RICE UNIVERSITY

MULTISPECTRAL OPTICAL IMAGING FOR THE DETECTION AND DELINEATION OF ORAL NEOPLASIA

by

Darren Michael Roblyer

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Rebecca Richards-Kortum, Stanley C. Moore Professor
Bioengineering

Ann M. Gillenwater, Associate Professor
The University of Texas M.D. Anderson Cancer Center

Tomasz Tkaczyk, Assistant Professor
Bioengineering & Electrical and Computer Engineering

Richard Baraniuk, Victor E. Cameron Professor
Electrical and Computer Engineering

HOUSTON, TEXAS
APRIL, 2009

ABSTRACT

Multispectral Optical Imaging for the Detection and Delineation of Oral Neoplasia

by

Darren Michael Roblyer

Despite the accessibility of the oral cavity to inspection, patients with oral cancer most often present at a late stage, leading to high morbidity and mortality. Multispectral widefield optical imaging has emerged as a promising technology to aid clinicians in screening and resection of oral neoplasia, but current approaches rely on subjective interpretation. This work focuses on the design, construction, and clinical testing of a novel multispectral widefield optical imaging device for objective screening and delineation of oral neoplasia.

The Multispectral Digital Microscope (MDM) acquires *in vivo* images of oral tissue in autofluorescence, narrow band reflectance, and orthogonal polarized reflectance modes that the diagnostic value of each modality may be qualitatively and quantitatively evaluated alone and in combination.

Using *in vivo* imaging data collected from 56 patients and 11 normal volunteers, combined with computer aided diagnostics, a sensitivity of 100% and a specificity of 91.4% was achieved for discriminating oral dysplasia and cancer from normal tissue in an independent validation set. A single feature calculated from the autofluorescence images at 405 nm excitation was used to achieve this performance. Disease probability maps

were constructed using this feature to help identify areas with a high probability of abnormality. Autofluorescence imaging at 405 nm excitation also provided the greatest image contrast which was significantly higher than that using standard white-light illumination. Features extracted from other imaging types did not appear to aid in diagnosis.

Ex vivo image data from the MDM was combined with image data from a high-resolution microendoscope (HRME) in order to determine if a synergistic relationship existed between these devices. The ability to objectively diagnose oral lesions substantially increased when using both devices in combination compared to using either alone. This combination of devices provides a practical means of screening the entire mucosal surface for suspicious regions, using the MDM, and then using the HRME for confirmation of diagnosis.

This work has demonstrated that widefield autofluorescence imaging at 405 nm excitation can be highly effective for the objective discrimination of oral lesions.

DEDICATION

To my Father, Michael Guy Roblyer.

ACKNOWLEDGEMENTS

I would like to thank everyone who has helped with this work including: Cristina Kurachi, Vanda Stepanek, Jency George, Tao Le, Leslie Zachariah, Kostia Sokolov, Michelle Williams, Adel K. El-Naggar, Ann Gillenwater, Crystal Weber, Richard Schwarz, Kelsey Rosbach, Mark Pierce, John Wright, and Vivian Mack. Special thanks to Tim Muldoon, together be blazed new roads in science and technology.

I would also like to thank Michelle Follen and the 2006-2007 PO1 group, Yvette Mirabel, as well as the wonderful people at the University Hospital in Ibadan for an enlightening and unique experience in Nigeria.

I want to thank Rebecca for being a fantastic mentor and advisor. You are brilliant and have an incredible balance in your life. Thank you for making my career development a priority as well as allowing me to pursue my less professional interests and responsibilities down at Valhalla.

Thanks to all my great friends and my best friend and girlfriend, Erin. Thanks to my Mother, of course you deserve most of the credit! Thank to both my Mother and Ken for all of their support and encouragement. Finally, thanks to my Father, my biggest advocate. I miss you.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1: INTRODUCTION

1.1 Overview

One in three Americans will be diagnosed with cancer within their lifetimes [1]. Although there have been significant improvements in treatment during the last 50 years, there has only recently been a slight decline in cancer mortality in the United States [2]. Early detection, however, can greatly increase a patient's chance of survival and often responds to simpler, less expensive, and less invasive interventions. When implemented correctly, early detection programs have the added benefit of reducing the burden on our national healthcare system. Early detection methods which are robust, relatively inexpensive, and require minimal expertise and training have the added potential to help combat rising mortality due to cancer in developing countries.

Oral cancer contributes significantly to the global cancer burden. In the US alone, there are an estimated 30,000 cases of oral cancer diagnosed and 7,320 deaths annually [3]. The incidence of disease increases dramatically in some developing nations and ranks as the sixth most common cancer globally [4]. The accessibility of the oral cavity makes is an ideal location for inspection and early detection, which can greatly improve survival. The five-year survival rate for oral cancer patients with advanced stages of disease is 17% but improves to 80% if diagnosed at local stages of development [5].

The most common detection methods for oral cancer in the US include visual inspection during routine dental visits and self referral by patients followed by invasive biopsy for confirmation. There are currently no widely implemented standards for oral cancer screening by dental practitioners and often decisions for further investigation of

suspicious areas are made based on experience. The situation is further complicated by the plethora of common benign conditions which can occur in the oral cavity. There are several screening aids commercially available but to date, all remain subjective in nature and confirmatory biopsies are still required.

An accurate and objective screening technique would provide great benefit to the global population and high risk populations in particular. Additionally, methods for improving the delineation of the extent or margin of malignant lesions in the oral cavity would aid in treatment methods such as surgical resection.

The vast majority of malignancies in the oral cavity begin in the epithelium at the junction between the epithelial and stromal tissue layers called the basement membrane. Growth of abnormal cells then progresses upward to the tissue surface before breaking through the basement membrane. The tissue depth of morphologic and biochemical changes in tissue (typically several hundred micrometers) which accompany malignant progression, coupled with the accessibility of the oral cavity, make oral cancer an ideal candidate for inspection with optical technologies. Visible light has a penetration depth in tissue which corresponds to the location of these changes. Recent advances in the understanding of the biological basis of the optical changes which occur in tissue during malignant progression, including autofluorescence signatures, help provide support to the idea that optical imaging, with appropriate illumination and collection conditions and wavelengths, can greatly improve the detection and delineation of oral premalignancies and malignancies [6].

1.2 Specific Aims

The overall goal of this dissertation was to elucidate the extent to which widefield multispectral optical imaging could improve diagnosis and delineation of oral neoplasia. The specific aims of the work were to:

Specific Aim 1: Design, construct, and characterize the performance of a widefield multimodal optical microscope capable of narrowband reflectance, polarized reflectance, and autofluorescence imaging. Use this system to study both endogenous elastic light scattering and autofluorescence features of precancerous and cancerous lesion in the oral cavity. Additionally, determine resolution, field of view, exposure times and other performance indicators of the MDM.

Specific Aim 2: Test the performance of the MDM for detecting pre-cancerous and cancerous lesions in the oral cavity using criteria such as sensitivity and specificity. Using data from an *in vivo* pilot clinical study of 72 consenting patients, extract quantitative image features and use computer aided diagnosis algorithms to determine performance. Intrinsic in this analysis are the determination of the imaging modalities and specific illumination and collection wavelengths which provide the greatest diagnostic utility. Additionally, determined to what extend the MDM can provide a specific graded diagnosis (i.e. grades of dysplasia). Finally, quantify image contrast to determine the modalities/wavelengths which provide the highest image contrast.

Specific Aim 3: Combine imaging results from the MDM with a high resolution microendoscope developed in our lab to reveal if this combination can provide increased diagnostic performance. Analyze data from resected tissue from 14 patients and utilize hybrid consensus techniques to determine diagnostic performance from the synthesis of two separate classification algorithms for the devices.

1.3 Chapter Summaries

This dissertation describes the development and usage of a novel, objective, oral cancer screening device based on multispectral optical imaging. The device was designed as a research grade scientific tool to explore the optical modalities and conditions which could provide improved information on the presence and location of oral lesions in vivo. Although previous studies have been conducted to investigate the ability of some specific imaging modalities to identify lesions, this is the first work to combine multiple modalities including autofluorescence, narrowband, and polarized imaging at multiple illumination and collection wavelengths and quantitatively compare results. Additionally, this work is the first to attempt quantification of the most common observations and implementation of computer aided diagnostics for the diagnosis of malignant lesions in the oral cavity with the aim of developing an objective screening and delineation tool. This work is extended by combining widefield multispectral imaging results with a high resolution optical imaging device, with the aim of further improving detection.

Chapter 2 provides background on the motivation for this work, and briefly reviews the biology of oral cancer. The biochemical and morphological changes which

occur during malignant progression are described as well as how these changes can be exploited by optical technologies for improved detection. Additionally, this chapter provides an overview of the current screening and detection methods available for oral cancer and limitations of these techniques. Finally, Chapter 2 provides a description of how computer aided diagnostics can be implemented with optical imaging.

Chapter 3 details the design criteria and construction of the device which will now be referred to as the Multispectral Digital Microscope (MDM), as well as initial qualitative observations from study subjects. This chapter describes the specific modalities of the MDM, including autofluorescence, narrow reflectance, and orthogonal polarized reflectance modes, which were used to acquired *in vivo* image data from oral tissue to allow evaluation of lesions which may not exhibit high contrast under standard white light illumination. The chapter continues with qualitative observations of the visual appearance of oral lesions using different imaging modalities. For example, changes in vascular patterns, contrast, and color are described on lesions and tissue from normal volunteers.

Chapter 4 provides an in depth analysis of autofluorescence imaging mode, which appeared to be highly useful based on results presented in Chapter 3. This chapter describes a quantitative method to extract image features and their use for objective, computer aided diagnosis of tissue regions. Data collected from an *in vivo* clinical study of 56 patients with oral lesions and 11 normal volunteers was utilized. From these images, 276 measurements from 159 unique regions of interest (ROI) sites corresponding to normal and confirmed neoplastic areas were identified. Data from ROIs in the first 46 subjects was used to develop a simple classification algorithm based on the ratio of red-

to-green fluorescence; performance of this algorithm was then validated using data from the ROIs in the last 21 subjects. This algorithm was applied to patient images to create visual disease-probability maps across the field of view. Histologic sections of resected tissue were used to validate the disease-probability maps.

Chapter 5 expands on the analysis conducted in Chapter 4 and includes the other imaging modalities and image types provided by the MDM in the computer aided diagnosis analysis. This chapter also quantifies the increase in optical image contrast seen using the different imaging modalities. Five new patients were added from Chapter 4 so that analysis was conducted on 61 patients with oral dysplasia and/or cancer and 11 normal volunteers. Image contrast of oral lesions compared to healthy tissue was computed and compared for all image types. Computer aided diagnosis was performed on the image data using a linear classifier and a decision tree classifier to explore the image types and extracted features which provide the best objective diagnosis.

Based on results from the previous chapters it was clear that the MDM could provide accurate identification of oral lesions in a tertiary care patient population. There are several questions and foreseeable limitations that this technology could encounter in a more diverse screening population. With this in mind, Chapter 6 explores combining the MDM analysis with results obtained from a high resolution microendoscope developed by our group. An underlying practical implication of combining these devices is that the MDM can provide rapid screening of the entire oral mucosal surface, or in this case, an entire resected specimen, and suspicious areas can be identified. The high resolution system can then interrogate these areas to help confirm a diagnosis.

This study involved a combined analysis including image data collected by the MDM and the high resolution microendocope system from *ex*-vivo tissue resected from 14 patients with oral precancers and cancer. Several methods of combining extracted image feature data from both devices were explored based on hybrid consensus techniques.

Chapter 7 provides a discussion of the implications of results from this dissertation work as well as potential future work. Chapter 8 provides a conclusion for the work.

CHAPTER 2: BACKGROUND

2.1 Motivation

Oral cancer poses a significant and largely preventable world health problem. Oral and pharyngeal cancer ranks as the sixth most common cancer globally [4]. In the US, there were an estimated 30,000 cases of oral cancer diagnosed and 7,320 deaths in 2005 [3]. The incidence is two times greater in men than woman and risk factors such as tobacco, including cigarettes, cigars, pipes, and smokeless tobacco, as well as alcohol, ultraviolet light, and poor nutrition greatly increase chances of developing oral cancers. Specific regional practices such as chewing of the betel nut in some parts of the world, and the areca quid chewing habit popular in Taiwan, can also increase incidence [7,8]. Oral cancer has a high mortality rate because it is often diagnosed during late stage disease, when symptoms normally occur. If not caught by routine dental or doctor's visit, oral cancers are often only diagnosed after a patient complains of pain or irritation resulting due to the cancer. In the US, the five-year survival rate for localized disease is 81% but for those with distant metastases is only 17% [5] showing that early detection has the ability to save a large number of lives.

2.2 Anatomy, Etiology, Histology and Progression of Disease

In the US, oral cancer occurs most frequently in the tongue (30% of cases) followed by lip (17%), floor of mouth (14%), gingiva, hard palate, soft palate, and buccal mucosa [4]. Over 90% of oral cancers are squamous cell carcinomas (SCCs) whose tissue of origin is the epithelial lining of the oral cavity or oropharynx. Differences in the epithelial lining of the regions of the oral cavity include a nonkeratinized mucosal lining

and loose lamina propria in the buccal, soft palate, floor of mouth and inner lip, and a superficial keratin layer on the dorsal tongue, hard palate, and gingiva. [9] The transformation from normal tissue to cancer may occur over a widely varying amount of time from several months to many years and includes abnormal cellular changes. Cellular transformations often begin at the basal layer and progress upward through the epithelium. This upward progression of cellular changes is called dysplasia. There is uncontrolled cell division and the typical progression from tall cells in the basal layer to flattened squamous cells on the surface may be replaced by disordered and abnormally shaped basal cells which are mitotic. There are several common cellular morphological and anatomical changes at this stage such as a variation in cell shape and size (pleomorphism), an increase of DNA content due to increased mitosis causing an increased nuclear to cytoplasmic ratio, and changes in chromatin texture which may cause nuclei to be deeply stained (hyperchromatism) [10-12]. Changes due to increased metabolic activity associated with malignant transformations occur and an increase of metabolic co-factors such as reduced forms of nicotinamide adenine dinucleotide (NADH), flavin adenine dinucleotide (FAD) [13]. Levels of aromatic amino acids such as tryptophan, tyrosine, and phenylalanine as well as porphyrins may also increase[14]. Stromal changes such as increase in collagen content or the breakdown of collagen crosslinks may occur during the transformation [10].

Since the progression of cancer is intimately linked to increased environmental stimuli and a cell's ability to replicate uncontrollably, certain biomarkers have been identified which are over expressed in cancers. In many cancers, including cancer of the oral cavity, it has been shown that epithelial growth factor receptor (EGFR), a tyrosine kinase

protein is expressed in proportion to the stage of malignant transformation [15,16]. Additionally, matrix-metalloproteases such as MMP-2 and MMP-9 have been shown through studies of immunohistochemical staining of oral cancer tissue sections to be overexpressed in a majority of oral cancers [17,18]

2.3 Optically Detectable Changes in Tissue

Optically detectable changes which occur in oral mucosal tissue during malignant transformation include changes in light scattering properties in bulk tissue due to underlying cellular changes, vasculature pattern changes due to angiogenesis, changes in native tissue fluorescence, and changes in the content of absorbers such as hemoglobin [19][6].

Modeling efforts have shown that the scattering cross section of dysplastic cells in the cervix is increased in comparison to normal cells due to nuclear morphology [12]. These changes may be analogous in the oral cavity and may be detectable using reflectance imaging modalities.

Changes in the autofluorescence of dysplastic and malignant tissue have been observed optically and there has been significant investigation into the cause of these changes. Changes in metabolic activity in premalignant tissue may be responsible for changes in the concentration of fluorescent metabolic byproducts. Among those substances, structures, or molecules suspected of being responsible for increased fluorescence are NADH, FAD, porphyrins, tryptophan, tyrosine, and phenylalanine. Tissue remodeling may also cause reduced fluorescence intensity due to the breakdown of collagen and elastin crosslinks. There has been significant research done into

discovering what excitation and emission wavelengths can be used to visualize these fluorophores. These wavelengths as well as the peaks of some of these fluorophores excitation and emission wavelengths are shown in table 1.

A study examining autofluorescence confocal imaging of freshly resected oral tissue by Pavlova et al. demonstrated that there is a strong autofluorescence signal from the stroma in normal tissue likely from structural components such as collagen and elastin crosslinks. This autofluorescence decreases significantly in both benign and dysplastic tissue. In the epithelium, dysplastic tissue showed an increase in autofluorescence whereas benign tissue showed a decrease [6].

Studies from dysplastic tissue from the cervix, with is histologically similar to oral tissue, showed a similar decrease in stromal autofluorescence. Pavlova et. al. showed that precancerous cervical tissue shows increased fluorescence in the epithelium due to NADH and mitochondrial FAD and decreased fluorescence in the stroma due to decreased density of matrix fibers [10]. Drezek et. al. showed increased epithelial fluorescence in cervical tissue sections due to NADH and a decrease in stromal fluorescence due to fewer collagen cross-links [20].

Table 1. Excitation and emission wavelengths for endogenous fluorophores and the tissue discussed in reference.

Fluorophore	Excitation λ	Emission λ	Tissue	Reference
NADH	290-370	440,460	Oral, Cervical,Breast	Pavlova et al. 2003, Drezek et al. 2003, Palmer et al. 2003, Richards-Kortum et al. 1996
FAD	430-460	515-520	Oral, Cervical,Breast	Pavlova et al. 2003, Drezek et al. 2003, Palmer et al. 2003, Richards-Kortum et al. 1996
Collagen & Elastin	300-370	400, 440	Cervical,Breast	Drezek et al. 2003, Palmer et al. 2003, Richards-Kortum et al. 1996
Tryptophan	280-290	340	Cervical,Breast	Ingrams et al. 1997, Palmer et al. 2003
Protoporphyrin IX	410	635	Oral	Ingrams et al. 1997, Betz et al. 2002
Tyrosine	270	300	Cervical	Drezek et al., 2003, Richards-Kortum et al. 1996

2.4 Current Detection Methods

Visual preclinical indicators of cancer in the oral cavity may include leukoplakia (white pathes) or erythroplakia (red patches) [4,21]. Because of the accessibility of the oral cavity to identify these conditions, visual inspection at a routine dental visit is currently the most common detection method. This method relies heavily on the experience and skill of the physician. Mistakes are easily made, however, because many benign conditions mimic the physical presentation of precancerous lesions. The sensitivity and specificity of visual oral examination has been systematically reviewed by Downer et al [22]. Downer identified eight prospective studies between 1980 and 2002 that involved conventional oral exam with gold standard verification provided by an expert observer. In four of the studies the screeners were general dentists and in four of the studies the screeners were trained health workers. Sensitivity ranged from 59% to 97%, specificity ranged from 75% to 99%, and meta-analysis resulted in a weighted pooled sensitivity of 85% and a specificity of 97%. Other reports of the performance of

visual oral screening include Sankaranarayanan et al (sensitivity 77%, specificity 76%) [23], Ramadas et al (sensitivity 82%, specificity 85%) [24], and Nagao et al (sensitivity 92%, specificity 64%) [25].

Invasive biopsy with histological examination is standard clinical practice but this method has drawbacks such as pain induced by physical biopsy and reluctance of physicians to perform biopsies for a condition that very well may be benign. To increase sensitivity of visual inspection, studies using vital dyes such as toluidine blue and Lugol's iodine solution have been performed with the drawback of having a high specificity [26,27].

Fluorescence and reflectance spectroscopy has been explored as a noninvasive detection method in several epithelial cancers including oral cancer [8,28-33]. The procedure generally involves collecting a spectrum from tissue using a point prove device either *in vivo* or *ex vivo* from a pre-malignant or malignant region and comparing it to a normal tissue spectrum from the same patient or study subject. This technique has shown promise in distinguishing normal tissue from dysplastic tissue but has the weakness of being able to sample only a small portion of. *In vivo* confocal and optical coherence tomography of the oral cavity to visualize malignant transformations at the cellular level are being investigated and show promise but may also be limited by a small field of view (typically 200-400μm) [34,35].

