Annals of Surgical Oncology, **11**(1):65–70
DOI: 10.1245/ASO.2004.03.031

Autofluorescence and Diffuse Reflectance Properties of Malignant and Benign Breast Tissues

Tara M. Breslin, MD, Fushen Xu, MD, Gregory M. Palmer, MS, Changfang Zhu, MS, Kennedy W. Gilchrist, MD, and Nirmala Ramanujam, PhD

Background: Fluorescence spectroscopy is an evolving technology that can rapidly differentiate between benign and malignant tissues. These differences are thought to be due to endogenous fluorophores, including nicotinamide adenine dinucleotide, flavin adenine dinucleotide, and tryptophan, and absorbers such as β-carotene and hemoglobin. We hypothesized that a statistically significant difference would be demonstrated between benign and malignant breast tissues on the basis of their unique fluorescence and reflectance properties.

Methods: Optical measurements were performed on 56 samples of tumor or benign breast tissue. Autofluorescence spectra were measured at excitation wavelengths ranging from 300 to 460 nm, and diffuse reflectance was measured between 300 and 600 nm. Principal component analysis to dimensionally reduce the spectral data and a Wilcoxon ranked sum test were used to determine which wavelengths showed statistically significant differences. A support vector machine algorithm compared classification results with the histological diagnosis (gold standard).

Results: Several excitation wavelengths and diffuse reflectance spectra showed significant differences between tumor and benign tissues. By using the support vector machine algorithm to incorporate relevant spectral differences, a sensitivity of 70.0% and specificity of 91.7% were achieved.

Conclusions: A statistically significant difference was demonstrated in the diffuse reflectance and fluorescence emission spectra of benign and malignant breast tissue. These differences could be exploited in the development of adjuncts to diagnostic and surgical procedures.

Key Words: Detection—Breast cancer—Fluorescence spectroscopy—Imaging.

Fluorescence spectroscopy is an optical method that can provide rapid differentiation between tumor and normal tissue in a variety of epithelial organ systems. When tissue is illuminated with specific wavelengths of ultraviolet (UV) or visible (VIS) light (excitation), fluorescent biological molecules (fluorophores) will absorb the energy and emit it as fluorescent light at longer wavelengths (emission). Furthermore, there are nonfluorescent light absorbers (such as hemoglobin) and scatterers (cells and subcellular organelles) in tissues that modulate the tissue fluorescence intensity at the excitation and emission wavelengths. Diffuse reflectance spectroscopy is another optical method that provides direct measurement of light absorption and scattering. The fluorophores can be either endogenous to the tissue or exogenous in the form of an injectable fluorescent molecule. The advantage of using exogenous fluorophores is that the photophysical and pharmacokinetic properties can be selected and are known. Furthermore, the exogenous fluorophores are more highly fluorescent than endogenous fluorophores. However, the disadvantage of using exogenous fluorophores is that issues relating to potential drug toxicity and timing of administration have to be addressed. Endogenous fluorophores in tissue include amino acids, structural proteins, enzymes and coenzymes, vitamins, lipids, and porphyrins. Each of these

Received March 7, 2003; accepted August 25, 2003.

From the Departments of Surgery (TMB), Anatomic Pathology (FX, KWG), Biomedical Engineering (GMP, NR), and Electrical and Computer Engineering (CZ), University of Wisconsin-Madison, Madison, Wisconsin.

Address correspondence and reprint requests to: Tara M. Breslin, MD, Department of Surgery, University of Wisconsin, 600 Highland Ave., H4/744 CSC, Madison, WI 53792-7375; Fax: 608-263-7652; E-mail: breslin@surgery.wisc.edu.

molecules has unique excitation and emission spectra in the UV/VIS spectral region.[1–3]

Fluorescence spectroscopy has been successfully used to provide fast and minimally invasive detection of cancers and precancers in a variety of organ systems in vivo, including the cervix, colon, bronchus, bladder, and oral mucosa.[3–8] These studies involved modification of existing endoscopic equipment to include fluorescence spectroscopy capability. Several groups have evaluated fluorescence spectroscopy for breast cancer detection in ex vivo studies. Alfano et al.[9] were the first to measure spectra of normal and malignant breast tissues from two patients at 488 and 457.9 nm excitation. Subsequently, Yang et al.[10–12] reported 93% sensitivity and 95% specificity rates for discrimination between malignant and benign tissue by using 300 nm excitation spectra.

