. In vivo tumor identification of colorectal liver metastases with diffuse reflectance and fluorescence spectroscopy. *Lasers Surg. Med.* **2016**, *48*, 820–827. [[CrossRef](#)]
95. Baltussen, E.J.M.; Snæbjörnsson, P.; De Koning, S.G.B.; Sterenberg, H.J.C.M.; Aalbers, A.G.J.; Kok, N.; Beets, G.L.; Hendriks, B.H.W.; Kuhlmann, K.F.D.; Ruers, T.J.M. Diffuse reflectance spectroscopy as a tool for real-time tissue assessment during colorectal cancer surgery. *J. Biomed. Opt.* **2017**, *22*, 106014. [[CrossRef](#)]
96. Baltussen, E.J.M.; Kok, E.N.D.; de Koning, S.G.B.; Sanders, J.; Aalbers, A.G.J.; Kok, N.F.M.; Beets, G.L.; Flohil, C.C.; Bruin, S.C.; Kuhlmann, K.F.D.; et al. Hyperspectral imaging for tissue classification, a way toward smart laparoscopic colorectal surgery. *J. Biomed. Opt.* **2019**, *24*, 16002. [[CrossRef](#)] [[PubMed](#)]
97. Baltussen, E.J.M.; de Koning, S.G.; Sanders, J.; Aalbers, A.G.J.; Kok, N.F.M.; Beets, G.L.; Hendriks, B.H.W.; Sterenberg, H.J.C.M.; Kuhlmann, K.F.D.; Ruers, T.J.M. Using Diffuse Reflectance Spectroscopy to Distinguish Tumor Tissue from Fibrosis in Rectal Cancer Patients as a Guide to Surgery. *Lasers Surg. Med.* **2019**, *52*, 604–611. [[CrossRef](#)] [[PubMed](#)]
98. Baltussen, E.J.M.; De Koning, S.G.B.; Sanders, J.; Aalbers, A.G.J.; Kok, N.F.M.; Beets, G.L.; Hendriks, B.H.W.; Sterenberg, H.J.C.M.; Kuhlmann, K.F.D.; Ruers, T.J.M. Tissue diagnosis during colorectal cancer surgery using optical sensing: An in vivo study. *J. Transl. Med.* **2019**, *17*, 333. [[CrossRef](#)] [[PubMed](#)]
99. Baltussen, E.J.M.; Sterenberg, H.J.C.M.; Ruers, T.J.M.; Dashtbozorg, B. Optimizing algorithm development for tissue classification in colorectal cancer based on diffuse reflectance spectra. *Biomed. Opt. Express* **2019**, *10*, 6096–6113. [[CrossRef](#)] [[PubMed](#)]
100. Langhout, G.C.; Spliethoff, J.W.; Schmitz, S.J.; Aalbers, A.G.J.; van Velthuysen, M.-L.; Hendriks, B.H.W.; Ruers, T.J.M.; Kuhlmann, K.F.D. Differentiation of healthy and malignant tissue in colon cancer patients using optical spectroscopy: A tool for image-guided surgery. *Lasers Surg. Med.* **2015**, *47*, 559–565. [[CrossRef](#)] [[PubMed](#)]
101. Langhout, G.C.; Spliethoff, J.W.; Aalbers, A.G.J.; Verwaal, V.J.; Hendriks, B.H.W.; Ruers, T.J.M.; Kuhlmann, K.F.D. Colorectal Cancer Identified Using Optical Spectroscopy. *Ann. Oncol.* **2014**, *25*, iv206. [[CrossRef](#)]
102. Kumashiro, R.; Konishi, K.; Chiba, T.; Akahoshi, T.; Nakamura, S.; Murata, M.; Tomikawa, M.; Matsumoto, T.; Maehara, Y.; Hashizume, M. Integrated endoscopic system based on optical imaging and hyperspectral data analysis for colorectal cancer detection. *Anticancer Res.* **2016**, *36*, 3925–3932.