Widefield optical imaging has been studied in the cervix, gastrointestinal tract, lung, and oral mucosa and this method has the advantage of a large field of view (several centimeters) with the potential to detect changes in scattering properties of tissue as well to detect native tissue autofluorescence [14,36-50]. In the oral cavity, sensitivity and

specificity of diagnosis of severe dysplasia and cancer have been reported as high as 98% and 100% respectively for small sample sizes using visual autofluorescence techniques [21]. The device used in this study is currently FDA approved and excites tissue in autofluorescence mode between 400 and 450 nm. Other studies from both *in-vivo and ex-vivo* tissue have reported optimal excitation wavelengths for imaging oral cavity autofluorescence as 350, 380, 400, 420-440 nm excitation [51-54]. Most studies of autofluorescence imaging report qualitative observations of a decrease in blue/green fluorescence signal on oral lesions and/or and increase in red fluorescence [55]. Currently, these qualitative changes have not been described quantitatively.

2.5 Computer Aided Diagnostics

2.5.1 Overview

Automated or computer aided diagnosis based on imaging data has been demonstrated in multiple organ sites using instrumentation systems such as optical coherence tomography, mammography, and colposcopic images [56,57][58]. Numeric features are extracted from images or regions of images either in an automated manner or with a user selected region of interest (ROI). These features can describe any number of properties of the image data including statistical measures, textural measures, frequency domain measures, and colormetric measures. Often these features attempt to quantify some qualitative observation made from the imaging data. For example, texture analysis attempt to quantify qualitative descriptors such as smooth or rough. A classification algorithm can be trained using these features extracted from collected measurements, and used to discriminate new measurements into two or more classes. Classes may be

diagnostic categories such as *normal* and *abnormal,* where the *abnormal* class could contain data collected from the disease tissue the user is attempting to detect. The gold standard for this case is often histopathology.

There are many different kinds of classification algorithms. In this work I will limit the background discussion to a linear discriminant classifier which is based on estimating the statistical distribution of the measurements, and a decision tree classifier, which is a non-parametric classifier. The feature values extracted from the each image or image region form a measurement vector. Measurements may be separated into a training set and a test set. The classification algorithm is trained on the training set and the performance of the algorithm is determined from the test set. In the case of a small sample size, resampling methods such as cross-validation may be used.

2.5.2 Classification Algorithms

The linear discriminant function utilizes the Bayes decision rule for minimum-error-rate classification [59-61]. Multivariate normal densities and equal covariance matrices are assumed for each class ω_i . A training set is used to estimate the mean μ_i for each class and a pooled covariance matrix $\mathbf{\Sigma}$ for all classes. *A priori* probabilities $P(\omega_i)$ are provided as input indicating the relative prevalence of each class. As an example, the percentage of normal and abnormal measurements determined by histopathology may be used. Posterior probabilities, $P(\omega_i|\mathbf{x})$, the probability of a measurement vector $\mathbf{x}$ belonging in class ω_i, is calculated from the Bayes formula for each measurement in the test set and this value is normalized to fall between [0 1]. For the 2 class case, a linear

discriminant function is then used to classify the test observations, **x**, into one of the two classes based on comparing the output of the two discriminant functions for $i = 1,2$:

$$g_i(x) = -\frac{1}{2}(x-\mu_i)' \Sigma^{-1}(x-\mu_i) + \ln P(w_i)$$

The two discriminant functions are then combined into a single discriminant function

$$g(x) \equiv g_1(x) - g_2(x),$$

so that observation, **x**, is assigned class ω_1 if $g(\mathbf{x}) > 0$ and assigned to class ω_2 otherwise. If the threshold for the discriminant function is varied, or equivalently, if the decision boundary in varied between [0, 1] based on the posterior probabilities calculated from the test set, one may construct a receiver-operating characteristic (ROC) curve.

An ROC curve is a plot of the sensitivity versus 1-specificity for the classifier. This is equivalent to counting and plotting the true positives (TP) and false positives (FP) as the decision threshold is varied. One can choose a threshold at which to operate the classifier. Often this point is chosen to maximize both the sensitivity and specificity. The area under the curve (AUC) of the ROC is an often used figure-of-merit (FOM) for 2-class classifiers. An AUC of .5 indicates no diagnostic performance while an AUC of 1.0 indicates perfect diagnostic performance. The AUC can be used to evaluate and compare the performance of different classifiers.

A limitation of the linear discriminant classifier is that normal distributions for the observations are assumed. If the parameters for the normal distributions are estimated from a small training set, this assumption may cause the classifier to poorly generalize to

additional new data. Decision trees are a commonly used nonparametric method for multiclass classification which help to classify the data in an intuitive manner and make no assumptions about the distributions of the observations [62]. Decision trees are structured from a collection of nodes connected by branches. Each node 'asks a question' concerning the value of a single feature, for example, whether the value of a particular feature if above or below a specific threshold. The tree begins with a root node. The root node is connected, by branches, to other nodes which eventually terminate at leaf nodes where observations are put into a class. For a particular decision tree, one must specify how the tree is constructed from the training data, including how splits are made (i.e. what feature to use to make new branches), and how to determine when to declare leaf nodes. Feature selection is implicit in the construction of the tree as a decision needs to be made about which feature to test at each node.

It is possible to allow the tree to grow until all training data is classified perfectly, but this is likely to generalize poorly to new data, especially with small training sets. In order to avoid overtraining, a decision tree may be pruned using cross-validation to find the smallest tree at which adding further nodes does not statistically decrease the cost of the tree. One way to define the cost of the tree is in the *zero-one* sense, where the cost of misclassifying an observation is 1 and the cost of correctly classifying an observation is 0.

It is not intuitive to create ROC curves using decisions tree classification methods and therefore the AUC criterion is not often used as a FOM for evaluating the classification performance. Sensitivity, specificity, positive predictive values, and

negative predictive values are commonly used to evaluate decision tree classification performance.

2.5.3 Feature Selection and Resampling methods

Using an excessive number of features with a small data set may result in overtraining a classification algorithm [59,63]. A feature selection algorithm can be used in order to determine the most important and diagnostically relevant features which can then be used for classification. An important attribute of a feature selection algorithm is that it preserves the feature set. This differs from feature extraction methods, such as principle component analysis (PCA), which transforms the original feature space in a new feature space.

The feature selection method used in this work is termed a forward sequential search (FSS) algorithm. This method starts with an empty feature set and adds features one at a time to maximize a specified criterion determined by a mining algorithm until some defined threshold or a specified feature set size is reached. This method is not exhaustive and risks missing optimal subsets but is much less computationally expensive than a complete search which searches all possible combinations of features [64].

An example of a criteria which may be used by a FSS algorithm is the area under the curve (AUC) of the receiver-operating characteristic (ROC) produced by a classifier. This type of criteria is termed a *dependent criteria* because it is utilizes a mining algorithm based on classification accuracy and therefore tends to pick features which achieve higher final performance compared to a search algorithm using *independent* criteria such as distance measures or correlation measures [64].

The FSS algorithm works as follows: Starting with a set of one feature, classification of the training set is performed with cross-validation utilizing the FSS algorithm to find the single feature which produces the most accurate classification. This is repeated with two features, then three, four, ect. The optimal number of features is determined when adding additional features does not increase the classifier performance. This entire procedure may be repeated many times to provide statistically significant results. The AUC values from each iteration may then be compared using statistical methods to determine when no statistically improvement is achieved by adding additional features. Once the optimal number of features is determined, the specific optimal feature subset can be determined by selecting the most commonly chosen features.

Resampling methods are useful for avoiding overly optimistic estimation of classifier performance due to overtraining. K-fold cross-validation is a commonly used method. It consists of separating an available data set into k, approximately equal sized, subsets or folds.

k-1 folds are then used to train a classification algorithm and the remaining fold is used for validation. This is repeated k times so that the validation set is a different fold at each iteration. The classifier performance is then averaged over the k-folds to produce a performance estimate with minimal optimistic bias [63].

CHAPTER 3: A MULTISPECTRAL OPTICAL IMAGING DEVICE FOR *IN VIVO* DETECTION OF ORAL NEOPLASIA[1]

3.1 Introduction

Oral cancer poses a significant world health problem, ranking as the sixth most common cancer globally [4], causing over 127,000 deaths worldwide each year [65]. In the United States alone, there were an estimated 30,990 cases of oral cancer diagnosed and 7,430 deaths in 2006 [3]. Oral cancer often goes undiagnosed until the later stages of development, resulting in high mortality. In the United States, the five-year survival rate for localized oral cancer is 81%, but this drops to only 30% for advanced disease [66], suggesting that early detection has the ability to save many lives. In addition, patients who survive an initial occurrence of oral cancer are known to be at an increased risk of developing a second malignancy highlighting the importance of continued surveillance of this population [67].

The most common screening method for oral cancer is visual inspection and palpation of the mouth. Physicians inspect for clinically evident oral lesions such as leukoplakia (white patches) or erythroplakia (red patches) which are associated with increased risk of oral cancer [4,21]. Visual examination relies heavily on the experience and skill of the physician to identify and delineate early premalignant and cancerous changes. Several benign conditions, such as lichen planus, inflammation and hyperkeratosis, mimic the clinical presentation of precancerous lesions, and visual

[1] This chapter has been published in the following journal article: Darren Roblyer, Sokolov KV, El-Naggar A, Williams M , Kurachi C , Gillenwater A, Rebecca Richards-Kortum, "*A Multispectral Optical Imaging Device for In Vivo Detection of Oral Neoplasia*", Journal of Biomedical Optics. 2008 Mar-Apr; 13(2)

inspection with standard white light illumination may not yield sufficient contrast between normal and abnormal tissues. Once a suspicious lesion is identified, biopsy and histological examination is required for definitive diagnosis. Biopsies are invasive, painful, costly, and require familiarity and skill, and are therefore typically limited to highly suspicious lesions. Additionally, many lesions are heterogeneous in morphology and visual appearance, and biopsy diagnosis may not be representative of the entire lesion due to the small sampling area [68]. Approaches which can help differentiate between normal and neoplastic areas and which can reduce the need for evaluator experience have the potential to greatly facilitate and improve early diagnosis of oral cancer.

Optical imaging has the potential to address these clinical challenges. Contrast between normal and neoplastic areas can be increased beyond that which is available with standard white light by tuning the illumination and detection conditions. Reflectance imaging can detect local changes in scattering and absorption of tissue and fluorescence imaging can probe changes in the biochemical composition of tissue by revealing levels of endogenous fluorophores. These changes have shown to be indicative of malignant progression [55] and are targeted in a variety of imaging modalities to detect and diagnose premalignant changes in different anatomic sites including the cervix [50], gastrointestinal tract [41], lung [69], and oral mucosa [21,38,44,53,55,70-72]. High resolution optical imaging techniques such as confocal microscopy, and optical coherence tomography have the potential to image changes in tissue architecture and cellular morphology. A drawback to *in vivo* cellular resolution microscopy however, is the small field-of-view (FOV) interrogated, making screening of an entire mucosal surface impractical without an initial means of guidance towards suspicious areas. Wide-

field optical imaging techniques allow clinicians to screen several centimeters at a time for identification of margins and optimal sites for further interrogation by small FOV imaging techniques or biopsy.

Several wide-field optical imaging systems have shown promise to improve detection of neoplastic lesions in the oral and oropharyngeal regions. Lane et al. presented a non-magnifying hand-held device for direct visualization of oral cavity tissue fluorescence [21]. The system uses a metal-halide lamp with emission peaks at 405 and 436 nm to excite tissue autofluorescence. Images obtained using this device showed a characteristic decrease of green fluorescence associated with oral precancer and cancer. Decreased green fluorescence distinguished normal tissue from severe dysplasia, carcinoma in situ, or invasive carcinoma in 50 biopsy sites from 44 patients with a sensitivity of 98% and specificity of 100%, using histology as the gold standard. This device is available commercially as the VELscope®, which is FDA approved and is currently in clinical use. De Veld et al. provide an excellent review of the status of *in vivo* autofluorescence imaging for oral oncology [55].

While fluorescence imaging alone has shown great promise for detection of neoplastic lesions in the oral cavity, incorporating additional imaging techniques may increase the effectiveness and sensitivity of wide-field optical devices. Techniques which are capable of detecting increases in vascular density and changes in light scattering properties may be particularly useful, since these features have previously been associated with the development of cancer, and are not easily detected by fluorescence imaging alone [73]. Additionally, reflectance imaging systems tend to be simpler in design compared to fluorescence devices, thereby reducing instrument complexity and

cost. In narrow band (NB) reflectance imaging, tissue is illuminated using a narrow wavelength band (e.g. 10-20 nm) and reflected light is imaged. The optical contrast of microvasculature close to the epithelial surface may be increased by selecting an illumination band which matches peaks in the absorption spectrum of hemoglobin [74]. Another approach to increase contrast is to control the polarization of the illumination source and detected light. In orthogonal polarization reflectance (OPR) imaging, tissue is illuminated with linearly-polarized light; a second linear polarizer is placed in front of the detector and oriented so that only remitted light with a polarization orthogonal to the illumination is captured. This has the effect of selectively detecting photons which have undergone multiple scattering events in the tissue (resulting in a reduction in degree of polarization), and rejecting those singly-scattered photons returning from the surface of the tissue. This technique enables observation of deeper or birefringent tissue structures and enhances the prominence of microvasculature due to increased hemoglobin absorption [75]. When both orthogonal and parallel polarization images are captured, they can be mathematically combined to form a composite image which has been shown useful for imaging skin pathologies [75]. Lindeboom et al. have described orthogonal polarization spectral (OPS) imaging of oral tumor sites using a magnifying handheld device coupled to a CCD camera, designed to provide maximum contrast between the microvasculature and surrounding tissue [76].

The Multispectral Digital Microscope, or MDM, was developed to investigate the diagnostic effectiveness of combining multiple imaging modalities in a single device compared to measuring reflectance or fluorescence signals alone. The MDM acquires high-resolution, wide-field images *in vivo*, using white light illumination, fluorescence

excitation, NB, and OPR modes. The MDM employs a wide range of illumination wavelengths to investigate which provides the greatest contrast between normal and neoplastic areas for each imaging mode. In this paper, we describe the design of the device, evaluate its optical performance, and present sample images obtained in human subjects with the MDM system.

3.2 Materials and Methods

3.2.1 Instrumentation

Figure 1 shows an optical schematic of the Multispectral Digital Microscope (MDM). The OPMI® pico dental microscope (Carl Zeiss, Jena, Germany) is used as the core optical component of the MDM. These is a single objective lens and the magnification setting is adjusted by turning a turret to one of five positions. Light is routed from the source to the microscope head by a fiber optic light guide. The working distance of the microscope is 25 cm. The field of view (FOV) of the device is dependent on the choice of magnification setting and ranges between 1 cm and 4 cm in diameter. With a monitor size of 17 inches, the system magnification ranges from 8.5 times to 34.7 times.

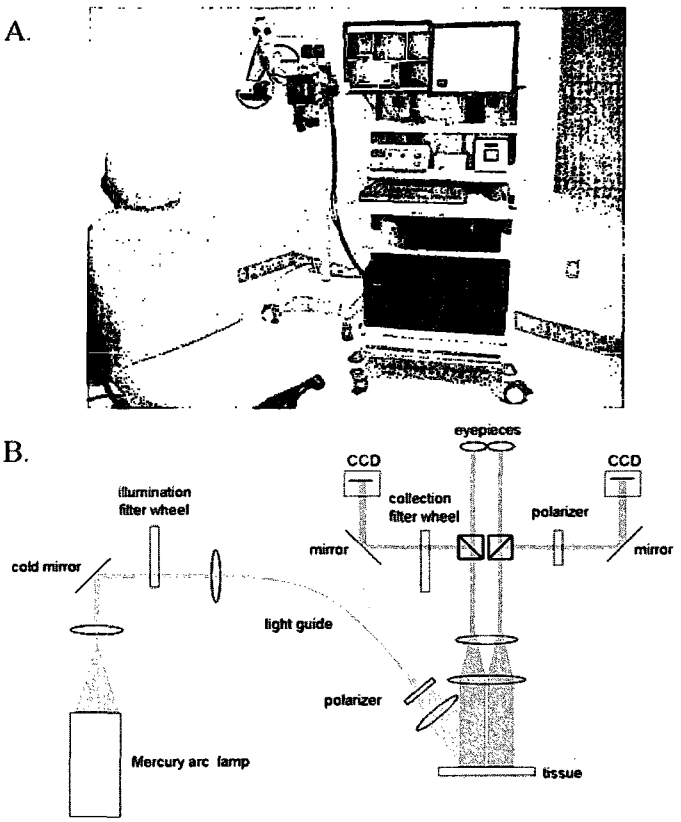


Figure 1 A. The Multispectral Digital Microscope (MDM) in the Head and Neck Clinic at the MD Anderson Cancer Center. B. System optical schematic.

We replaced the original light source with an air-cooled 100 Watt mercury arc lamp (Photon Technology International, Birmingham, NJ), housed in a custom light box. The lamp and ellipsoidal reflector produce a converging beam which is collimated using a 25 mm plano-convex lens (Lambda Optics, Costa Mesa, CA). To protect both the patient and optical components from excessive heating, a custom cold mirror (Barr Associates, Westford, MA) eliminates unwanted IR radiation from the lamp, and an absorbing glass filter removes UV light at wavelengths below 350 nm. The beam then passes through a ten-position excitation filter wheel (Sutter Instruments, Novato, CA) and

through a second plano-convex lens which focuses the beam onto the proximal face of a 5 mm diameter fused quartz light guide (Fiberoptic Systems, Simi Valley, CA), which routes the light to the microscope head. Since the microscope was not originally designed to support near UV excitation, a custom machined adaptor was implemented to bypass the original illumination optics and to hold the distal end of the light guide and a fused quartz focusing lens (Newport Optics, Stratford, CT). This adaptor provides a significant increase in near UV excitation power and subsequently reduces exposure times required for fluorescence imaging. The adaptor also holds a linear polarizer (Chroma Technologies, Rockingham, VT) which is servo controlled and engaged by the LabView (Austin, TX) interface.

The illumination filter wheel contains narrow band illumination filters at 420 nm, 430 nm, 530 nm, and 600 nm for NB imaging. These wavelengths were selected to match hemoglobin absorption features, and also to provide multiple wavelengths throughout the visible spectrum to study the effects of wavelength on penetration depth. The illumination filter wheel also contains narrow band illumination filters at 365 nm, 380 nm, 405 nm, and 450 nm for autofluorescence excitation. These excitation wavelengths were chosen based on previous work by Heinzelman et al. and Utzinger et al. [53,77]. The full-width at half-maximum (FWHM) transmission bandwidth for the excitation filters ranges from 20 nm for NB imaging, to 50 nm for UV illumination (Chroma Technologies, Omega Optics). A 1.0 neutral density filter is used to attenuate the white light illumination power. A light-blocking disk occupies one position in the excitation filter wheel and is used for measurement of the background light level.

Table 2. Irradiance and typical exposure times for each illumination condition.

Illumination λ (nm)	Imaging Modality	Irradiance (mW/cm^2)	Typical Exposure (ms)
365	fluor.	11.15	100
380	fluor.	8.86	160
405	fluor.	7.00	400
450	fluor.	5.10	500
white	reflect.	12.02	16
420	reflect.	3.13	64
430	reflect.	9.62	13
530	reflect.	1.91	24
600	reflect.	4.17	64
white polarized	polarized	3.47	100

The emission arm of the MDM consists of a Zeiss objective lens and 50/25/25 beam splitter which passes 50% of the emission light to the microscope eyepieces and 25% each to two CCD detectors. An emission filter wheel (Oriel Optics, Stratford, CT) is located in one detection arm and contains longpass filters and a linear polarizer. A longpass filter with a 50% transmission at 410 nm is used with the 365 nm and 380 nm excitation, one centered at 430 nm is used with 405 nm excitation wavelength, and one with a 50% transmission at 475 nm is used with 450 nm excitation. An additional linear

polarizer is placed in the opposing detector arm as shown in Figure 1. In one arm, the polarizer is oriented parallel to the excitation polarizer and in the other, the polarizer is oriented perpendicular. The opposing polarizers allow for simultaneous OPR and parallel polarization imaging; this feature is useful for reducing changes in FOV which can occur due to patient movement.