Diffuse reflectance spectroscopy for breast cancer detection has also been studied. Using wavelengths between 330 and 750 nm, Bigio et al.[13] measured diffuse reflectance spectra through a core biopsy needle and during breast cancer surgery and showed that this technique can differentiate tumor from normal tissue with a sensitivity of 60% to 70% and a specificity of 85% to 95%.

Despite the promising results of previous work, significant gaps in knowledge remain. The main limitations of the previous studies are that fluorescence spectra were obtained only at one excitation wavelength or several excitation wavelengths, and the utility of combining fluorescence and diffuse reflectance spectroscopy has not been evaluated. In addition, the challenge exists of adapting this technology, which is ideally suited to evaluating epithelial surfaces, to a three-dimensional organ system. The primary goal of this study was to characterize the multiexcitation fluorescence spectra (at nine excitation wavelengths in the UV/VIS range) and UV/VIS diffuse reflectance spectra of benign and malignant breast tissue and to identify the optimal spectral features for breast cancer diagnosis. We hypothesized that this technique would be able to demonstrate a statistically significant difference between benign and malignant breast tissue on the basis of their unique fluorescence and reflectance properties. In a previous report,[14] we described the development of the optical system and a novel nonparametric algorithm for statistical analysis and sample classification. This report focuses on the clinical application of this exciting technology to breast cancer diagnosis and surgery.

PATIENTS AND METHODS

The Institutional Review Board at the University of Wisconsin approved the conduct of this study, and all subjects gave written, informed consent before study participation. Eligible subjects included women with a diagnosis of invasive breast cancer who were scheduled for definitive surgery and women undergoing reduction mammoplasty. After gross evaluation by the pathologist (F.X. or K.W.G.), samples of tumor and adjacent normal-appearing tissue measuring approximately 1 cm^3 were harvested from the mastectomy or lumpectomy specimens, and normal-appearing breast tissue was harvested from the reduction mammoplasty specimens. Optical spectroscopic measurements were performed on the freshly excised samples within 2 hours after surgical excision. Normal tissue was measured $\geq$1 cm away from the grossly visible tumor margin to minimize the potential for measuring adjacent ductal carcinoma-in-situ.

A Skinscan spectrofluorometer (J. Y. Horiba, Edison, NJ) was used for all measurements. This instrument consists of a fiberoptic probe with a central collection region and an outer ring of excitation fibers. Fluorescence emission spectra were recorded in 20-nm increments at excitation wavelengths of 300 to 460 nm. At each excitation wavelength, fluorescence emission was recorded in 5-nm increments, beginning at a wavelength 10 nm longer than the excitation wavelength, up to 600 nm (e.g., 310 to 600 nm for a 300-nm excitation). Fluorescence emission spectra were thus obtained at nine excitation wavelengths. Diffuse reflectance was measured by illuminating and collecting at the same wavelength ranging from 300 to 600 nm in 5-nm increments.

After each measurement, the probe position on each tissue sample was inked (TMD-BK; Triangle Biomedical Sciences, Durham, NC), and the specimen was formalin-fixed and processed for routine histopathology. Microscopic evaluation of each histological section was performed (F.X. and K.W.G.) and a consensus diagnosis was reached. When a sample exhibited a heterogeneous diagnosis at the site of measurement, the worst-case diagnosis was used (e.g., for samples that contained both normal glandular tissue and malignant tissue, the diagnosis was coded as malignant). In cases in which normal adipose and fibrous/glandular tissues were present, the histology was determined by the predominant tissue type at the measurement site.