103. Han, Z.; Zhang, A.; Wang, X.; Sun, Z.; Wang, M.D.; Xie, T. In vivo use of hyperspectral imaging to develop a noncontact endoscopic diagnosis support system for malignant colorectal tumors. *J. Biomed. Opt.* **2016**, *21*, 16001. [[CrossRef](#)]
104. Yuan, X.; Zhang, D.; Wang, C.; Dai, B.; Zhao, M.; Li, B. Hyperspectral Imaging and SPA-LDA Quantitative Analysis for Detection of Colon Cancer Tissue. *J. Appl. Spectrosc.* **2018**, *85*, 307–312. [[CrossRef](#)]
105. Ge, Z.; Schomacker, K.T.; Nishioka, N.S. Identification of colonic dysplasia and neoplasia by diffuse reflectance spectroscopy and pattern recognition techniques. *Appl. Spectrosc.* **1998**, *52*, 833–839. [[CrossRef](#)]
106. Rodriguez-Diaz, E.; Huang, Q.; Cerda, S.R.; O'Brien, M.J.; Bigio, I.J.; Singh, S.K. Endoscopic histological assessment of colonic polyps by using elastic scattering spectroscopy. *Gastrointest. Endosc.* **2015**, *81*, 539–547. [[CrossRef](#)] [[PubMed](#)]
107. Chen, H.; Lin, Z.; Wu, H.; Wang, L.; Wu, T.; Tan, C. Diagnosis of colorectal cancer by near-infrared optical fiber spectroscopy and random forest. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2015**, *135*, 185–191. [[CrossRef](#)]
108. Chen, H.; Lin, Z.; Mo, L.; Tan, C. Identification of colorectal cancer using near-infrared spectroscopy and adaboost with decision stump. *Anal. Lett.* **2017**, *50*, 2608–2618. [[CrossRef](#)]
109. Chen, H.; Tan, C.; Wu, H.; Lin, Z.; Wu, T. Feasibility of rapid diagnosis of colorectal cancer by near-infrared spectroscopy and support vector machine. *Anal. Lett.* **2014**, *47*, 2580–2593. [[CrossRef](#)]
110. Ehlen, L.; Zabarylo, U.J.; Speichinger, F.; Bogomolov, A.; Belikova, V.; Bibikova, O.; Artyushenko, V.; Minet, O.; Beyer, K.; Kreis, M.E.; et al. Synergy of Fluorescence and Near-Infrared Spectroscopy in Detection of Colorectal Cancer. *J. Surg. Res.* **2019**, *242*, 349–356. [[CrossRef](#)]
111. Claridge, E.; Hidović-Rowe, D. Model based inversion for deriving maps of histological parameters characteristic of cancer from ex-vivo multispectral images of the colon. *IEEE Trans. Med. Imaging* **2013**, *33*, 822–835. [[CrossRef](#)]

112. Lee, L.C.; Liong, C.-Y.; Jemain, A.A. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: A review of contemporary practice strategies and knowledge gaps. *Analyst* **2018**, *143*, 3526–3539. [\[CrossRef\]](#) [\[PubMed\]](#)
113. Brereton, R.G.; Lloyd, G.R. Partial least squares discriminant analysis: Taking the magic away. *J. Chemom.* **2014**, *28*, 213–225. [\[CrossRef\]](#)
114. Gromski, P.S.; Muhamadali, H.; Ellis, D.I.; Xu, Y.; Correa, E.; Turner, M.L.; Goodacre, R. A tutorial review: Metabolomics and partial least squares-discriminant analysis—a marriage of convenience or a shotgun wedding. *Anal. Chim. Acta* **2015**, *879*, 10–23. [\[CrossRef\]](#)
115. ElMasry, G.; Nakauchi, S. Prediction of meat spectral patterns based on optical properties and concentrations of the major constituents. *Food Sci. Nutr.* **2016**, *4*, 269–283. [\[CrossRef\]](#)
116. Maruo, K.; Yamada, Y. Near-infrared noninvasive blood glucose prediction without using multivariate analyses: Introduction of imaginary spectra due to scattering change in the skin. *J. Biomed. Opt.* **2015**, *20*, 47003. [\[CrossRef\]](#) [\[PubMed\]](#)
117. Tromberg, B.J.; Shah, N.; Lanning, R.; Cerussi, A.; Espinoza, J.; Pham, T.; Svaasand, L.; Butler, J. Non-invasive in vivo characterization of breast tumors using photon migration spectroscopy. *Neoplasia* **2000**, *2*, 26–40. [\[CrossRef\]](#) [\[PubMed\]](#)
118. Wang, C.-Y.; Chen, C.-T.; Chiang, C.-P.; Young, S.-T.; Chow, S.-N.; Chiang, H.K. Partial least-squares discriminant analysis on autofluorescence spectra of oral carcinogenesis. *Appl. Spectrosc.* **1998**, *52*, 1190–1196. [\[CrossRef\]](#)
119. Nogueira, M.S.; Cosci, A.; Kurachi, C. Assessment of oxidative stress and metabolic rates in liver grafts using time-resolved fluorescence spectroscopy. In Proceedings of the Biophotonics: Photonic Solutions for Better Health Care VI, Strasbourg, France, 22–26 April 2018; Volume 10685. [\[CrossRef\]](#)