Two Retiga Exi (QImaging, Burnaby, BC Canada) 12-bit color cooled cameras are used to collect image data. This camera model was chosen because of its low light sensitivity and compatibility with the LabView programming environment. The sensor is a Sony ExHAD ICX285 progressive-scan high-sensitivity CCD chip, providing adequate spectral sensitivity at chosen wavelengths of interest, from 400 nm to 650 nm. The chip is covered by a Bayer Mask mosaic for color imaging. Image data is saved in raw Bayer format and color interpolation is performed in post-processing based on color balancing standards.

Control of the MDM system is by accomplished by a custom LabView graphical interface. The interface provides the operator with the ability to image with one excitation/emission wavelength pair at a time, or to collect a rapid sequence of image data over the range of the MDM's capabilities. The interface also allows the user to store information such as study number and patient data which are linked and saved with the optical data. Imaging parameters are recorded for all data collected, and background measurements are taken with each sequence.

The MDM takes about one minute to collect a complete image sequence. The sequence is composed of white light reflectance, fluorescence, NB, OPR, parallel polarization reflectance, and background data sets. A predefined set of exposure and gain

settings was established for each image in the sequence based on measurements of normal volunteers. In order to account for inter-patient variability, measurement site variability, and changes due to disease, the predefined set of exposure times could be manually scaled up or down in 20 percent intervals during imaging. Additionally, four duplicate images were taken for each illumination/emission setting at lower and higher exposure times relative to the initial choice. Table 2 shows typical exposure times based on the most commonly used exposure settings used in the pilot clinical trial.

A set of standards was chosen to quantify the performance of the MDM as well as to track any changes occurring in the device over time. Positive reflectance standards include 99%, 75%, and 50% reflectance spectralon disks (Labsphere, North Sutton, NH), and red, green, and blue portions of a Macbeth Color Chart (GretaMacbeth LLC, New Windsor, NY). A 2% spectralon disk serves as a negative reflectance standard. Positive fluorescence standards include a set of four fluorescent slides (Microscopy/Microscopy Education, microscopyeducation.com) and the negative fluorescence standard is a frosted quartz disk (Mark Optics, Santa Ana, CA).

For white light illumination, color balance was achieved by imaging a white balance sheet and adjusting the RGB ratio in software so that equal pixel intensity values were obtained in the red, green, and blue channels. For fluorescence imaging, a constant color balance was used throughout the study to maximize qualitative diagnostic ability and to reflect the natural blue-green tissue fluorescence color. The fluorescence color slides were used as a standard to track any changes in color response of the camera over time in fluorescence mode.

3.2.2 Pilot Clinical Study

Several sites within the oral cavity of normal subjects were imaged using the MDM as part of an IRB approved protocol at Rice University. In a separate, ongoing pilot clinical study conducted at The University of Texas MD Anderson Cancer Center (MDACC), the MDM was used to image clinically abnormal and normal oral mucosal sites in patients. This protocol was approved by the Institutional Review Boards at both Rice University and MDACC.

The measurements were performed either in the outpatient clinic or in the operating room on patients under general anesthesia prior to surgery. The areas to be imaged were chosen by the clinician and included areas that had previously been identified as abnormal by clinical examination and/or biopsy. The physician first evaluated these targeted areas using the MDM under white light illumination, adjusting the focus and FOV. Several measurement sequences were then taken from each subject including abnormal sites, and when possible, a corresponding contralateral normal site. Excess saliva was suctioned prior to imaging. Because teeth exhibit particularly strong autofluorescence, attempts were made to cover any teeth in the FOV with low-fluorescence gloves or a mouth guard. The head of the patient was positioned and held by the physician during imaging in order to reduce motion artifacts. In patients seen in the clinic, biopsies were obtained from clinically abnormal sites and a contralateral normal site. In patients who underwent surgical excision, the resected specimens were histologically examined. A clinical diagnosis for normal and abnormal-appearing areas (graded as normal, abnormal but not suspicious for neoplasia, suspicious for neoplasia, or

cancer) was rendered by the clinician. The histopathology diagnosis of biopsy sites and resected tissue from imaged areas were considered the gold standard for diagnosis.

Four observers (DR, CK, AG, RRK) examined the images to identify features that differed in abnormal sites compared to clinically normal areas, and which might be further explored for diagnostic relevance. Qualitative observations of image data were correlated with clinical observations regarding peripheral extent of abnormal changes and clinical diagnostic category. Pathology results from resected tissue and biopsies allowed comparison between MDM image data and histopathologic diagnosis of lesions.

3.3 Results

3.3.1 Instrument Performance

The resolution of the MDM was determined by imaging a USAF spatial resolution target, as shown in Fig. 2. A line spacing of 15.6 μm (group 5, element 1) can clearly be discriminated at the highest magnification. The appearance of shadowing behind each element is due to the finite thickness of the glass substrate, and is not intrinsic to the MDM system. Measurements of white light illumination on a 99% reflectance spectralon standard (Labsphere, North Sutton, NH) show a 4 cm diameter illumination spot size. This is adequate to illuminate all but the edges of the FOV at the most commonly used magnification. The illumination pattern is Gaussian due to light guide coupling with the highest intensities in the center of the FOV.

To confirm the absence of excitation filter leakage and adequacy of performance of longpass filters, a 2 inch diameter frosted quartz disk was imaged in fluorescence mode. Exposure times and gain settings matching or exceeding those used for normal

tissue fluorescence imaging were used at each excitation/emission wavelength pair to confirm the absence of signal from the non-fluorescent frosted quartz disk. The ratio of autofluorescence signal from tissue to frosted quartz was always greater than 10:1.

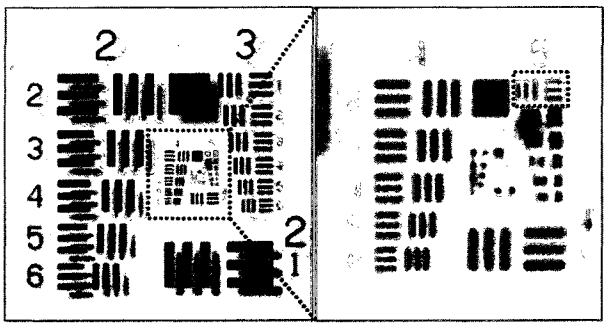


Figure 2. A USAF Resolution target imaged at high magnification. Element 5 group line spacings are 15.6μm (shown in the dashed box) are easily discriminated in the zoomed-in image on the right.

Intra-sequence image registration was affected by patient movement, generally causing some changes in the FOV during the course of the image sequence. Figure 4 shows several images taken from a typical sequence in a clinical setting at the most common magnification setting. Using user-defined selection of points common to both images, the linear translation was less than 60 pixels (less than 5% of the FOV) between any two of the images in the sequence in both the x and y directions.

UV radiation during all human measurements is well below the American Conference of Governmental Industrial Hygienists maximum allowed rating [78].

3.3.2 MDM Images of Normal Oral Sites

Sites imaged in the oral cavity from four normal volunteers included tongue, buccal, lip, gingiva, hard palate, soft palate, and floor of mouth. Fine vasculature was clearly identifiable in images acquired from the lip, floor of mouth, hard palate and soft palate using white light, NB, OPR and fluorescence techniques. These images serve as further examples of the resolution capabilities of the MDM. Figure 3 shows an image of the lower lip of a normal volunteer under several illumination conditions. The white light image (3A) shows microvasculature from a variety of depths beneath the epithelial surface. The OPR image illustrates similar vascular patterns; specular reflection is no longer visible in the image, and spatial resolution is somewhat reduced because the OPR technique selectively records photons that have undergone more scattering events in the tissue. The NB image obtained with 420 nm illumination (3B) shows only the superficial, fine vasculature due to the reduced penetration of this wavelength; vessel contrast is also increased because this wavelength matches the Soret absorption band of hemoglobin. As the illumination wavelength is increased from green (3D) to red (3F), vessels deeper within the tissue are visible in the image. The NB image obtained with 600 nm illumination shows a loss of contrast and spatial resolution due to the increased ratio of scattering to absorption and increased mean free path between scattering events at this wavelength. Vasculature was not as apparent in the buccal and tongue.

Tissue autofluorescence is predominantly blue at UV excitation wavelengths (3E) and blue-green at the longer 450 nm excitation. The hard palate, soft palate, floor of mouth, and buccal mucosa provided a higher fluorescence signal than the tongue, gingiva, and lip. The midline of the hard palate was particularly bright. Teeth were

highly fluorescent and tooth fluorescence could be seen through portions of the gingival mucosa. Blood vessels appeared dark under fluorescence mode compared to the surrounding tissue. Red fluorescence occasionally appeared on the dorsal tongue.

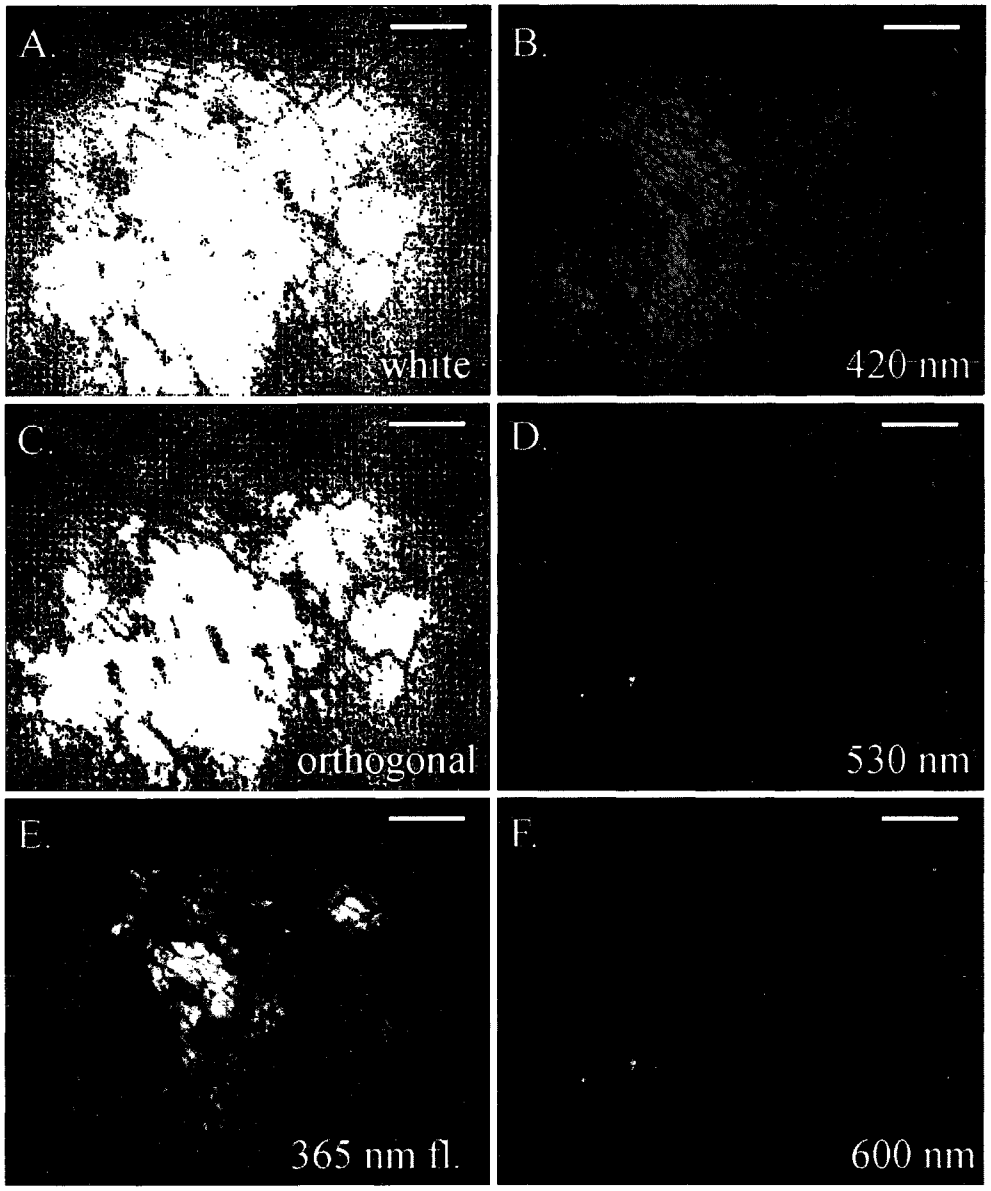


Figure 3. Images of normal volunteer inner lip. A. White light illumination image. B. 420 nm NB image. C. OPR image. D. 530 nm NB image. E. 365 nm excited fluorescence image. F. 600 nm NB image. Note increased contrast of vasculature in C. compared to white light image. Fine vasculature is visible in B., whereas, deep, larger diameter vasculature is visible in F. Scale bars indicate 2.5 mm.

3.3.3 MDM Images of Oral Dysplasia and Cancer

Images of a premalignant lesion and an invasive cancer are presented as representative examples to demonstrate the capabilities of the MDM system for examination of the oral cavity. Some images were cropped to only show relevant areas.

Figure 4 shows images obtained from a thin leukoplakia lesion (indicated by arrows) on the right mandibular gingiva. The thin leukoplakia in the OPR image (4C) is less evident demonstrating the ability of OPR imaging to selectively detect photons that have scattered more deeply into the tissue. The blue NB image (4D) shows the leukoplakia as brighter in reference to surrounding tissue compared to the white light image.

Figure 5 shows images acquired from a subtle lesion on the right lateral tongue. The clinical impression of the lesion was leukoplakia, but not overly suspicious for dysplasia or cancer. Following imaging, a 2.4 x 1.0 x 0.4 cm^3 volume was surgically resected and histopathology showed moderate squamous dysplasia (Fig. 5H). The standard white light image (5A) shows some patchy irregularities in the mucosal surface. Using the green NB (5E) and the OPR imaging conditions (5C), an apparent increase in visual contrast was observed between the lesion and surrounding normal areas (indicated by arrows in each case). A decreased blue/green autofluorescence (DA) was observed in the area of the lesion in fluorescence images at 365 nm, 380 nm and 405 nm excitation (5B, D, F). An image of a contralateral normal area imaged using white light is shown in 5G for comparison.

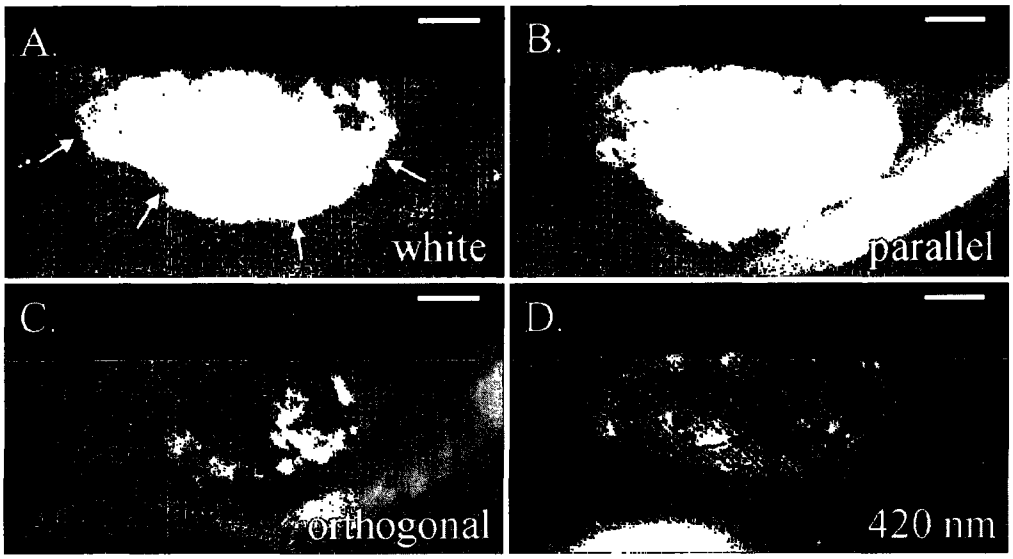


Figure 4. Images of leukoplakia lesion on right gingiva. A. White light illumination image. Arrows indicate area of leukoplakia. B. Parallel polarization image. C. OPR image. D. 420 nm NB image. Note that the OPR (C) allows visualization of tissue below the thin leukoplakia and the 420 nm OPR image (D) accentuates areas of leukoplakia. Scale bars indicate 2.5 mm

Figure 6 shows images acquired from a premalignant lesion on the left lateral tongue. The clinical impression was erythroplakia; a reddish lesion associated with a high risk of dysplasia or early carcinoma. Histopathology from a biopsy of the lesion indicated severe squamous dysplasia with a focal ulceration and chronic inflammation (6H). Both the OPR image (6C) and the green NB image (6E) showed an area of abnormality (appearing darker red on OPR and darker on NB images as indicated by arrows) which is more extensive in peripheral extent and has increased contrast as viewed against the surrounding mucosa. In the images obtained using all four fluorescence

excitation wavelengths a comparative decrease or loss of blue or green fluorescence was seen in the abnormal site. The total area of DA was larger at 405 nm excitation (6F) than at 365 nm (6B) and 380 nm excitation (6D). In addition, a striking ring of increased red fluorescence surrounding the lesion was observed in the fluorescence images, and is most apparent at 405 nm excitation. This red fluorescence is consistent with the presence of endogenous porphyrins. A white light image obtained from a contralateral normal area is shown in 6G for comparison.

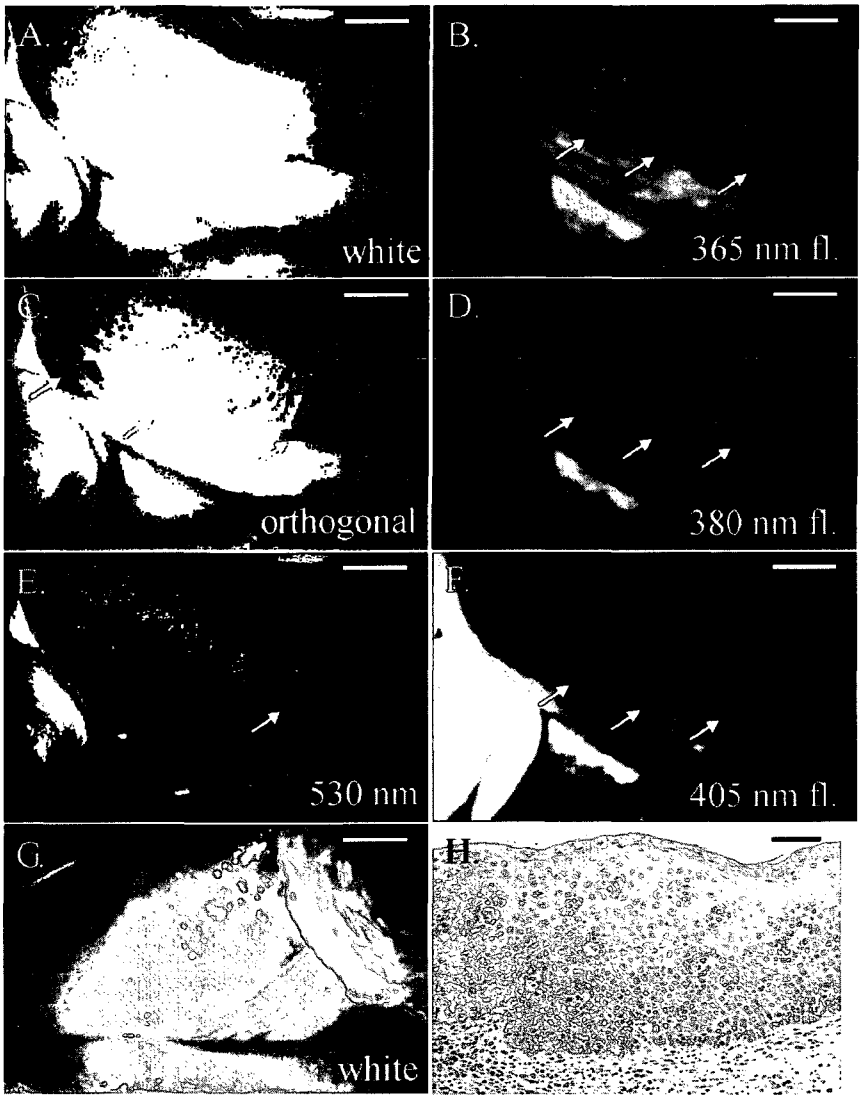


Figure 5. Images from tongue of patient with leukoplakia. A-F show a FOV containing areas along the lateral aspect of the tongue which were resected and found to contain moderate dysplasia. A. White light illumination image. B. 356 nm excited fluorescence image. C. OPR image. D. 380 nm excited fluorescence image. D. 530 nm NB image. F. 405 nm excited fluorescence image. G. White light illumination image of contralateral normal site. H. H&E stained tissue from abnormal area indicating moderate dysplasia.