Data analysis was performed with the Matlab (Mathworks Inc., Natick, MA) software package. Principal component analysis (PCA) was used as a data-reduction technique. PCA characterizes a majority of the variance while greatly reducing the input data set into a few orthogonal variables. The principal components (PCs) are extracted such that the first PC (PC1) accounts for the largest amount of the total variance of the input data. The second PC (PC2) accounts for the second largest amount

of the variance while being orthogonal to PC1, and so on. There are two advantages to this transformation: (1) the input data can be represented by a few subsets of PCs with minimal mean square error, which reduces the dimensionality of the data set; and (2) the projection onto the PC subspace maximizes the separation of data clusters.[15,16] The primary drawback of this technique is that it is not trivial to determine which parts of the data input (i.e., spectra) are diagnostically useful. To balance these concerns, the PCA was performed individually on all fluorescence emission spectra, one excitation wavelength at a time. A similar process was performed on the diffuse reflectance data, thus yielding a set of PCs to characterize each spectrum.

To determine which of the PCs represented variance due to malignancy (rather than normal variability), a Wilcoxon ranked sum test ($P < .0005$) was used to test which PCs demonstrated significant differences between malignant and nonmalignant samples. Because PCA was performed on each excitation wavelength separately, the wavelength responsible for each PC could be determined, and the most diagnostically useful wavelengths were deduced.

A support vector machine (SVM) algorithm was used for classification.[17,18] SVM is a classification algorithm based on statistical learning theory. The principal idea of an SVM is to determine an optimal separating hyperplane that maximizes the margin between two classes in a multidimensional data space. With the largest separation of the two data clusters, the SVM classifier gives a lower expected risk, which means that future error can be minimized if more data are added to the sample pool. The final step in data analysis involved comparing the sensitivity and specificity of those PCs found to be statistically significant with the gold standard histopathological diagnosis. For comparison with the classification scheme, the samples were labeled as either tumor or benign (fibrous/glandular and adipose). The rationale for grouping adipose tissue with benign (fibrous/glandular) tissue despite the observation that they are histologically distinct, with different optical properties, was based on our previous work. In developing the statistical classification scheme, we found that the algorithm is not yet able to discriminate between adipose and nonadipose tissue with 100% accuracy. As a result, the overall classification was more accurate with a one-step process as opposed to a two-step process.[14]

A full cross-validation was performed on the classification scheme by sequential removal of individual samples from the training datasets used in the PCA, Wilcoxon, and SVM algorithms. The accuracy of the resulting classification scheme was then tested on the deleted sample. This method allows for unbiased evaluation of the diagnostic algorithm when the sample size is too small for division into separate training and testing datasets. This procedure was performed for all 56 samples; each case had 55 training samples and 1 testing sample.

RESULTS

Forty subjects were enrolled onto the study, and 56 tissue samples from 32 patients were used for this analysis. A prototype device was used for the first six cases, and data from these samples was not included in this analysis because of potential measurement differences between the two instruments. Device failure occurred in 1 case, and histopathologic correlation was not available in another case, leaving a total of 32 subjects. Paired tumor and normal specimens were obtained from 11 participants. Because of limitations imposed by small surgical specimens, small tumors, or a noncancer diagnosis in the remaining 21 subjects, only either a normal or tumor sample was available for the measurement.

The mean age of all participants was 48.4 years and was 51.5 years for those with cancer. Twenty-seven subjects had a diagnosis of infiltrating carcinoma, and five underwent reduction mammoplasty for benign conditions. For each sample, the histopathologic diagnosis was classified as tumor, fibrous/glandular, or adipose. Table 1 summarizes the histopathologic diagnosis of the tissue samples used for spectroscopic analysis.