120. Nogueira, M.S.; Rosa, R.G.T.; Pratavieira, S.; D'almeida, C.D.P.; Kurachi, C. Assembly and characterization of a fluorescence lifetime spectroscopy system for skin lesions diagnostic. *Biophotonics S. Am.* **2015**, 9531, 95313D. [\[CrossRef\]](#)
121. Almeida, C.D.P.D.; Campos, C.; Nogueira, M.S.; Kurachi, C. Time-resolved and steady-state fluorescence spectroscopy for the assessment of skin photoaging process. In Proceedings of the Biophotonics South America, Rio de Janeiro, Brazil, 23–25 May 2015; Volume 9531. [\[CrossRef\]](#)
122. Baker, M.J.; Byrne, H.J.; Chalmers, J.; Gardner, P.; Goodacre, R.; Henderson, A.; Kazarian, S.G.; Martin, F.L.; Moger, J.; Stone, N.; et al. Clinical applications of infrared and Raman spectroscopy: State of play and future challenges. *Analyst* **2018**, *143*, 1735–1757. [\[CrossRef\]](#) [\[PubMed\]](#)
123. Coelho, M.C.D.M.S.; Leal, L.B.; Nogueira, M.S.; Castro, P.A.; Peralta, F.; Zezell, D.M. Biochemical characterization of saliva of smoking and non-smoking patients by using Fourier-transform infrared spectroscopy. In Proceedings of the Biomedical Vibrational Spectroscopy 2022: Advances in Research and Industry, San Francisco, CA, USA, 22 January–28 February 2022; Volume 11957, p. 119570E. [\[CrossRef\]](#)
124. Pakiet, A.; Kobiela, J.; Stepnowski, P.; Sledzinski, T.; Mika, A. Changes in lipids composition and metabolism in colorectal cancer: A review. *Lipids Health Dis.* **2019**, *18*, 29. [\[CrossRef\]](#)
125. Bravo, J.J.; Paulsen, K.D.; Roberts, D.W.; Kanick, S.C. Sub-diffuse optical biomarkers characterize localized microstructure and function of cortex and malignant tumor. *Opt. Lett.* **2016**, *41*, 781–784. [\[CrossRef\]](#)
126. Kirkpatrick, N.D.; Brewer, M.A.; Utzinger, U. Endogenous optical biomarkers of ovarian cancer evaluated with multiphoton microscopy. *Cancer Epidemiol. Prev. Biomark.* **2007**, *16*, 2048–2057. [\[CrossRef\]](#)
127. Petrik, V.; Loosemore, A.; Howe, F.A.; Bell, B.A.; Papadopoulos, M.C. OMICS and brain tumour biomarkers. *Br. J. Neurosurg.* **2006**, *20*, 275–280. [\[CrossRef\]](#)
128. Georgakoudi, I.; Quinn, K.P. Optical imaging using endogenous contrast to assess metabolic state. *Annu. Rev. Biomed. Eng.* **2012**, *14*, 351–367. [\[CrossRef\]](#) [\[PubMed\]](#)
129. Slaby, O. Non-coding RNAs as biomarkers for colorectal cancer screening and early detection. In *Non-Coding RNAs in Colorectal Cancer*; Springer: Boston, MA USA, 2016; pp. 153–170.
130. Custodio, A.; Barriuso, J.; De Castro, J.; Martínez-Marín, V.; Moreno, V.; Rodríguez-Salas, N.; Feliu, J. Molecular markers to predict outcome to antiangiogenic therapies in colorectal cancer: Current evidence and future perspectives. *Cancer Treat. Rev.* **2013**, *39*, 908–924. [\[CrossRef\]](#) [\[PubMed\]](#)
131. Song, H.; Wang, L.; Liu, H.-L.; Wu, X.-B.; Wang, H.-S.; Liu, Z.-H.; Li, Y.; Diao, D.-C.; Chen, H.-L.; Peng, J.-S. Tissue metabolomic fingerprinting reveals metabolic disorders associated with human gastric cancer morbidity. *Oncol. Rep.* **2011**, *26*, 431–438. [\[PubMed\]](#)
132. Gadducci, A.; Guerrieri, M.E.; Greco, C. Tissue biomarkers as prognostic variables of cervical cancer. *Crit. Rev. Oncol. Hematol.* **2013**, *86*, 104–129. [\[CrossRef\]](#) [\[PubMed\]](#)
133. Beger, R.D. A review of applications of metabolomics in cancer. *Metabolites* **2013**, *3*, 552–574. [\[CrossRef\]](#) [\[PubMed\]](#)
134. Langhout, G.C.; Bydlon, T.M.; van der Voort, M.; Müller, M.; Kortsmits, J.; Lucassen, G.; Balthasar, A.J.R.; van Geffen, G.-J.; Steinfeldt, T.; Sterenborg, H.J.C.M.; et al. Nerve detection using optical spectroscopy, an evaluation in four different models: In human and swine, in-vivo, and post mortem. *Lasers Surg. Med.* **2018**, *50*, 253–261. [\[CrossRef\]](#)
135. Salomatina, E.; Yaroslavsky, A.N. Evaluation of the in vivo and ex vivo optical properties in a mouse ear model. *Phys. Med. Biol.* **2008**, *53*, 2797. [\[CrossRef\]](#)