Arrows indicate areas with loss of fluorescence or increased contrast. Scale bars indicate 5 mm in A-G and 200 μm in H.

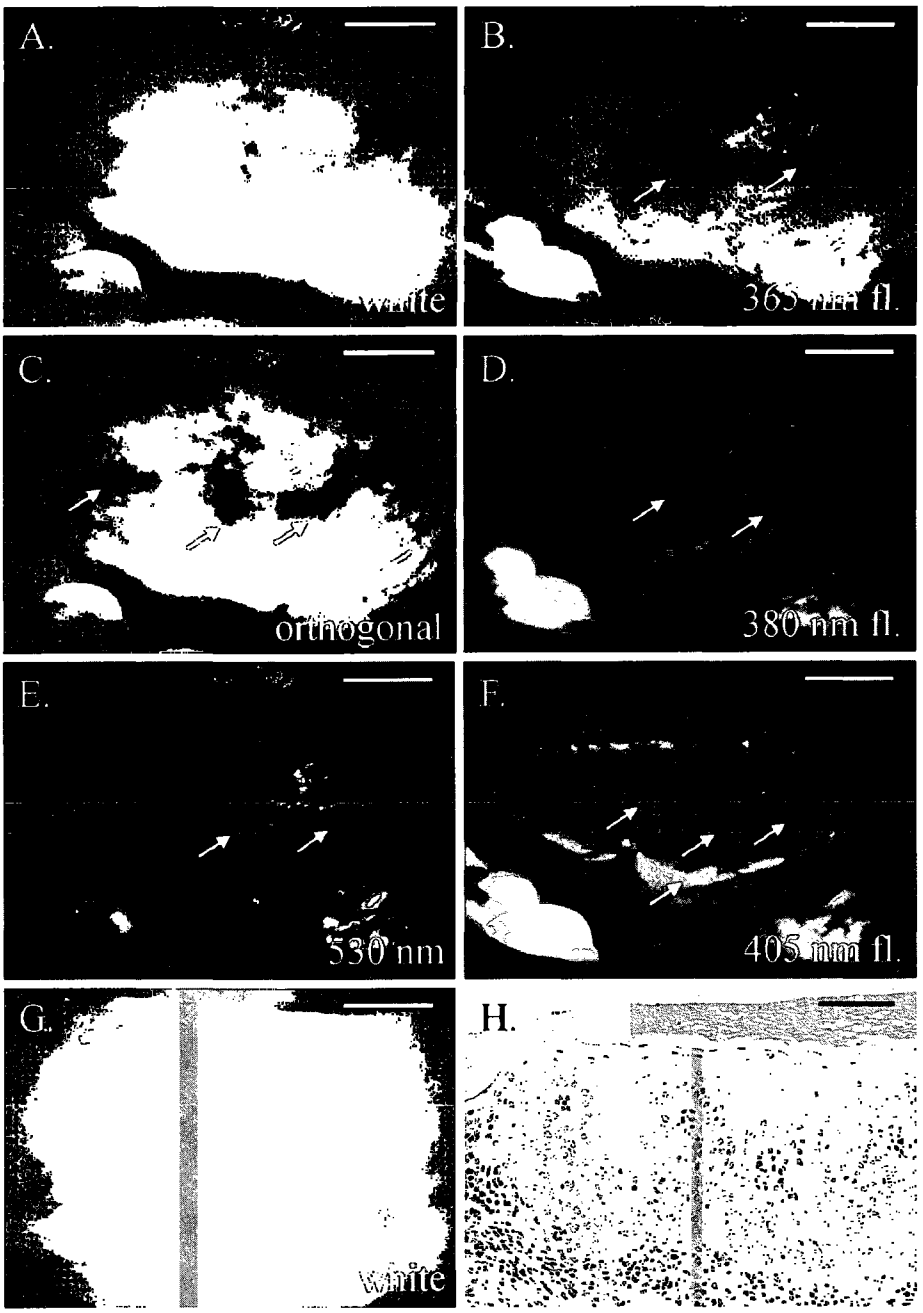


Figure 6. Images from ventral tongue of patient with erythroplakia. A-F show a FOV of the left lateral tongue containing erythroplakia from clinical appearance and severed dysplasia from

histophathology of biopsy site. A. White light illumination image. B. 356 nm excited fluorescence image. C. OPR image. D. 380 nm excited fluorescence image. D. 530 nm NB image. F. 405 nm excited fluorescence image. G. White light illumination image of contralateral normal site. H. H&E stained tissue from abnormal area indicating severe dysplasia. Arrows indicate areas with loss of fluorescence or increased contrast. Scale bars indicate 5 mm in A-G and 200 μm in H.

Figure 7 shows images from the right lateral tongue in a subject with histologically confirmed carcinoma. A previously biopsied and ulcerative lesion is shown in the center of the FOV. In the OPR image (7B), the blue NB image (7C), and the green NB image (7C), increased contrast is noted by arrows in the tissue surrounding the ulcer. The blue NB image (7C) shows larger dark areas in the ulcerative lesion and in the bottom right compared to the white light image (7A). The red NB image (7G) shows smaller dark areas compared to white light. The fluorescence images show a decrease of blue/green fluorescence surrounding the ulcerative lesion, and this DA is more apparent for 405 nm and 450 nm excitation (7F, 7H) than at 365 nm excitation (7ED). Following imaging a 3.2 x 2.5 x 1.0 cm^3 portion was resected and was determined by histopathology to contain invasive squamous carcinoma centrally, with dysplasia near the margins of resection.

In these representative examples, we observe one or more of the following features in areas histologically determined to be abnormal: a decrease of blue/green fluorescence, increased red fluorescence, and an increase of contrast in highly vascular regions. Blood on the surface of the tissue also appeared dark compared to white light under blue and green narrow band illumination. Teeth and gloves occasionally interfered with fluorescence imaging of tissue due to their autofluorescence, as can be seen in the lower left corner of the fluorescence frames in Figure 5.

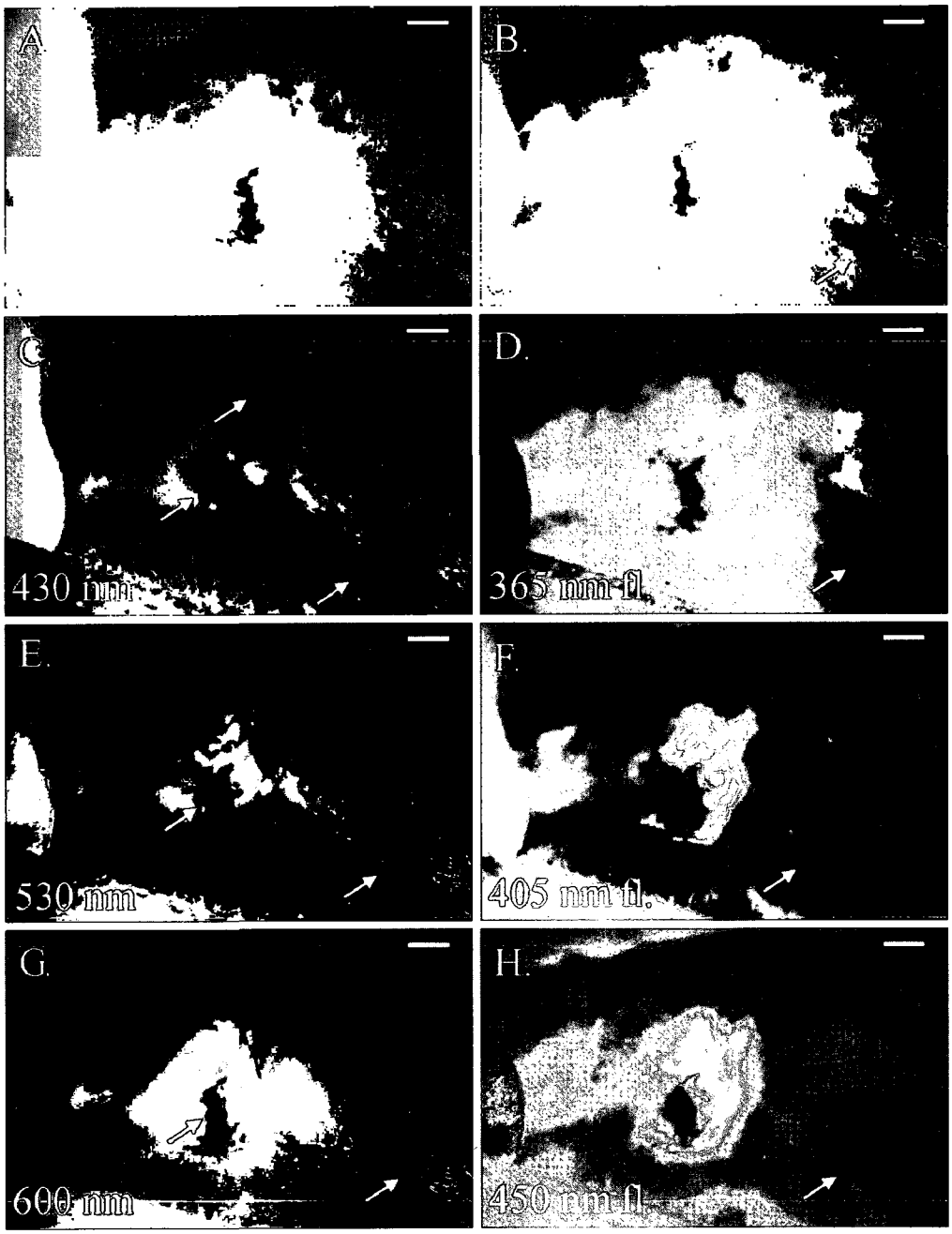


Figure 7. Images from cancer on right lateral tongue. A. White light illumination image. B. OPR image. C. 430 nm NB image.. D. 365 nm excited fluorescence image. E. 530 nm NB image F. 405 nm excited fluorescence image. G. 600 nm NB image. H. 450 nm

excited fluorescence image. Arrows indicate increased contrast or decreased autofluorescence. Scale bars indicate 2.5 mm.

3.4 Discussion

In this study, DA was observed in many clinically abnormal areas compared to surrounding normal tissue, and this finding was common to all four fluorescence excitation wavelengths. DA was also observed surrounding submucosal, exophytic or hyperkeratotic abnormal areas as well as in clinical lesions lacking the white keratotic surface associated with leukoplakia. Areas with DA did not always correspond to areas where increased contrast was seen using other modalities. Increased red fluorescence was observed in about half the patients imaged to date. It appeared most often on cancerous regions and on necrotic tissue but occasionally appeared in clinically dysplastic or precancerous areas. Red fluorescence did not usually correlate with DA or with increased contrast from NB or OPR imaging.

Blue and green NB images tended to increase contrast of visible blood vessels and areas suspected of containing a high density of superficial microvasculature. Blue illumination revealed fine vasculature which appeared superficial whereas illumination at green or red wavelengths revealed larger diameter, deeper vasculature. 530 nm illumination was most useful in increasing negative contrast in lesions in study patients. Blood on the surface of the tissue also appeared dark under blue and green narrow band excitation. OPR had the effect of reducing specular reflection and increasing negative contrast in some abnormal regions compared to imaging under conventional white light illumination. Both OPR and NB appear useful for detecting areas with high

microvascular density, and in many cases, the resulting images showed extended margins of clinically abnormal areas compared to white light imaging. In most cases, discrete vessels were observed in normal contralateral measurements and at the periphery of abnormal areas, while diffuse, homogeneous red areas were observed on abnormal lesions with flat contours. Elevated lesions in cancerous or precancerous areas often obscured vascular detail.

The characteristic DA associated with neoplasia has been extensively explored to determine abnormal areas in previous imaging studies in the oral cavity [21,38,72,79]. It has been suggested that a decrease of collagen cross-links and decreased density of matrix fibers is the major contributor to loss of green fluorescence; these are important features that may be associated to malignant progression. Pavlova et al. and Drezek et al. have shown that precancerous cervical tissue exhibits decreased stromal fluorescence due to a decreased density of matrix fibers [10]. Lane et al., who used excitation light from 400 nm to 460 nm, suggest collagen-related DA as a major mechanism enabling contrast between the appearance of neoplastic areas and normal tissue using the Velscope® [21,80].

Increased red fluorescence in the oral cavity is well documented for carcinoma, and in previous imaging studies [38,39,70,71], red fluorescence was best observed when excited at approximately 410 nm. Inaguma and Hashimoto found the presence of red fluorescence in eighty-five percent of the 78 lesions of oral carcinomas they investigated [70]. Ingrams et al. have obtained 90% sensitivity and 91% specificity for the discrimination of normal and dysplastic/malignant mucosa using the presence of red fluorescence [54]. This red fluorescence is generally attributed to the presence of

porphyrins whose origins in biological tissue are disputed. It is unclear whether they are a byproduct of abnormal metabolism in tissue or if it is produced by bacteria contaminating the surface of the tissue.

Changes in contrast using NB and OPR have been previously observed in oral tissue and are likely due to increased microvasculature density and structure associated with malignant transformation. Pazouki et al. have shown an increase in vascularity in oral tissue throughout malignant progression [73]. By choosing reflectance excitation wavelengths which correspond to hemoglobin spectral features, the narrow band reflectance images maximize microvasculature visualization and contrast. Subhash et al. used this technique to detect abnormal tissue in the oral cavity. They showed that the ratio of two oxygenated hemoglobin dips at 540 nm and 575 nm decreased in excised malignant tissue compared to normal [74]. De Felice et al. speculated that observed optical changes in the color spectrum observed in bronchopulmonary dysplasia may be caused by an altered microvasculature network [81]. Wavelength dependent photon penetration depth utilized in narrow band reflectance imaging may also provide useful depth dependent diagnostic information.

OPR techniques enabled us to look deeper into tissue by utilizing the ability to reject singly-scattered photons that have retained their polarization state. Photons that have had their polarization state altered either by multiple scattering events or by birefringent properties of the tissue are selectively detected. This can have the effect of providing negative contrast against surrounding tissue when imaging vasculature, and of making vasculature networks lying deep to the surface appear visible. Lindeboom et al. used orthogonal polarization spectral imaging to observe vascular patterns in patients

with oral SCC [76]. They found that in 80% of the cases from ten patients, capillary density was increased in tumorous areas and in 90% of the cases, disarrangement of capillary morphology correlated with tumorous areas.

This pilot trial was trial was designed to demonstrate the capabilities of the MDM as well as to make preliminary qualitative observations of abnormal lesions in humans. Further studies of larger groups of patients is required to determine the utility of each imaging technique compared to white light inspection, as well as to determine which combinations of techniques afford the most useful predictive information. Contributions from inflammation or benign conditions to DA or increased contrast in NB imaging will affect the specificity of detection and it will be important to determine these effects. Additionally, precise pathology correlations will be key to future studies. Correlations in this study were made using histopathologically-diagnosed biopsy or resected tissue data which was known to be within the FOV during image collection. Biopsy data provides information on only a small fraction of the imaged FOV and represents a limitation of this pilot study. Improved techniques for spatial correlation between histopathologic sections and image data will be needed in future work.

The MDM described here can obtain high quality images of the oral cavity using multiple imaging modalities. To facilitate comparison between modalities, images were collected in rapid sequences from the same field of view. Pilot clinical data indicate that additional imaging modalities such as NB and OPR may provide information not available in fluorescence mode alone, and may be useful in discriminating precancerous and cancerous tissue from normal and benign or inflammatory regions.

CHAPTER 4: OBJECTIVE DETECTION AND DELINEATION OF ORAL NEOPLASIA USING AUTOFLUORESCENCE IMAGING[2]

4.1 Introduction

Head and neck cancer, including cancers of the oral cavity, currently ranks as the sixth most common malignancy in the world. There were more than 270,000 new cases of oral cancer reported in 2002 [65]. Approximately 60% of these individuals present with stage III or IV disease, and about half will die within five years of diagnosis [7]. Screening individuals at risk for oral cancer and its precursors has the potential to improve early detection, providing the opportunity to intervene when treatment is most effective. In addition, surveillance of patients who have survived their initial oral cancer is important to identify local recurrences and second primary oral tumors, which occur at a higher rate than for any other tumor [82,83].

Conventional oral examination using incandescent white light is the current standard of care for screening and surveillance for oral cancer and precancerous lesions. The sensitivity of visual examination is limited by several factors including the experience and index of suspicion of the examiners. In primary care situations, cases of malignancy may be seen rarely and clinicians may have difficulty discriminating the sometimes subtle mucosal changes associated with premalignant lesions and early

[2] This chapter has been published in the following journal article: Darren Roblyer , C. Kurachi , V. Stepanek , M. Williams , A. El-Naggar , J. Jack Lee , A. Gillenwater, R. Richards-Kortum, "Objective Detection and Delineation of Oral Neoplasia Using Autofluorescence Imaging," Accepted for Publication in The Journal of Cancer Prevention Research, February 2008

cancers from more common benign inflammatory conditions [84]. Furthermore, it can be challenging to delineate the boundaries of neoplastic lesions using conventional oral examination making the choice of a biopsy location difficult.

Several new approaches have been proposed to address the limitations of the conventional oral examination, including the use of toluidine blue, brush cytology, reflectance visualization after acetic acid application, and illumination with a chemi-luminescent light source. While useful in certain situations, each of these approaches is associated with a high rate of false-positives [84-87]. Recently, several studies have demonstrated that autofluorescence imaging may improve the ability to distinguish normal from premalignant and malignant oral tissue [21,38,45,51,55,88,89]. When tissue is illuminated in the ultraviolet-visible region, a portion of photons are absorbed by molecules within the tissue called fluorophores which then emit lower energy photons that can be detected as fluorescence from the mucosal surface. Examples of fluorophores which produce autofluorescence signals in tissue include NADH and FAD in the epithelial layer, and collagen and elastin crosslinks in the stroma [19]. In comparison to normal oral tissue, neoplastic lesions are associated with a decrease of green fluorescence when excited with ultraviolet (UV) or near-UV light [14,51,55,88] that is attributed to decreased signal from collagen crosslinks in the stroma [6]. Increased red fluorescence has also been observed by several groups in oral lesions and is frequently attributed to porphyrins [90]. Several groups have proposed that this perceived loss of green fluorescence and increase in red fluorescence can be useful as a diagnostic aid to help detect and diagnose early neoplastic disease in several anatomic sites including the oral cavity, bronchus, cervix, esophagus and colon [21,91-94]. In addition, the changes in

fluorescence may aid in surgical resection by delineating the extent of neoplastic changes beyond the clinically apparent margins [55,88].

Recently, the U.S. Food and Drug Administration approved an autofluorescence imaging device for early detection of oral neoplasia. The device, marketed as the VELscope® (LED Dental, Inc., White Rock, BC, Canada), uses a blue/violet light (400 – 460 nm wavelengths) to illuminate oral tissue and long pass and notch filters to enable clinicians to directly visualize fluorescence in the oral cavity [21,88]. The VELscope and other proposed fluorescence imaging devices rely on qualitative observations to detect and delineate neoplastic oral lesions and therefore reliable screening with these instruments necessitates well-defined and standardized image interpretation criteria, and appropriate user training. This may not be feasible in many primary care situations. We hypothesize that the application of digital image processing techniques to autofluorescence imaging of oral tissue will provide the ability to objectively identify and delineate the peripheral extent of neoplastic lesions in the oral cavity. This will provide a powerful tool in patient care locations where experts are not available or where physicians encounter few cases of malignant and premalignant neoplasia. Low-cost digital cameras with sufficient sensitivity to record tissue autofluorescence in near real time are now readily available [95], making clinical application of such automated image processing feasible.

The primary goal of the present study was to evaluate the use of quantitative autofluorescence imaging for the detection and delineation of oral neoplastic lesions. We demonstrate that a simple, objective method can be used to accurately classify regions of interest within an autofluorescence image with 100% sensitivity and 91.4% specificity

relative to histopathology. This method can delineate the presence and extent of neoplastic lesions within a field of view and provide results which correlate with the histopathologic assessment of extent of disease. Thus, quantitative autofluorescence imaging may provide a non-invasive and objective method to improve screening and margin delineation of oral cancers and precancers.