Fluorescence excitation-emission matrices (EEMs) were measured for tumor, fibrous/glandular, and adipose tissues. These plots allow for the display of fluorescence intensities acquired at multiple excitation wavelengths, where the florescence spectrum at each excitation wavelength is represented by a horizontal section through the plot. The EEMs for the tumor and benign/fibrous sample

TABLE 1. *Histopathology of samples used for fluorescence spectroscopy*

Histology	No. Samples
Tumor	20
Infiltrating ductal carcinoma	16
Infiltrating lobular carcinoma	2
Ductal carcinoma-in-situ	1
Infiltrating ductal and ductal carcinoma-in-situ	1
Fibrous/glandular	15
Normal fibrous tissue	8
Adenosis	2
Reparative change	2
Fibrocystic change	1
Fibroadenoma	1
Cystic	1
Adipose	21

sets are demonstrated in Fig. 1A and B. Four dominant peaks occurring at excitation-emission wavelength pairs of 300 to 340 nm, 340 to 390 nm, 360 to 460 nm, and 440 to 520 nm are visible. Adipose tissue (Fig. 1C) shows distinct differences relative to tumor and fibrous/glandular tissue. Most notably, the peak at 340 to 390 nm is absent, and the peak at 360 to 460 nm is shifted to 360 to 520 nm. There are valleys in both EEMs at 360 to 420 nm, 420 to 480 nm, and 420 to 580 nm; these are likely due to the Soret absorption of hemoglobin. It should be noted that spectral line shape differences among normal, benign, and malignant tissues can be observed only by examining individual spectra. The multivariate statistical algorithm indicates that the line shape of fluorescence spectra at several excitation wavelengths shows the most significant differences among the three tissue types.

Figure 2 depicts the average diffuse reflectance spectra with SDs for the three tissue classifications. The average tumor and fibrous/glandular spectra have similar intensities, but there are subtle differences in spectral line shape at 300 to 350 nm, 400 to 450 nm, and 525 to 575 nm. Adipose tissue shows, on average, a decreased intensity at all wavelengths and a different line shape between 425 and 525 nm when compared with tumor and benign (fibrous/glandular) tissues.

After the PCA of the fluorescence and diffuse reflectance measurements was completed, nine PCs were identified that demonstrated a statistically significant difference ($P < .05$) between tumor and benign (fibrous/glandular or adipose) tissue. To limit the number of measurement parameters (and, therefore, the time required for measurement in a clinical setting), the significance cutoff was decreased to $P < .0005$ and the remaining PCs were used for the classification with the SVM algorithm. Because the most significant PCs were all derived from the fluorescence spectra, only the fluorescence spectra were included in the analysis at this significance level.

After determining which PCs demonstrated statistically significant differences between tumor and benign tissue, the spectra were analyzed by using the SVM algorithm to classify samples as tumor or benign. These results were compared with the histopathologic diagnosis (gold standard) to determine the sensitivity and specificity of the technique. Table 2 depicts the diagnostic data in 2×2 format and demonstrates a sensitivity of 70.0% and specificity of 91.7% for fluorescence spectroscopy alone. Six tumor samples and three benign samples were misclassified, yielding a positive predictive value of 82.3% and a negative predictive value of 84.6%.