4.2 Methods

4.2.1 Human Subjects

Study subjects were enrolled in a clinical protocol reviewed and approved by the Internal Review Boards at The University of Texas MD Anderson Cancer Center and Rice University. Patients were eligible and recruited if they were 18 years of age or older and had known or suspected precancerous or cancerous squamous lesions located in the oral mucosa. Patients may have had previous surgical, radiation, or chemotherapeutic treatments. Normal volunteers were eligible and recruited if they were 18 years of age or older and had no history of oral pathology. All subjects enrolled in the study gave written informed consent. The average age of patients in this study was 59. 42% of the patients were female and 58% were male. The average age of normal volunteers in this study was 27.4. 27% were female and 73% were male.

4.2.2 Imaging Procedure

Autofluorescence images were obtained from the oral cavity of 56 patients with clinically abnormal lesions and 11 normal volunteers. Data were divided into a training set and a validation set. Data acquired from the first 39 patients and 7 normal volunteers

imaged between June 2006 and January 2008 were allocated to the training set and used to develop an algorithm for detection of neoplasia. Data acquired from the subsequent 17 patients and 4 normal volunteers imaged between March and June 2008 formed a validation set and were used to test the performance of this algorithm relative to histopathology.

White light and autofluorescence images were obtained at 365 nm, 380 nm, 405 nm, and 450 nm excitation using a Multispectral Digital Microscope (MDM). This device is described in detail elsewhere [96] but briefly, the MDM is a wide-field optical microscope which collects digital autofluorescence and reflectance images with a color CCD camera from a variable field of view, ranging in size from approximately 1 to 7 cm. Patients were imaged either in an outpatient clinic or in the operating room under general anesthesia prior to surgery. A physician positioned the patient and microscope so that the suspicious lesion or area of interest was clearly in the field of view of the device. Clinically normal areas distant from or contralateral to the lesion were also imaged. Following imaging in the clinic, suspicious lesions were biopsied. In the operating room, previously biopsied lesions were surgically resected.

4.2.3 Histopathologic Correlation

Biopsies and resected tissues were evaluated using standard histopathologic analysis by a board certified pathologist (either AEN or MDW). The location of biopsies and resected lesions was recorded using digital photography so that pathology results could later be correlated to multispectral imaging results. In addition, the locations of gross anatomical features were noted in both autofluorescence images and

histology specimens to aid in correlation. The resulting histopathology sections were evaluated to provide a diagnosis along the entire length of the epithelium, also noting any submucosal abnormalities in each slide. Histopathology diagnosis included the following categories: normal, mild dysplasia, moderate dysplasia, severe dysplasia/carcinoma in situ, and invasive carcinoma. For the purposes of diagnostic algorithm development, two major categories were defined: normal tissue (including inflammation and hyperplasia) and neoplastic tissue (including dysplasia, carcinoma in situ and cancer).

4.2.4 Analysis and Statistical methods

Images were preprocessed to subtract signal from ambient room light and translated so that white light and fluorescence images of the same field of view were spatially registered. 276 measurements corresponding to 159 unique regions of interest (ROIs) sites of clinically normal and suspicious regions of tissue were selected from white light images by a head and neck surgeon (AMG) blinded to the results of the autofluorescence imaging. In some cases, repeat measurements were obtained from the same ROI site to help ensure image data was collected without motion artifacts; often both the first and repeat measurements were included in the analysis. These repeat measurements account for the difference between the number of measurements and the number of ROI sites. Heterogeneity in pathologic diagnoses may occur within relatively small areas of diseased oral mucosa [68,97] so ROIs were stringently selected from suspicious areas using one of following four criteria: 1) areas corresponding to the same size and location as a biopsy with a pathological diagnosis, 2) ROIs from locations which could be correlated to a histopathology slide with a corresponding pathological

diagnosis, 3) areas within well-defined exophytic tumors confirmed by pathological diagnosis and 4) ROIs from a location which was clinically normal and deemed by the physician to be sufficiently distant from the lesion.

Autofluorescence images from the training set were analyzed to determine whether specific image features could be used to classify a measurement site as normal or neoplastic. The autofluorescence images and white light images were spatially registered so that the ROIs chosen in the white light images corresponded to the same region of tissue in the autofluorescence images. The training set included data from the first 39 patients and 7 normal volunteers and included measurements from 173 measurements from 102 unique ROIs. Qualitatively, neoplastic ROIs were associated with a decrease in average green fluorescence intensity and often an increase in red fluorescence intensity. The mean ratio of red-to-green pixel intensities inside each of the ROIs was calculated from the fluorescence images at each excitation wavelength. Red and green pixel intensities were obtained from the collected Red-Green-Blue color images, created by the Bayer color mask on the CCD detector. A classifier was developed to distinguish neoplastic and normal ROIs using linear discriminant analysis with the single input feature of average ratio of red-to-green fluorescence. When more than one measurement corresponded to a ROI site, the mean of the feature values was used for classification. The classifier was trained using all of the ROI sites in training set and the prior probability input into the classifier was chosen to represent the percentage of abnormal to normal measurements in the data set. The classifier was developed after images were acquired from patients in the training set but before measurements were acquired from patients in the validation set. Classifier accuracy in the training set was assessed by

plotting the receiver operating characteristic (ROC) curve, the area under the ROC curve (AUC), and the sensitivity and specificity at a particular operating point on the ROC curve [59,63,98]. The positive and negative predictive values were also calculated at the operating point. Confidence intervals were calculated for operating characteristics using the Wilson 'score' method including a continuity correction [99].

The algorithm was then applied to data from the validation set using the red-to-green ratio threshold found to produce the highest combination of sensitivity and specificity in the training set. The validation set was designed to rigorously test the algorithm and for most patients, ROI and biopsy pairs were collected on the clinical margins of the lesion in addition to directly on the lesion and in clinically normal areas. The validation set included 103 measurements from 57 unique ROIs in a second group of 17 patients and 4 normal volunteers.

An additional analysis step was explored to increase the performance of the classifier by normalizing the red-to-green ratio measurements for each patient. An additional unique and non-overlapping ROI of clinically normal tissue was chosen from the same anatomical site and in the same field of view for each of the ROIs described above. At each excitation wavelength, the mean red-to-green autofluorescence ratio was calculated in this ROI; the mean red-to-green ratios from the other ROIs were normalized by this value. This method provides a way to compensate for anatomical and patient to patient variations in red-to-green fluorescence intensity ratio. Identical statistical analysis was performed using this measured feature with both the training set and the validation set. The method utilizing the magnitude of the red-to-green fluorescence intensity ratio is

termed the *raw red-to-green method* and the method utilizing a normalized red-to-green fluorescence intensity ratio is termed *normalized red-to-green method.*

4.2.5 Disease Probability Maps

The classification algorithms described above provided a relationship between the magnitude of the red-to-green fluorescence intensity ratio for a particular region of interest within the image and the probability of that region having a diagnosis of abnormal. This relationship was used to predict the probability of a diagnosis of dysplasia or cancer for each pixel in an image, given the red-to-green fluorescence intensity ratio at that pixel. The posterior probability values at each pixel in the image were computed and pixels which corresponded to a 50% or greater probability of being classified as dysplastic or cancerous were color coded and digitally overlaid onto the white light images. This method provides a means to illustrate areas of tissue with the highest probability of being neoplastic. The assumption was made that the region of interest method described above could be generalized on a pixel by pixel basis. Disease probability maps were compared to histologic images of tissue resected from the field of view to confirm the accuracy of this method.

Table 3. Anatomic sites of ROIs in the training and validation set. Note: Percentages may not add up to 100 % because of rounding.

Anatomical Site	No. of sites in training set (%)	No. of sites in validation set (%)
Tongue	37 (36.3)	19 (33.3)
Buccal mucosa	12 (11.8)	15 (26.3)
Floor of mouth	22 (21.6)	4 (7.0)
Gingiva	2 (2.0)	7 (12.3)
Lip	14 (13.7)	4 (7.0)
Palate	15 (14.7)	8 (14.0)
Total	102 (100)	57 (100)

Table 4. Pathology diagnosis of ROI sites in training and validation set.

Diagnosis	No. of sites in training set (%)	No. of sites in validation set (%)
Normal	53 (52.0)	35 (61.4)
Mild dysplasia	11 (10.8)	5 (8.8)
Moderate dysplasia	6 (5.9)	4 (7.0)
Severe dysplasia/CIS	12 (11.8)	6 (10.5)
Invasive carcinoma	20 (19.6)	7 (12.3)
Total	102 (100)	57 (100)

4.3 Results

Tables 3 and 4 summarize the anatomic site and histopathologic diagnoses of the 159 sites included in this analysis. The most common sites were tongue, buccal mucosa and floor of mouth, followed by palate, lip, and gingiva. The training set contained 52% normal, 28% dysplastic, and 20% invasive carcinoma sites while the validation set contained 61% normal, 26% dysplastic, and 12% invasive carcinoma sites. The normal histopathologic category could include tissue with hyperkeratosis, hyperplasia, and/or inflammation as long as there was no dysplasia or carcinoma. The normal sites in the training set, based on available pathology (not including normal volunteers and normal sites where no biopsy was taken), included 7 sites (13.2% of normal sites) with hyperplasia and hyperkeratosis, 4 sites (7.5% of normal sites) with hyperkeratosis, and 3 sites (5.7% of normal sites) with hyperplasia and/or fibroadipose tissue. The validation set included 3 sites (8.6% of normal sites) with hyperplasia and hyperkeratosis, 1 sites (2.9% of normal sites) with hyperplasia, 1 site (2.9% of normal sites) with a submucosal hemorrhage, and 1 site (2.9% of normal sites) with marked inflammation and osteonecrosis. The abnormal histopathology category could include dysplasia and carcinoma. In the training set 59.2% of the abnormal sites were premalignant (mild, moderate, or severe dysplasia), in the validation set 68.2% of the abnormal sites were premalignant.

Figure 8 shows white light and autofluorescence images from the buccal mucosa of a patient with pathologically confirmed invasive carcinoma. The white light image (Fig. 1A) shows two ROIs, one which corresponds to a pathologically confirmed invasive carcinoma, and the other which was clinically normal and outside of the pathologically

confirmed clear resection margin. Figures 8B-8D show autofluorescence images at different excitation wavelengths that were taken before surgery from the same field of view. The autofluorescence image obtained at 405 nm excitation qualitatively shows the greatest visual contrast between the normal and neoplastic ROI. This observation was typical for study patients.

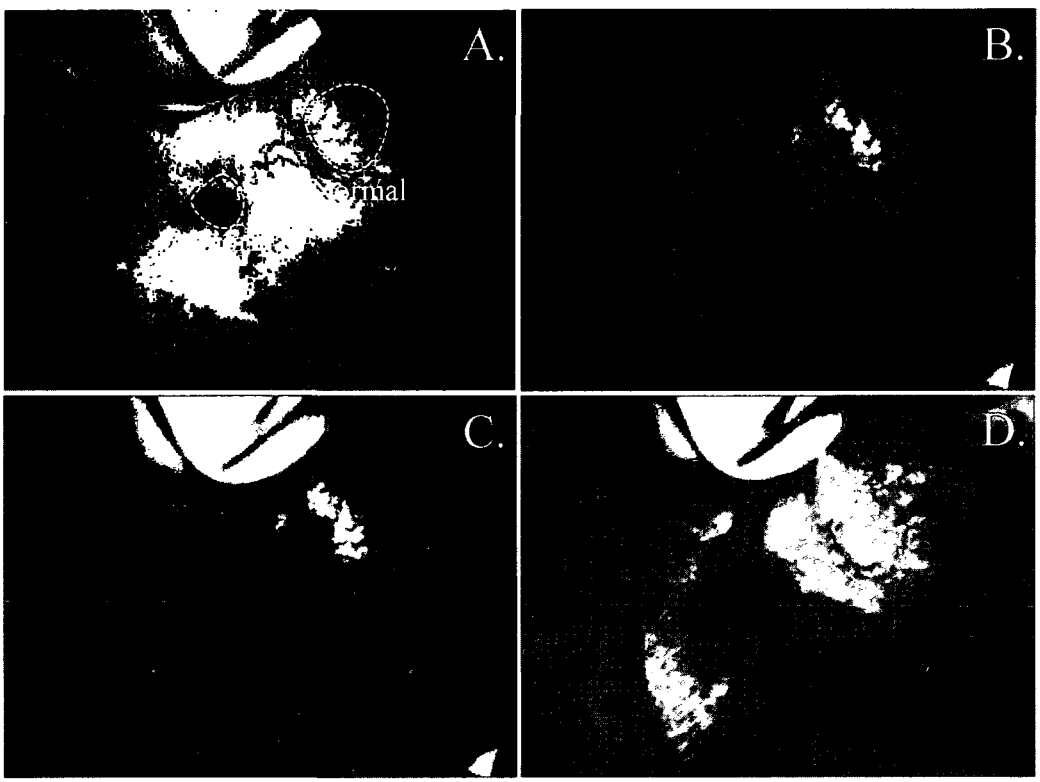


Figure 8. Autofluorescence and white light images of the buccal mucosa of a typical study patient. A. White light image showing regions of interest of histopathologically confirmed normal tissue and invasive carcinoma. B. Fluorescence image at 365 nm excitation. C. Fluorescence image at 405 nm excitation. D. Fluorescence image at 450 nm excitation.

Table 5 summarizes the performance of both diagnostic algorithms, based on either the raw or the normalized mean red to green fluorescence intensity ratios, for classifying lesions in the training set. At each excitation wavelength, the classifier that used the normalized red-to-green fluorescence intensity ratio (Normalized R/G ratio) had slightly higher AUC than the algorithm based on the raw red/green fluorescence intensity ratio (Raw R/G ratio). In all cases, the highest AUC was obtained at 405 nm excitation. The sensitivity and specificity values at the point on the ROC curve nearest the gold standard (Q-point) are also reported in Table 5.

Table 5. Classification results at each fluorescence excitation wavelength using both the Raw R/G Ratio method and the Normalized R/G ratio method in the training set.

Fluorescence excitation wavelength	Raw R/G ratio			Normalized R/G ratio		
	AUC	Sensitivity (%)	Specificity (%)	AUC	Sensitivity (%)	Specificity (%)
365 nm	.799	91.1	60.7	.858	79.5	86.9
380 nm	.871	75.9	83.6	.902	84.0	88.5
405 nm	.966	89.3	95.1	.982	93.8	95.1
450 nm	.887	83.9	82.0	.897	84.0	90.2

A scatter plot of the normalized red-to-green ratio at 405 nm excitation for each of the 102 sites in the training set, as well as the threshold of 1.19 used in the classification algorithm is shown in Figure 9A. Of the 102 sites, 4 were misclassified including one site of fibroadipose tissue on the lower lip misclassified at abnormal, one hyperkeratotic site on the right buccal misclassified at abnormal, one cancer site on the right lateral tongue misclassified as normal, and one site on the left soft palate with focal ulceration

and dysplasia misclassified as normal. Figure 9B shows the ROC curve for this classifier; the AUC is 0.988, and at the Q-point, the sensitivity is 95.9% (95% confidence interval (CI) 84.9% - 99.3%) and the specificity is 96.2% (95% CI 85.9% - 99.3%). The positive predictive value is 95.9% (95% CI 84.9% - 99.3%) and the negative predictive value is 96.2% (95% CI 85.9% - 99.3%). This operating point is indicated on the ROC curve.

The algorithm using the normalized red-to-green fluorescence intensity ratio at 405 nm excitation was applied to the validation set. In Figure 9C a scatterplot of the normalized R/G ratio for each site in the validation set is shown along with the threshold that had been previously selected for the training set. Figure 9D depicts the ROC curve with the operating point selected for the training set indicated. A 100% sensitivity (95% CI 81.5% - 99.6%) and 91.4% specificity (95% CI 75.8% - 97.8%) and an AUC of .987 were achieved at this operating point for the validation set. The positive predictive value is 88.0% (95% CI 67.7% - 96.9%) and the negative predictive value is 100% (95% CI 86.7% - 99.7%). Of the 57 sites in the validation set, 3 were misclassified as abnormal including one site on the left buccal with hyperplasia, one site on the right buccal, and another site on the left buccal.

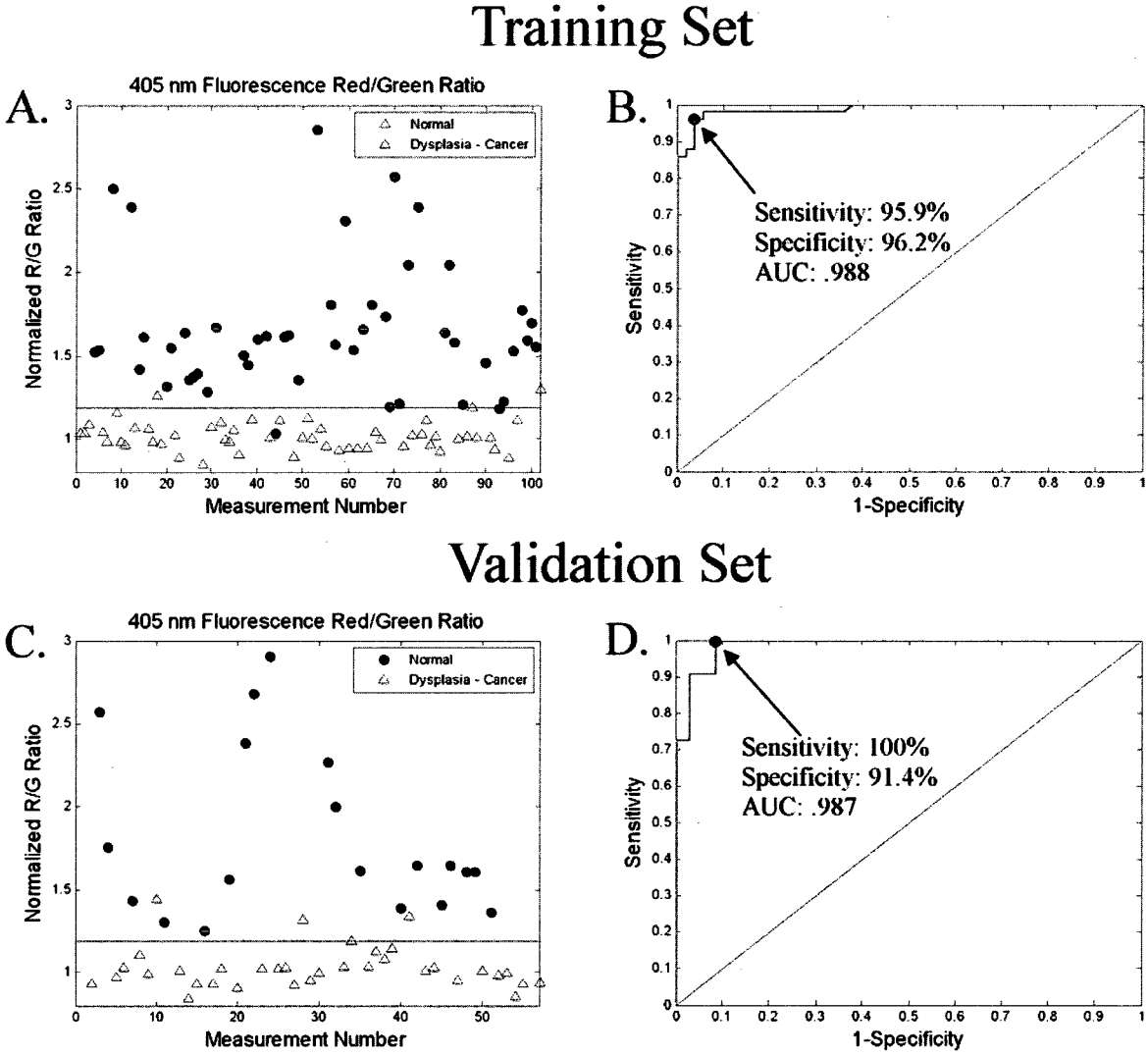


Figure 9. A. Scatter plot of normalized red-to-green ratios at 405 nm excitation for the 102 ROI sites in the training set. The horizontal line indicates the threshold used to obtain 95.9% sensitivity and 96.2% specificity. Note that 2 additional abnormal data points had a red-to-green fluorescence intensity ratio greater than 3 but are not shown on this plot. B. Receiver-operating characteristic (ROC) curve of the classifier based on the normalized red-to-green ratio. The operating point used for classification is indicated by a dot and

arrow. C. Scatter plot of the red-to-green ratio for the 57 sites in the validation set with threshold selected from the training set indicated. Note that 3 additional abnormal data points had a red-to-green fluorescence intensity ratio greater than 3 but are not shown on this plot. D. ROC curve obtained for the validation set. The operating point is indicated and corresponds to the threshold chosen from the training set.

Figure 10 shows white light and 405 nm excited autofluorescence images from a study patient with moderate dysplasia and carcinoma in situ located in the floor of mouth. The white light image is also shown with an overlay of the calculated disease probability map; regions corresponding to a predictive probability of a neoplastic lesion greater than 50% are shaded as indicated by the color bar. The disease probability map indicates the probability that a particular pixel in the image corresponds to a neoplastic area of tissue. Histologic sections obtained at six areas in the tissue are also shown. Only one of these areas was included in the previous classification analysis. The disease probability map shows qualitative agreement with the presence of dysplasia and cancer in the areas corresponding to the histologic sections.

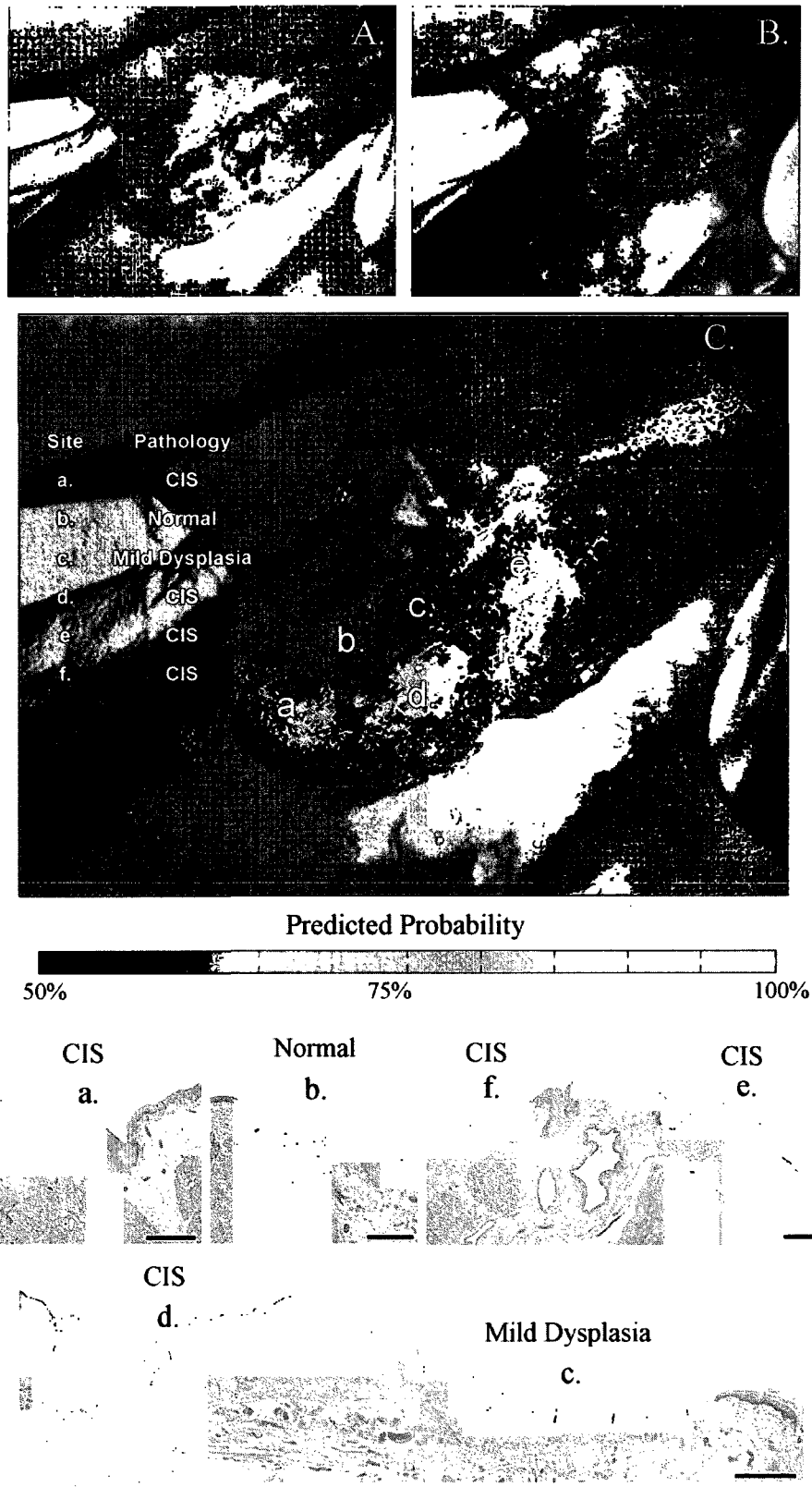

A.
B.
C.
Site
Pathology
a.
CIS
b.
Normal
c.
Mild Dysplasia
d.
CIS
e.
CIS
f.
CIS
Predicted Probability
50%
75%
100%
CIS
a.
Normal
b.
CIS
f.
CIS
e.
CIS
d.
Mild Dysplasia
c.

Figure 10. A. White light image of floor of mouth with histopathologically confirmed dysplasia and carcinoma in situ. B. 405 nm excitation fluorescence image showing areas with deceased autofluorescence. C. White light image with disease probability map showing the predictive probability of a neoplastic lesion superimposed. Letters indicate specific locations were pathology is known. The key to the right of C. indicates pathology. The histology slides below show tissue sections from these areas. Marking bar at the lower right-hand corner = 1 mm.

Figure 11 shows representative white light images with and without superimposed disease probability maps from four study patients. Images in the first three rows correspond to patients with histologically confirmed neoplasia, while the image in the bottom row is from a normal volunteer with no clinically suspicious lesions. Although the lesion in Figure 11A is obvious, those in Figures B and C are less so, highlighting the potential to aid clinicians in identifying the presence of neoplasia and identifying optimal sites for further evaluation with biopsy. Images in Figures 11A and B are from a patient with an invasive carcinoma in the floor of mouth. Images in Figures 11C and D are from a patient with a region of severe dysplasia on the tongue. The images in Figures 11E and F are from a patient with a region of moderate dysplasia on the gingiva. In all three cases, the disease probability map delineates the suspicious regions identified clinically by an oral cancer specialist blinded to the results of the autofluorescence imaging and are consistent with histopathologic sections obtained. Figures 11G and H are from the inner lip of a normal volunteer and the disease probability map does not indicate any lesions.

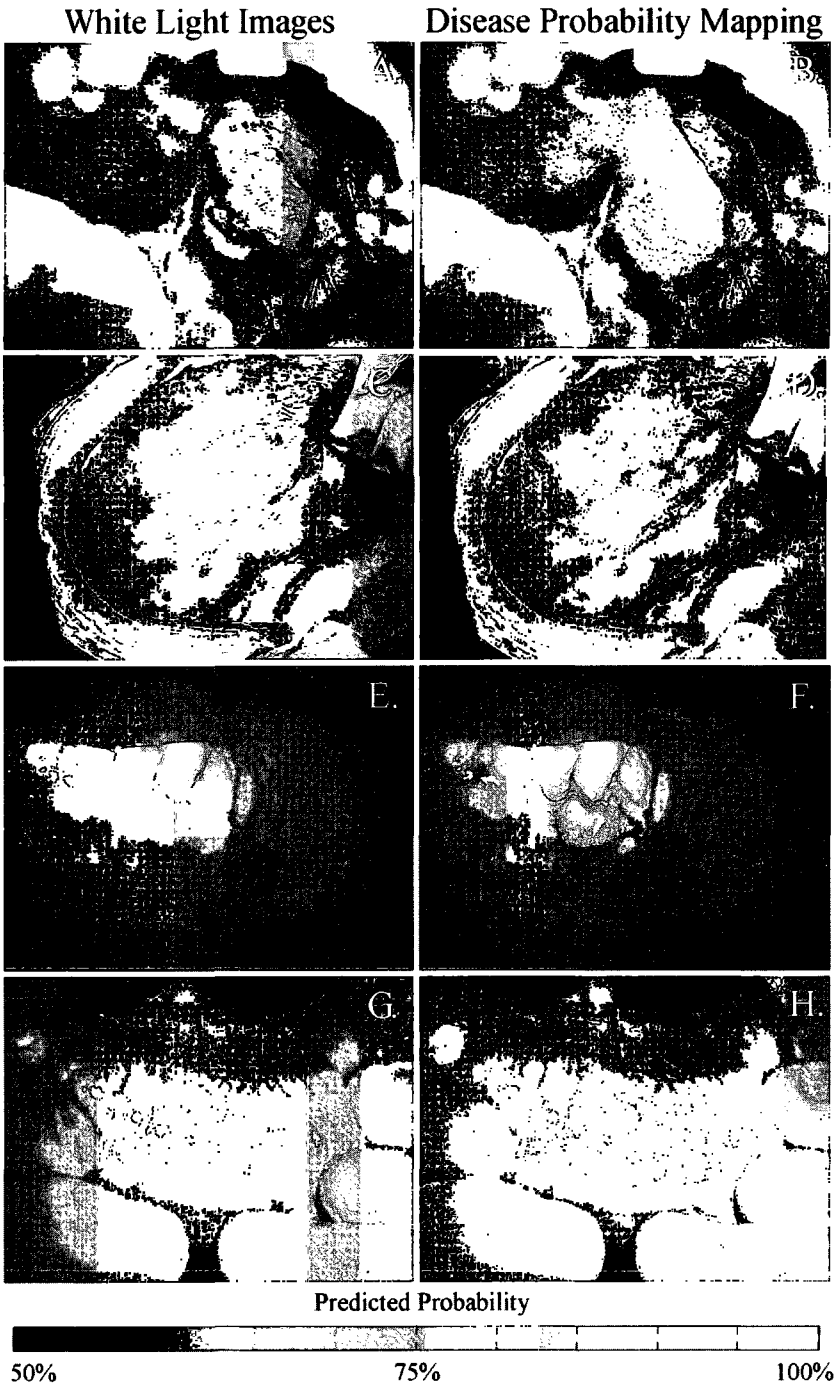


Figure 11. A. and B. show images from a patient with an invasive carcinoma in the floor of mouth. A. White light image B. White light image with disease probability mapping showing the predictive probability of a neoplastic lesion. C. and D. show images from a

patient with a region of severe dysplasia on the tongue. E and F show images from a patient with a region of moderate dysplasia on the gingiva. G. and H. show images from inner lip of a normal volunteer.

4.4 Discussion

Our results illustrate how autofluorescence imaging may enhance the ability of clinicians to detect and delineate areas of oral dysplasia and carcinoma. Although all four illumination conditions tested allowed visualization of changes in autofluorescence with neoplasia, illumination with 405 nm wavelength produced the highest discriminatory capability. This corresponds to previous findings comparing illumination wavelengths for autofluorescence imaging in freshly resected oral cancer surgical specimens [14]. While subjective interpretation of loss of autofluorescence has been shown to be useful [21,88], there are several important advantages associated with objective and quantitative analysis of changes in autofluorescence signal. First, quantitative analysis methods provide a rigorous and repeatable way to determine the threshold for demarcating a lesion, even for providers with less experience. Second, digital imaging allows the operator to save and process images, directly comparing data from multiple patients in a series or from a single patient over time. Third, ratios of fluorescence intensity values provide a way to reduce variations in images associated with spatial non-uniformities in illumination.

In the present study, the performance of a simple classifier based on the ratio of red-to-green autofluorescence intensity at 405 nm excitation was tested and found to

discriminate neoplastic and non-neoplastic tissue with a sensitivity and specificity of 96% in the training set and 100% sensitivity and 91.4% specificity in the validation set. These results compare favorably with the performance of visual oral examination, which has been systematically reviewed by Downer et al [22]. Downer identified eight prospective studies between 1980 and 2002 that involved conventional oral exam with gold standard verification provided by an expert observer. In four of the studies the screeners were general dentists and in four of the studies the screeners were trained health workers. Sensitivity ranged from 59% to 97%, specificity ranged from 75% to 99%, and meta-analysis resulted in a weighted pooled sensitivity of 85% and a specificity of 97%. Other reports of the performance of visual oral screening include Sankaranarayanan et al (sensitivity 77%, specificity 76%) [23], Ramadas et al (sensitivity 82%, specificity 85%) [24], and Nagao et al (sensitivity 92%, specificity 64%) [25]. The classifier in this study can be applied to entire images of the oral cavity to visualize areas with a high probability of being neoplastic; disease probability maps are consistent with histologic sections obtained from tissue in the field of view.

Autofluorescence imaging has shown great promise for enhancing visualization of neoplastic areas in recent studies [14,21,38,45,55,88] . In a study of 44 patients, Lane et al. demonstrated high sensitivity and specificity at discriminating normal oral mucosa from severe dysplasia, carcinoma in situ, or invasive carcinoma based on visual assessment of loss of autofluorescence in diseased mucosa at excitation wavelengths between 400 nm and 460 nm [21]. In another study by the same group, the potential for autofluorescence imaging to enhance delineation of the margins of neoplastic changes was demonstrated. In some cases fluorescence loss extended as far as 25 mm beyond the

clinically apparent margin [88]. Autofluorescence endoscopic imaging technologies for lung and the GI tract utilizing ratios of red and green signal have been available for over a decade and have greatly increased sensitivity of disease detection in these organ systems [41,91,100,101]. The LIFE^{TM} (laser induced fluorescence emission) system is an autofluorescence bronchoscopy device which provides the user with a real time image in which changes in hue correspond to suspicious and/or abnormal areas. Users of the device must be trained in order to interpret these changes in image hue [100].

A potential confounding factor which may limit specificity of classifiers based on the red-to-green fluorescence intensity ratio for automated image analysis software is the frequent presence of red fluorescence on normal papillae of the dorsal aspect of the tongue. At 405 nm excitation, increased fluorescence above 600 nm emission has been observed in oral lesions and is thought to originate from porphyrins, although it is uncertain whether the origins of these porphyrins are intrinsic or derived from bacterial contamination [70]. Red porphyrin-like fluorescence has been shown to correlate with neoplastic disease in other regions of the oral cavity [55,70]. However, normal red fluorescence is limited to the dorsal tongue; it is not observed on normal areas on the lateral and ventral tongue where there is a much higher propensity for developing neoplastic disease [102]. In this study, a single site in the training set was imaged on the dorsal tongue and corresponded to cancer. Bright red fluorescence was visible at this site.

Our results demonstrate the potential of quantitative fluorescence imaging as an objective approach to non-invasively identify and delineate the mucosal extent of neoplastic lesions in the oral cavity. It should be noted that the images were obtained

with a research-grade device and the disease probability maps described here were constructed subsequent to the image acquisition and compared to only a limited number of sites with diagnosis confirmed with histopathology. Unfortunately, it is difficult to obtain pathology diagnosis for the entire surface of a resected specimen. Additionally, in order to provide optimal benefit to clinicians both for detection and margin delineation of oral dysplasia and carcinoma, these disease probability maps need to be available to view in real or near real-time. We are currently making software improvements and developing quantitative fluorescence imaging devices that can show false color disease-probability maps based on red/green fluorescence intensity ratios at 405 nm excitation at the time of the examination. In addition, we recognize that our results were obtained from a small group of subjects with disease prevalence that might be expected in a treatment population rather than a screening population. Future studies are planned to evaluate this method of quantitative fluorescence imaging in community settings in a larger subject group with a wide spectrum of oral pathologies including oral dysplasia and early invasive disease, benign conditions and inflammation.

In summary, the present study provides proof-of–principle for the use of a practical tool for the quantitative fluorescence imaging as an objective technique for detection and delineation of oral neoplasia. The use of objective disease probability maps represents an important advance toward integration of optical imaging technologies into the clinical practice of dentists and primary health care workers. Development of non-invasive and objective diagnostic aids based on these findings may facilitate early detection and diagnosis of oral cancer and its precursors by less experienced personnel at the point of care.

CHAPTER 5: COMPUTER AIDED DIAGNOSIS OF ORAL CANCER AND PRECANCER USING MULTISPECTRAL WIDEFIELD OPTICAL IMAGING[3]

5.1 Introduction

There is an important need to improve the ability to identify and delineate the presence and extent of precancerous oral lesions and cancers so that they can be adequately treated [7,68,82,83]. Multispectral widefield imaging is emerging as an attractive, non-invasive means to visualize the mucosal surface at risk and identify early oral cancer and precancers. Direct autofluorescence visualization has been shown to reveal biochemical changes associated with oral precancer and cancer [6,21,38,39,42,44,51,87,89]. Reflectance imaging, including narrowband illumination imaging and polarized imaging, has been shown to aid in visualizing vasculature in the oral cavity [74,76] which increases during malignant progression [73]. It is important to identify the imaging modalities, illumination and collection conditions, and/or combinations of modalities which provide the most useful diagnostic information. This information is valuable both for direct visualization of the tissue with multispectral devices and for computer aided diagnosis of lesions based on multispectral digital images.

---

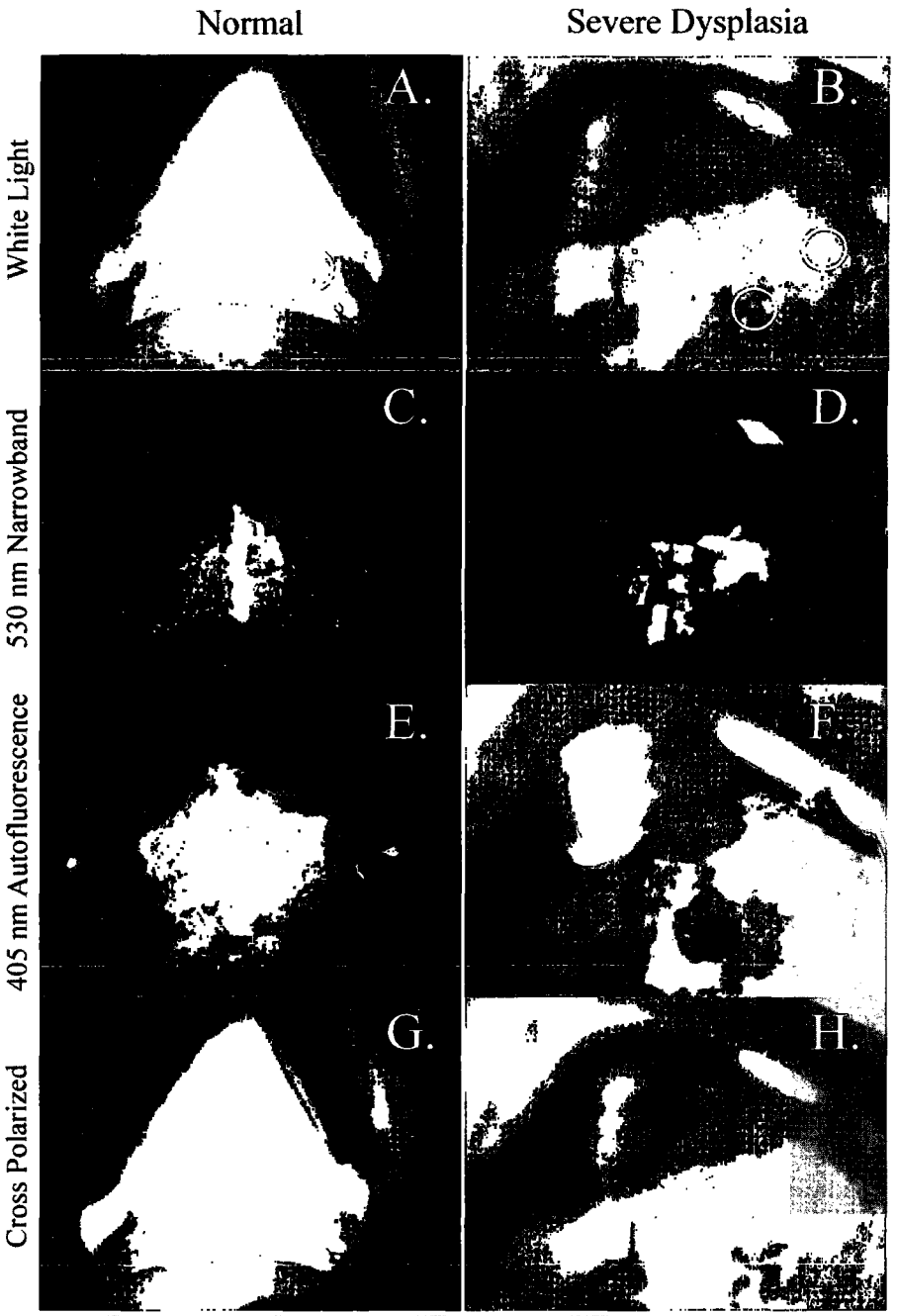
[3] The content for this chapter is being prepared for publication for the journal IEEE Transaction on Medical Imaging. The following authors have contributed: Darren Roblyer, Cristina Kurachi, Vanda Stepanek , Richard A. Schwarz, Michelle D. Williams, Adel K. El-Naggar, J. Jack Lee, Ann M. Gillenwater, Rebecca Richards-Kortum.