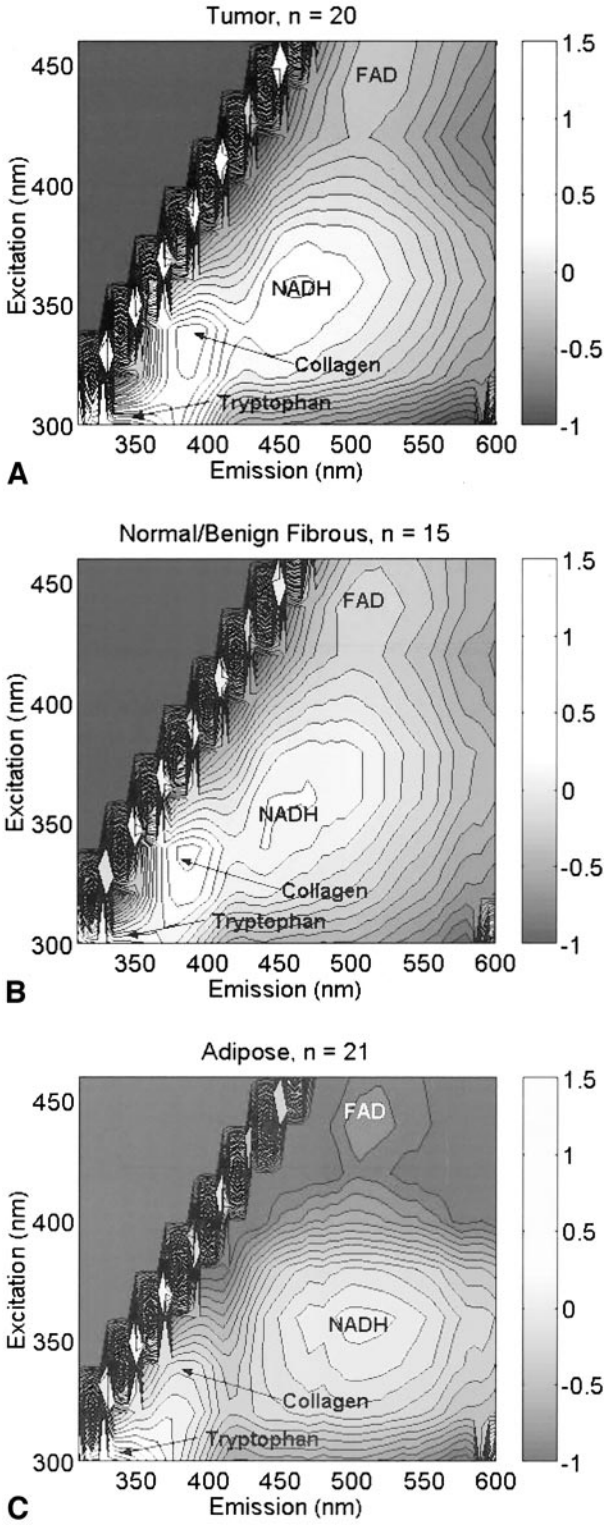

FIG. 1. (**A**) Excitation-emission matrix for tumor tissue. (**B**) Excitation-emission matrix for benign (fibrous/glandular) tissue. (**C**) Excitation-emission matrix for adipose tissue. FAD, flavin adenine dinucleotide; NADH, nicotinamide adenine dinucleotide.

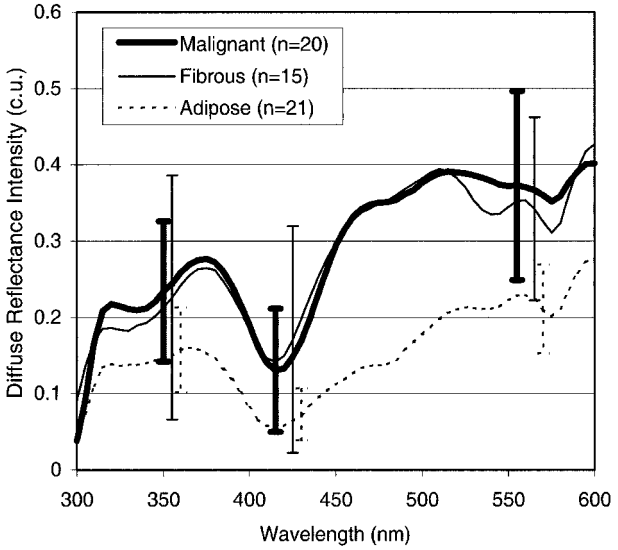

FIG. 2. Average diffuse reflectance spectra for tumor, benign (fibrous/glandular), and adipose tissue. c.u., calibrated units.

TABLE 2. *Comparison of classification accuracy by using autofluorescence and diffuse reflectance spectroscopy with histopathologic diagnosis*[a]

	Gold standard	
Test result	Histology = tumor	Histology = benign
Tumor	14	3
Benign	6	33

[a] Sensitivity is the proportion of those with disease according to the gold standard who are labeled positive by the test
$14/(14 + 6) = .700$.
Specificity is the proportion of those who are disease free according to the gold standard who are labeled negative by the test
$33/(33 + 3) = .917$.