An important goal of multispectral optical imaging is to increase the visual contrast between normal and neoplastic tissue beyond that available using conventional white-light screening techniques. Several studies have shown that illuminating tissue with blue light and viewing the autofluorescence from the tissue enhances the ability to visually identify neoplastic lesions based on loss of autofluorescence [14,21,55,88]. While these results are encouraging, to date, only qualitative observations of image contrast between normal and neoplastic oral lesions have been reported; quantitative assessment across multiple modalities is needed to optimize performance.

Computer aided diagnostics (CAD) have been successfully implemented to aid automated diagnosis in a variety of organ systems including the prostate [103], pancreas [104], cervix [56], liver [105], bladder [106], skin [107], and ovary [108]. CAD techniques have been previously applied to analyze white light images of the oral cavity to discriminate oral lichenoid reactions and luekoplakia and to segment these lesions based on color and morphologic features [109,110]. Our group has previously demonstrated that the normalized ratio of red-to-green autofluorescence intensity at 405 nm excitation is useful to objectively discriminate normal oral tissue and dysplastic and cancerous lesions [111].

The goals of this study were twofold. The first was to quantify and compare optical image contrast available using autofluorescence, narrowband reflectance, and cross polarized reflectance imaging modalities for neoplastic lesions in the oral cavity. Results of this analysis can help guide the choice of imaging modalities and illumination/collection wavelengths for visual examination to detect neoplastic oral lesions. The second goal was to develop and evaluate computer aided diagnostic

algorithms to detect neoplastic lesions from analysis of multispectral images of the oral cavity. Results will aid in determining the modalities and features which contain the most diagnostic information and will inform the design and construction of new imaging devices for objective diagnosis and delineation of neoplastic oral lesions.



Feature	Normal	Severe Dysplasia
White light Michaelson Contrast	.048	.13
White light Red to Green Ratio	1.01	1.08
NB 530 nm Michaelson Contrast	.052	.091
405 nm Michaelson Contrast	.017	.37
405 nm Red to Green Ratio	1.02	1.23
Cross Michaelson Contrast	.026	.095

Figure 12. A., C., E., and G. are white light, narrowband 530 nm reflectance, 405 nm excitation autofluorescence, and cross polarized images of the palate of a normal volunteer. Approximate ROIs are shown. B., D., F., and H are images from the palate of a patient with severe dysplasia. ROIs of the biopsy location and a corresponding clinically normal area are shown. Several feature values from these images, computed from the indicated ROIs are shown in the table.

5.2 Methods

5.2.1 Instrumentation and Data Acquisition

Widefield images were collected in four different modalities from patients and normal volunteers using a Multispectral Digital Microscope (MDM) [96]. The MDM consists of a commercially available surgical microscope modified to capture digital images in white-light reflectance, autofluorescence, narrowband reflectance, and cross polarized imaging modalities. Autofluorescence images were acquired at four different excitation wavelengths: 365 nm, 380 nm, 405 nm, and 450 nm. Narrowband reflectance images were acquired using illumination wavelengths of 430 nm and 530 nm. In a subset of patients, narrowband reflectance images were also acquired at 575 nm illumination. White-light reflectance and cross polarized white-light reflectance images were also captured. Cross polarized polarization reflectance imaging is achieved by illuminating the tissue with linearly-polarized light and collecting remitted light through a second linear polarizer, oriented orthogonal to the illumination polarization. In total, there were

9 image types which could be collected by the MDM, accounting for the different excitation or illumination wavelengths in each modality.

The MDM captures images of tissue with a rectangular field-of-view of approximately 5 X 7 centimeters. The CCD camera utilizes a Bayer mask to collect color images. Images are collected as 12-bit RGB tiff files. Camera integration times and gain settings were standardized for each subject so that quantitative intrapatient image analysis could be performed.

In this study, the MDM was used to acquire images from patients with pathologically confirmed oral lesions and normal volunteers with no history of oral lesions, under a protocol reviewed and approved by the Institutional Review Boards at Rice University and the University of Texas MD Anderson Cancer Center. Each measurement sequence consisted of a serial collection of the aforementioned image types. One or more measurements were taken from each study participant, at the lesion site and, when possible, at contralateral or distal clinically normal sites. To ensure that high quality measurements were obtained, measurements of sites were usually repeated one time.

For purposes of algorithm development and evaluation, data were divided into a training set and a validation set. Data from the training set were collected prior to the validation set.

5.2.2 Preprocessing

Images were collected *in vivo* and there was often some subject movement between image capture of different image types. To account for this, image registration

was performed for each measurement sequence. An affine transformation was used to translate all of the images taken in a sequence to the white-light *base* image using up to 8 common reference points chosen manually.

Regions of Interest (ROIs) corresponding to normal and neoplastic areas were chosen from the white-light images by an expert physician (AG) blinded to the other image types. These ROIs were chosen to correspond to one of the following three categories: (1) a histopathologically confirmed lesion, (2) a histopathologically confirmed non-neoplastic region, or (3) a clinically determined normal region either from a normal volunteer or determined by expert physician (AG) to be sufficiently distal to a lesion in a patient. ROIs were the approximate size of a 2 mm circular biopsy if biopsy was used for histopathology and for clinically determined normal areas, and were the surgical lesion size if surgical tissue resection occurred. For each patient, an additional clinically normal ROI was selected; image data from these ROIs were used to explore whether normalizing data from the aforementioned ROIs to account for interpatient variability could increase image contrast and/or diagnostic performance. These normalization ROIs were selected either from the contralateral anatomic site or from an area which was sufficiently distal from any lesion on the same anatomic site.

5.2.3 Optical Image Contrast

The optical image contrast between histopathologically determined lesions and normal tissue was calculated and compared for each image type. The contrast was computed from ROIs chosen from lesions with a pathologic diagnosis of dysplasia or

cancer relative to the corresponding clinically normal ROI. Grayscale versions of the collected images were used to compute the contrast.

Several methods have been proposed to compute image contrast [112,113]; in this study we explored four different contrast metrics. *Simple* contrast was calculated as the ratio of the mean grayscale pixel values from the abnormal and normal ROIs. *Difference* contrast was calculated as the difference between the mean grayscale pixel value from the abnormal and normal ROIs. We calculated the *Weber* contrast as the mean grayscale pixel value from the abnormal ROI divided by the sum of the mean grayscale pixel values from the abnormal and normal ROI. *Michaelson* contrast was calculated as the difference divided by the sum of the mean grayscale pixel values from the abnormal and normal ROI.

The four contrast metrics were computed from dysplastic and cancerous lesion sites. The contrast metrics for the autofluorescence, narrowband reflectance, and cross polarized reflectance image types are reported relative to the same contrast metric for the white-light image obtained from that site. For each image type, the percentage of lesions with a higher optical contrast than white-light was computed. One-way ANOVA was used to compare the mean contrast metrics for each image type to that for white light.

In order to explore the relationship between optical contrast and pathologic grade, a scatter plot of the contrast metrics versus pathologic diagnosis was made for image types with the greatest image contrast. Unbalanced one-way ANOVA was used to compare the statistical difference in mean optical contrast between diagnostic categories.

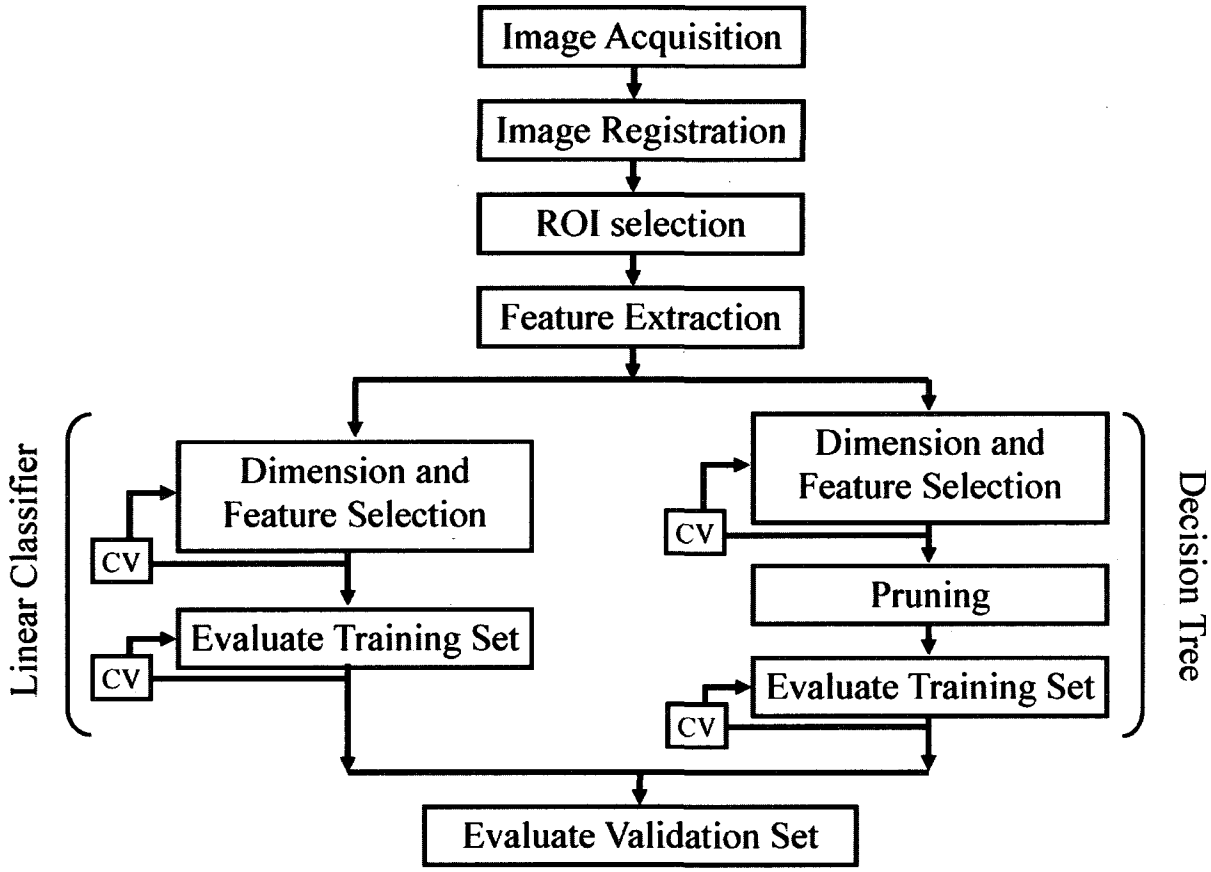


Figure 13. Flowchart of computer aided diagnostics procedure. CV denotes cross validation.

5.2.4 Computer Aided Diagnosis

Figure 13 provides an overview of the algorithm development and evaluation procedure used in this study. After image acquisition, registration, and ROI selection, image features were extracted from each ROI. In cases where image acquisition was repeated from the same tissue location, feature values extracted from the repeat images were averaged and considered as a single measurement site. Image features were designed to quantify statistical measures, texture, and frequency content

In order to determine the image types and modalities or combination of modalities most capable of discriminating neoplastic lesions, algorithms were developed and evaluated using features extracted from five groups, including: (1) Features obtained from white-light reflectance images; (2) Features obtained from narrowband reflectance images at 430 nm and 530 nm illumination; (3) Features obtained from cross polarized reflectance images; (4) Features obtained from autofluorescence images at 365 nm, 380 nm, 405 nm, and 450 nm excitation; and (5) Features obtained from all of the modalities and image types.

Two different supervised classifiers (a linear classifier and a decision tree classifier) were then used to construct diagnostic algorithms based on features from images in the training set. For each classifier, features were chosen using a feature selection algorithm. The results from the two methods were compared to the gold standard of histopathology and the algorithm with the highest performance was then used to classify data from the validation set.

Algorithms were designed first to classify the ROIs into one of two diagnostic categories: non-neoplastic and neoplastic. The neoplastic class included lesions diagnosed histopathologically as mild, moderate, or severe dysplasia, carcinoma in situ, or invasive carcinoma. The non-neoplastic class included clinically normal sites in patients and normal volunteers. We then attempted the more difficult problem of designing classifiers to assign ROIs into one of three diagnostic categories: non-neoplastic, dysplasia, and cancer.

5.2.5 Feature Extraction:

For each ROI site, 98 image features were computed for each image type, as described in detail below.

The Michaelson and Weber contrast were computed from the grayscale images and from each color channel of the RGB images, resulting in 8 features for each image type.

18 first-order statistical features were calculated using the grayscale pixel values from the ROIs, including the mean, standard deviation, entropy, variance, skewness, and kurtosis. These features were calculated for each ROI and as normalized by features from the normalization ROIs selected from clinically normal areas. Normalization was performed in two ways; the first was calculated as the difference between the ROIs (difference-normalization), and the second was the ratio of the ROIs (ratio-normalization).

18 features were obtained using the color channels of measurements. The mean values of the red, green, and blue channels of each ROI were used as features. The ratio of the mean red-to-green, red-to-blue, and green-to-blue pixel values were also utilized. Both normalized and non-normalized feature values were calculated.

Features representing texture in the images were obtained by using grayscale-level co-occurrence matrices (GLCM). GLCMs are useful for quantifying how pixel intensities vary spatially. A pixel separation, d, and angle, θ, are specified for a particular GLCM. The size of the GLCM is determined by the number of discrete intensity values contained in the grayscale image. Each entry *(i,j)* in the GLCM is a count of the number of times a pixel of intensity *i* occurred at a distance d and angle θ away from a pixel with

intensity j. Statistical measures including contrast, correlation, energy, and homogeneity were computed from the GLCMs. More detail is provided in Argenti et al. [114].

24 features were created based on these statistical measures from GLCMs where d varied from 1 to 6. The features were averaged at angles $\theta = 0^{\circ}$, 45°, 90°, and 135° to account for the fact that these multispectral images do not have a specific spatial orientation.

A 2-D DFT was performed on a rectangular region whose center corresponded to the approximate center location of the selected ROIs. This 2-D DFT was converted into a 1-D plot of frequency content by integrating the pixel intensities at discrete radii from the origin. The 1-D plot was then partitioned into 10 frequency ranges and the frequency content was integrated inside each range. The contribution of each partition was calculated by dividing by the total integrated 1-D plot so that the sum for all 10 partitions added to unity.

30 features were computed using the relative frequency content for the partitions. Normalized and non-normalized feature values were included. Variations of this method have been used by Gossage et al. and Srivastava et al. [57,108]

5.2.6 Linear Classifier

We implemented a linear classifier (LC) based on Bayesian parameter estimation. This method assumes multivariate normal densities and equal covariance for each class. The LC is trained on a data set which is used to estimate the mean μ_i for each class and a pooled covariance matrix Σ for all classes. *A priori* probabilities are determined from the relative proportion of each class in the training set; posterior probabilities are output for

each measurement and used by a linear discriminant function to separate the measurements into classes.

5.2.7 *Decision Tree Classifier*

We utilized a decision tree classification method based on the widely used Classification and Regression Tree (CART) induction technique. This method has the attractive attribute of classifying data without the need to make assumptions about the underlying statistical distributions of the observations [62]. We used the *Gini impurity* to determine splits [59,62]. To help avoid overtraining the decision tree in the training set it was pruned to find the smallest tree at which adding further nodes does not statistically decrease the cost of the tree. The cost of the tree is defined in the *zero-one* sense, where the cost of misclassifying an observation is 1 and the cost of correctly classifying an observation is 0.

5.2.8 *Feature Dimension and Selection*

A forward sequential search (FSS) algorithm based on the wrapper model was used on the training set to determine the optimal feature dimension: the minimum number of features needed to maximize a chosen classifier performance criterion without overtraining [59,63]. Starting with one feature, classification of the training set was performed with 5-fold cross-validation using the FSS algorithm to find the single feature which maximized the criterion value. The area under the curve (AUC) of the receiver-operating characteristic (ROC) was used as the criteria for the FSS algorithm for the LC. For the decision tree, the sum of the sensitivity and specificity was used. This was

repeated with additional features until classifier performance did not increase. This entire procedure was repeated 25 times to provide statistically significant results. One-way ANOVA with multiple comparison tests were used to determine the optimal feature dimension. The final features sets for each classifier were determined by the most commonly chosen features by the FSS algorithm in the 25-iterations. This entire procedure was repeated for each of the five imaging modality-based feature subsets.

5.2.9 Classification Performance

The classifier performance for the training set was determined using 5-fold cross-validation. For 2-class classification, we utilized sensitivity and specificity as the figures-of-merit (FOM) to evaluate and compare the performance of the classifiers on the training set for each of the feature subsets. For 3-class classification we utilized the FOM of total correct classification rates and correct classification rates for each class. Based on these results, the best performing classifier for the 2-class and 3-class problem was retrained on the entire training set and then applied without change to evaluate the validation set.

5.3 Results

5.3.1 Data Collection

In total, images were acquired from 72 subjects, including 61 patients with pathologically confirmed oral lesions and 11 normal volunteers. From these images, we defined 350 ROI measurements sites in the study. There were 93 non-neoplastic ROIs, and 82 neoplastic ROIs each with a corresponding normalization ROI. The neoplastic

ROIs consisted of 22 ROIs with mild dysplasia, 13 ROIs with moderate dysplasia, 16 ROIs with severe dysplasia or carcinoma in situ, and 31 ROIs with invasive cancer.

The ROIs were divided into training and validation sets. Not including normalization ROIs, there were 102 ROIs from 46 patients in the training set and 73 ROIs from 26 patients in the validation set. The training set was collected on dates before the validation set. No data was repeated in both sets.

5.3.2 Optical Contrast

Figure 12 illustrates selection of ROIs from multispectral images from the palate of a normal volunteer and the palate of a patient with severe dysplasia. White-light, 530 nm narrowband reflectance, 405 nm autofluorescence, and white-light cross polarized reflectance images are shown. ROI locations are indicated by circles in the white-light images. The image in 1E shows red fluorescence on the dorsal tongue of the normal volunteer. This was a common observation on the dorsal tongue of both normal volunteers and patients. A chart of example features and contrast metric values, which will be described below, are shown. In this example, all contrast and feature values are increased in the patient with severe dysplasia.

The results obtained using different definitions of image were similar (data not shown); thus, we report only results from a single contrast metric. Figure 14A shows a box plot of the Michaelson contrast metric by image type; contrast for each image type is reported relative to the contrast achieved using white-light illumination. For each image type, Fig. 3A also indicates the percentage of lesions where contrast was greater in that image type than in white light. On average, the contrast for each of the image types is

greater than that available in white light mode. Autofluorescence imaging at 405 nm showed the greatest average increase in contrast and the greatest percentage of abnormal lesions with increased contrast relative to white-light imaging. The mean contrast value for autofluorescence at 365 nm, 380 nm, 405 nm, and 450 nm excitation was found to be statistically different than that of white light imaging, using one-way balanced ANOVA with a 95% confidence interval.