DISCUSSION

Optical spectroscopy using fluorescence spectroscopy in the UV/VIS spectral range demonstrates statistically significant differences between tumor and nontumor human breast tissue in an ex vivo system. By measuring multiple excitation wavelength fluorescence spectra and diffuse reflectance spectra and analyzing each separately, those spectra that yielded the most diagnostically useful information were identified. Of the 10 measured spectra, only 4 are required to maximize classification accuracy. These include fluorescence spectra at excitation wavelengths of 300, 400, 420, and 460 nm. Minimizing the number of wavelengths analyzed is advantageous clinically because it lends speed to the process and should require a less complex, more economical instrument.

Diffuse reflectance exploits some of the same optical features as fluorescence spectroscopy. The most notable similarities are biological chromophores that result in absorption, such as oxygenated and deoxygenated hemoglobin and β-carotene, and tissue scattering properties. β-Carotene is abundant in adipose tissue and may provide a unique indicator for this tissue type. The primary advantage to using diffuse reflectance is that it can be performed at a fraction of the cost of fluorescence spectroscopy. However, with a similar classification technique, diffuse reflectance spectroscopy alone was not capable of discriminating between tumor and nontumor tissue with the same accuracy as the combined technique.[14] Therefore, there is a potential advantage to using fluorescence spectroscopy instead of diffuse reflectance. This advantage is likely due to additional chemical specificity arising from characterization of the many intrinsic fluorophores present in human tissue. Future studies will focus on performing measurements with only those fluorescence and reflectance spectra found to be statistically significant by PCA.

A nontrivial issue with using endogenous fluorophores as used in this study is that is difficult to determine the true biologic basis of the spectroscopic differences observed between tumor and nontumor tissue. There are, however, known biologic fluorophores that contribute to the fluorescence spectra at the optimum excitation wavelengths as determined by this study. The excitation wavelengths used, ranging from 300 to 460 nm, allow for characterization of a number of biologic fluorophores, including tryptophan, nicotinamide adenine dinucleotide, flavoproteins, and collagen, all of which are present in tissue systems.[19] The fluorescence excitation-emission wavelengths identified by the PCA/SVM algorithm suggest that tryptophan, nicotinamide adenine dinucleotide, and flavoproteins are important for breast cancer diagnosis. Hemoglobin is another important absorber that may contribute to differences between tumor and nontumor tissue.

The statistical algorithm used for this study was developed for analysis of this dataset, taking into account both the small sample size and the large number of individual data points. Previously published studies using cervical biopsies had a much larger sample size, which allowed for splitting the data into training and testing subsets. In one study, Ramanujam et al.[20] divided the samples into training and testing datasets and used PCA followed by logistic discrimination to classify tissue types and calculate the posterior probability of a correct diagnosis. This technique allows for a more robust evaluation of the classification algorithm. As our dataset enlarges, we plan to develop independent training and testing datasets and explore alternate classification schemes by using neural networks.

These exciting preliminary data derived from measurements performed on ex vivo breast tissues form an essential background for performing future in vivo experiments. The spectral differences demonstrated to exist between tumor and nontumor tissue could be exploited in the development of diagnostic adjuncts to core biopsy, fiberoptic ductoscopy, or evaluation of surgical margins. Future challenges include issues such as the development of a sterilizable probe appropriate for intraparenchymal or intraductal use, determining the optimum measurement depth, and elucidating the biologic mechanisms behind these observed optical differences.

ACKNOWLEDGMENTS

The acknowledgments are available online in the full-text version at www.annalssurgicaloncology.org. They are not available in the PDF version.

REFERENCES

1. Ramanujam N. Fluorescence spectroscopy of neoplastic and non-neoplastic tissues. *Neoplasia* 2000;2:89–117.
2. Richards-Kortum R, Sevick-Muraca E. Quantitative optical spectroscopy for tissue diagnosis. *Annu Rev Phys Chem* 1996;47:555–606.
3. Wagnieres GA, Star WM, Wilson BC. In vivo fluorescence spectroscopy and imaging for oncological applications. *Photochem Photobiol* 1998;68:603–32.
4. Mitchell MF, Cantor SB, Ramanujam N, Tortolero-Luna G, Richards-Kortum R. Fluorescence spectroscopy for diagnosis of squamous intraepithelial lesions of the cervix. *Obstet Gynecol* 1999; 93:462–70.
5. Mycek MA, Schomaker KT, Nishioka NS. Colonic polyp differentiation using time-resolved autofluorescence spectroscopy. *Gastrointest Endosc* 1998;48:390–4.
6. Zheng W, Lau W, Cheng C, Soo KC, Olivo M. Optimal excitation-emission wavelengths for auto fluorescence diagnosis of bladder tumors. *Int J Cancer* 2003;104:477–81.
7. Hirsch FR, Prindiville SA, Miller YE, et al. Fluorescence versus white-light bronchoscopy for detection of preneoplastic lesions: a randomized study. *J Natl Cancer Inst* 2001;93:1385–91.
8. Gillenwater A, Jacob R, Ganeshappa R, et al. Noninvasive diagnosis of oral neoplasia based on fluorescence spectroscopy and native tissue autofluorescence. *Arch Otolaryngol Head Neck Surg* 1998;124:1251–8.
9. Alfano RR, Pradhan A, Tang GC, Wahl SJ. Optical spectroscopic diagnosis of cancer and normal breast tissues. *J Opt Soc Am B* 1989;6:1015–23.
10. Yang Y, Katz A, Celmer EJ, Zurawska-Szczepaniak M, Alfano RR. Optical spectroscopy of benign and malignant breast tissues. *Lasers Life Sci* 1996;7:115–27.
11. Yang Y, Katz A, Celmer EJ, Zurawska-Szczepaniak M, Alfano RR. Fundamental differences of excitation spectrum between malignant and benign breast tissues. *Photochem Photobiol* 1997;66: 518–22.
12. Yang Y, Celmer EJ, Koutcher JA, Alfano RR. UV reflectance spectroscopy probes DNA and protein changes in human breast tissue. *J Clin Laser Med Surg* 2001;19:35–9.
13. Bigio IJ, Bown SG, Briggs G, et al. Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results. *J Biomed Opt* 2000;5:221–8.
14. Palmer GM, Zhu C, Breslin TM, Xu F, Gilchrist KW, Ramanujam N. Comparison of multiexcitation fluorescence and diffuse reflectance spectroscopy for the diagnosis of breast cancer. *IEEE Trans Biomed Eng* 2003;50:1233–42.
15. Etemad K, Chellappa R. Separability-based multiscale basis selection and feature extraction for signal and image classification. *IEEE Trans Image Processing* 1998;7:1453–65.
16. Yue W, Lan L, Freedman MT, Sun Yuan K. Probabilistic principal component subspaces: a hierarchical finite mixture model for data visualization. *IEEE Trans Neural Networks* 2000;11:625–36.
17. Burges C. A tutorial on support vector machines for pattern recognition. *Data Mining Knowledge Discovery* 1998;2:121–67.
18. Gunn S. Support vector machines for classification and regression. Available at: http://www.ecs.soton.ac.uk/~srg/publications/pdf/SVM.pdf. Southampton, UK: University of Southampton, Department of Electronics and Computer Science: 1998.
19. Ramanujam N. Fluorescence spectroscopy in vivo. In: Meyers R, ed. *Encyclopedia of Analytical Chemistry*. New York: Wiley, 2000:20–56.
20. Ramanujam N, Mitchell MF, Mahadevan-Jansen A, et al. Cervical precancer detection using a multivariate statistical algorithm based on laser-induced fluorescence spectra at multiple excitation wavelengths. *Photochem Photobiol* 1996;64:720–35.