A subset of 26 patients was imaged with narrowband illumination at 575 nm. From this subset there were 41 lesion sites. The 575 nm narrowband imaging provided an increase in Michaelson contrast over white light in 80% of these lesions. The median increase was 1.15, which was similar to the other narrowband illuminations but was significantly less than that achieved from the autofluorescence modality.

Autofluorescence images at 405 nm excitation show the greatest increase in optical contrast compared to white light images when discriminating non-neoplastic from neoplastic tissue. In order to explore whether optical contrast in this image type increased with increasing grade of dysplasia or cancer, we plotted the contrast of each lesion by diagnostic category (Fig. 3b). The means for the diagnostic categories are shown in the shaded boxes. The mean contrast of normal tissue is a significantly lower than that of dysplasia and cancer, when calculated using one-way unbalanced ANOVA with 95% confidence interval. However, the mean contrast values for each of the dysplastic categories and cancer were not statistically different from each other.

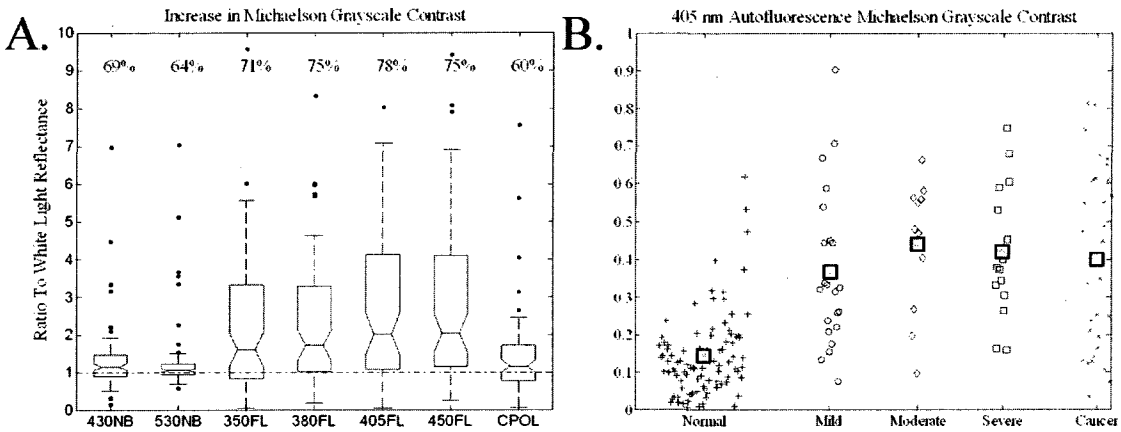


Figure 14. A shows a boxplot of the increase in Michaelson contrast compared to white light when using narrowband reflectance imaging, autofluorescence imaging, and cross polarized reflectance imaging. The imaging modality is indicated on the x-axis. For each modality, the percentage of lesions where the contrast was increased over white light is indicated near the top of the plots. The three horizontal lines on each of the boxes represent the lower quartile, median, and upper quartile of the data from bottom to top. The whiskers extending from the box indicate the rest of the data except for outliers, which are indicated with the "." symbol. 2B shows the Michaelson grayscale contrast computed from the 405 nm autofluorescence images. The contrast values are displayed by graded diagnostic category. The shaded squares in each diagnostic cluster indicate the location of the mean contrast value. There is a statistically significant difference in contrast between normal and the other diagnostic categories, but not between the grades of dysplasia or the carcinoma.

5.3.3 Computer Aided Diagnosis

5.3.3.1 2-class classification

Table 6 lists the number of features selected and summarizes the classifier performance for both the LC and decision tree methods using 5-fold cross validation for the training set. For both classifiers, the best performance was obtained using the autofluorescence feature subset. Furthermore, the only feature chosen by the FSS in the combined feature subset was a single feature from the autofluorescence modality subset.

Features extracted from white-light images provided the second best performance after autofluorescence for both of the classifiers on the training set; classification required five features for the LC and 3 features for the decision tree. Narrowband reflectance provided the third best classification with 4 features. Cross-polarized provided the worst classification with 6 features for the LC and 3 features for the decision tree.

Figure 15 shows ROC curves produced from the LC on the training set using 5-fold cross-validation and the same features used above. The autofluorescence feature subset produced the highest AUC followed by white-light, narrowband, and cross polarized features.

The autofluorescence feature chosen for the LC was the ratio of red-to-green intensity (difference-normalized). For the training set, this single feature provided an AUC of .981 and a sensitivity of 93.9% and a specificity of 98.1%. The feature chosen from the decision tree was very similar, the ratio of ratio of red-to-green intensity (ratio-normalized). For the training set, this single feature provided a sensitivity of 95.9% and a specificity of 92.5%.

The LC was used on the validation set because it provided a slightly higher sum of sensitivity and specificity than the decision tree. Using these same features, the algorithm was retrained on the entire training set and applied to the validation set yielding an AUC of .949, a sensitivity of 100%, a specificity of 85.0%, a PPV of 84.6%, and a NPV of 100%.

Table 6. Two-class classification results of the training data for the linear classifier (LC) and the decision tree classifiers. The area under the curve (AUC) of the receiver operating characteristic curve, sensitivity, and specificity are shown for the LC. Sensitivity and specificity are indicated at the q-point on the ROC curve. The sensitivity and specificity are shown for the decision tree classifier. The number of features chosen is indicated for each feature subset.

Feature Subset	LDA				Decision Tree		
	Num of Feats.	AUC	Se (%)	Sp (%)	Num of Feats.	Se (%)	Sp (%)
White Light	5	.931	87.8	88.7	3	87.8	84.9
Narrowband	4	.847	81.6	75.7	4	83.7	83.0
Cross Polarized	6	.898	83.7	75.5	3	85.7	77.4
Autofluorescence	1	.981	93.9	98.1	1	95.9	92.5
Combined	1	.981	93.9	98.1	1	95.9	92.5

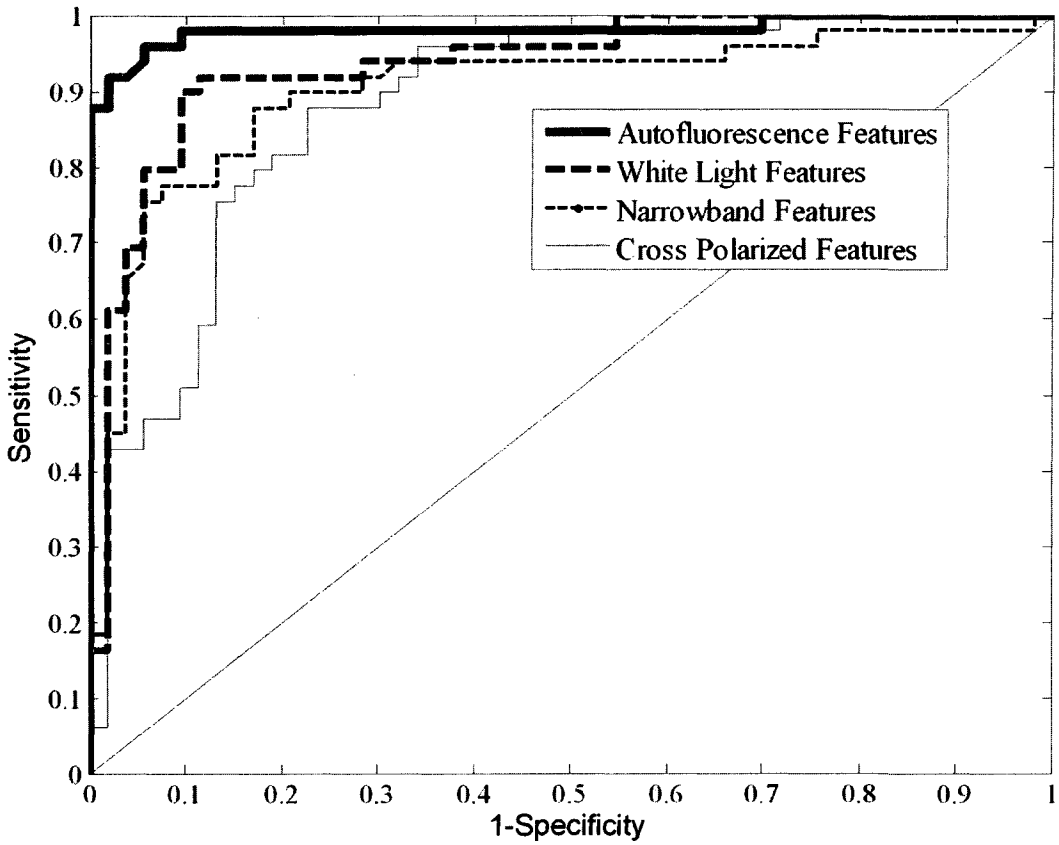


Figure 15. ROC curves from the 2-class LC for different feature subsets on the training set. When considering the AUC of the ROC plots, the autofluorescence features performed the best, followed by white light, narrowband, and cross polarized. Note that these ROC plots are from a single iteration of LC classification on the training set and therefore the parameters may not match exactly to table 6.

5.3.3.2 3-class classification

Table 7 lists the number of features chosen for the three-class LC and the decision tree classifier using 5-fold cross-validation. Algorithm performance is also summarized in Table 7, as the percent of sites correctly classified for all sites and for each of the three diagnostic categories.

The best performance from both classifiers was from the autofluorescence feature subset using the decision tree; in the training set this classifier correctly classified 94.3% of the non-neoplastic ROIs, 75.9% of the dysplasia ROIs, and 80% of the cancer ROIs. The first feature chosen was the ratio of red-to-green intensity (ratio normalized) from the 405 nm autofluorescence image; this was the same feature chosen for the 2-class problem by the decision tree. The second feature was the GLCM homogeneity with d=6 calculated from the 380 nm autofluorescence image. For both classifiers in the training set, autofluorescence features and white-light features were chosen in the combined feature subset, but this combination did not improve performance, based on overall correct classification rate, over the autofluorescence only feature subset.

The decision tree classifier was then retrained on the entire training set and applied to the validation set. Results are summarized in Table 8. In the training set, this method misclassifies 5.7% (3 of 53) non-neoplastic ROIs as dysplasia, 17.2% (5 of 29) dysplasia ROIs as cancer, 3.4% (1 of 29) dysplasia ROIs as non-neoplastic, and 20% (4 of 20) cancer ROIs as dysplasia. No cancerous ROIs were misclassified as non-neoplastic and no non-neoplastic ROIs were misclassified as cancer.

In the validation set, 20% (8 of 40) non-neoplastic ROIs were misclassified as dysplasia and 81.8% (9 of 11) cancerous ROIs were misclassified at dysplasia. 100% (22

of 22) dysplastic ROIs were correctly classified. No cancerous ROIs were misclassified as non-neoplastic and no non-neoplastic ROIs were misclassified as cancer.

Table 7. Three-class classification results of training set for LC and decision tree classifier. All results are computed using 5-fold cross-validation.

Feature Subset	Linear Classifier					Decision Tree				
	Num of Feats.	Total Correct (%)	Normal Correct (%)	Dys. Correct (%)	Cancer Correct (%)	Num of Feats.	Total Correct (%)	Normal Correct (%)	Dys. Correct (%)	Cancer Correct (%)
White Light	4	81.4	92.5	65.5	75.0	2	72.6	81.1	75.9	45.0
Narrowband	5	65.6	88.7	37.9	45.0	2	72.6	79.3	69.0	60.0
Cross Polarized	5	61.8	90.6	34.5	25.0	3	69.6	84.9	58.6	45.0
Autofluorescence	4	84.3	98.1	79.3	55.0	2	86.3	94.3	75.9	80.0
Combined	4	77.5	94.3	69.0	45.0	2	83.3	96.2	79.3	55.0

Table 8. Confusion matrices of 3-class classification results from the decision tree using two autofluorescence features. Numeric values indicate the number of classified measurement sites. Results from both the training set and validation set are shown without cross-validation.

Decision Tree Results		Training Set			Validation Set		
		Truth			Truth		
		Normal	Dysplasia	Cancer	Normal	Dysplasia	Cancer
Predicted	Normal	50	1	0	32	0	0
	Dysplasia	3	23	4	8	22	9
	Cancer	0	5	16	0	0	2

5.4 Discussion

In this study multispectral widefield imaging data from 72 study subjects were analyzed to determine the imaging modalities, illumination/detection conditions, and image features which provide the greatest optical image contrast between oral lesions and surrounding non-neoplastic tissue, and to determine if objective classification algorithms could be used to diagnostically classify oral lesions.

78% of the lesions imaged showed greater contrast when imaged in autofluorescence mode at 405 nm illumination that in white light mode. These lesions included all grades of dysplasia and carcinoma. When the contrast was analyzed by diagnostic category, contrast was significantly greater for all pathologies compared to non-neoplastic areas but contrast was not significantly different for images of lesions with different pathologic grades. This suggests the contrast observed using autofluorescence may be due to changes which occur early in malignant transformation. Narrowband images and cross-polarized images showed an increase in contrast over white light in the majority of images but not to the extent of the autofluorescence images.

2-class classification performance using both the LC and decision trees was excellent in the training set, providing sensitivities and specificities exceeding 90% for both classifiers. It is clear from the analysis that color-based features extracted from autofluorescence images at 405 nm were highly dominant, demonstrated by the fact that a single feature of this type was chosen using both classifiers, even with access to the entire feature set. The combination of features from multiple subsets did not improve 2-class classifier performance above that from the autofluorescence feature subset alone. The LC was applied to data from the validation set, yielding 100% sensitivity and 85%

specificity, helping to confirm the diagnostic capability of the 405 nm autofluorescence image type.

For the 3-class problem, the decision tree modestly outperformed the LC in the training set. Two features based on autofluorescence were selected after pruning in the two-step decision tree. The first feature, the ratio of red-to-green intensity (ratio-normalized) at 405 nm excitation, was used to classify non-neoplastic sites from dysplastic and carcinoma sites, just as in the 2-class problem. The second feature was then used to separate the dysplastic lesions from the carcinoma lesions. This second feature was produced by texture analysis and was from the 380 nm excitation autofluorescence image. In the validation set, 0% (0 of 22) of the dysplastic lesions were misclassified but 81.8% (9 of 11) of the carcinoma lesions were misclassified as dysplasia. This suggests that the second feature may not generalize well to new data to discriminate dysplasia from cancer.

These data suggest that it may not be feasible for widefield multispectral optical imaging to discriminate different abnormal diagnostic categories using the specific modalities and wavelengths tested here. Based on results from Pavlova et al., it may be necessary to selectively probe the epithelial in order to obtain a more specific diagnosis using optical techniques [6,115]. Results from both the contrast analysis and classification portions of this study indicate that dysplastic and cancerous oral lesions can be discriminated from normal tissue using the autofluorescence imaging modality at 405 nm excitation. It was shown that the reflectance white light, narrowband, and cross-polarized images included in this analysis did not improve the diagnostic ability available from the autofluorescence images. This is somewhat surprising given that the signal from autofluorescence and

reflectance modalities are likely dominated by two different biological phenomenon during malignant progression: the breakdown of the supporting tissue structure in the stroma [6], and the increase in vascularity in and around lesions [73] respectively.

Several groups have previously demonstrated that increased contrast of lesions observed using autofluorescence imaging with blue excitation light (400 nm – 460 nm) is an effective diagnostic means for high-grade oral lesions [14,51,55]. Poh et al. reported that an observed visual decrease in autofluorescence signal often extended beyond clinical tumor margins up to 25mm [88]. Lane et al. showed a sensitivity of 98% and a specificity of 100% for the discrimination of severe dysplasia and carcinoma from direct autofluorescence visualization [21]. This work has achieved comparable diagnostic performance using an objective discrimination technique while including low-grade as well as high-grade lesions.

Multispectral widefield optical imaging has the potential to improve upon current oral cancer screening and delineation methods. The use of 405 nm excitation autofluorescence imaging of the oral mucosa can increase the optical image contrast of lesions above that observed using white light and can provide objective classification of neoplastic lesions with high accuracy.

CHAPTER 6: WIDEFIELD MULTISPECTRAL IMAGING AND HIGH RESOLUTION MICROENDOSCOPE IMAGING FOR DETECTION OF ORAL NEOPLASIA[4]

6.1 Introduction

Squamous carcinoma of the head and neck is associated with poor five-year survival, usually a consequence of late-stage disease at diagnosis [1]. Treatment of oral cancer is associated with severe morbidity; surgical removal of oral tissue often results in disfigurement and impairs speech and the ability to swallow. Diagnosis of head and neck cancers at an earlier stage can improve reduce morbidity and improve efficacy of therapy, leading to improved patient survival and quality of life.

The current standard-of-care to screen for oral cancer is white light inspection of the oral cavity by an expert clinician, followed by biopsy and histologic analysis of suspicious lesions. Subjective visual examination requires adequately trained personnel, as well as the availability of histopathology labs to analyze biopsy specimens of suspicious lesions. Even in developed nations, lesions are commonly not discovered until a late stage because of the relative lack of screening.

Optical methods offer a means to improve screening for oral neoplasia at the point of care without the need for extensive infrastructure. Autofluorescence visualization is a promising tool to detect early neoplastic changes in the oral epithelium. The Velscope is

[4] The content for this chapter is being prepared for publication. Both I and Timothy Muldoon have contributed equally to the work presented in this chapter. The following authors have also contributed: Kelsey Rosbach, Vanda Stepanek, Michelle D. Williams, Ann M. Gillenwater, Rebecca Richards-Kortum.

a commercial system that uses blue light to excite autofluorescence in the oral mucosa; a clinician can then visually inspect the mucosa and identify suspicious areas based on loss of autofluorescence [21,80]. It has been demonstrated that precancer and cancer are associated with loss of stromal collagen fluorescence [6].

Computer-aided tools and quantitative classification algorithms can be used to further improve the screening process and provide an objective diagnosis. Digital imaging of the autofluorescence signal from the oral mucosa allows quantifiable image parameters to be calculated; the red-to-green pixel intensity ratio has been of particular importance in distinguishing between neoplastic and non-neoplastic sites. A recent study by our group using this technique achieved a sensitivity of 100% and specificity of 91.4% in an independent validation data set [111].

A potential limitation of widefield imaging and analysis of the bulk tissue autofluorescence signal is that this modality does not directly sample some of important pathologic features, such as changes in cellular and nuclear morphology. High-resolution imaging may be used to complement the diagnostic capability of widefield imaging. High-resolution imaging can be used to directly observe changes in epithelial cell morphology and epithelial architecture over small fields-of-view. These changes include increased nuclear-to-cytoplasmic ratios, pleomorphism, and alterations in the gradation of cellular differentiation throughout the epithelium. A variety of high resolution optical imaging modalities are available to sample these morphologic and architectural changes, including confocal micrscopy, optical coherence microscopy and high resolution microendoscopy [116-118].

Multimodal imaging approaches which combine images from wide field and high-resolution modalities have the potential to improve early detection of neoplastic changes. Wide field imaging allows for rapid inspection of nearly the entire oral cavity, while high-resolution imaging can be used to interrogate suspicious regions with higher spatial resolution.

The goal of this study was to explore the complementary diagnostic ability of widefield autofluorescence imaging ad high resolution microendoscopy for early detection of oral neoplasia. Widefield autofluorescence images were obtained from resected oral tumors; high resolution images were then obtained following topical application of a fluorescent contrast agent to visualize changes in nuclear morphometric and density. The ability to classify oral tissue as normal or neoplastic was explored using data from widefield imaging along, high resolution imaging alone and the combination of the two modalities. Multimodal approaches yielded the greatest sensitivity and specificity relative to the gold standard of histopathology, supporting the complementary role of these two imaging approaches.

6.2 Methods

6.2.1 Data Acquisition and Instrumentation

Data was collected under a clinical protocol approved by the Institutional Review Boards at the University of Texas MD Anderson Cancer Center and Rice University. All participating subjects gave written informed consent. Patients were eligible if they were 18 years old or older and were scheduled for surgical resection of a known precancerous or cancerous lesion in the oral cavity.

Following resection, tissue was imaged in the laboratory using first a widefield multispectral digital microscope (MDM) and then a high resolution microendoscope (HRME) system. The MDM is described in detail elsewhere [96]. Briefly, the MDM system is a modified clinical microscope which collects digital images with an approximate 5 cm by 7 cm field of view (FOV). The system is capable of collecting data in several different imaging modalities but for this study only white light reflectance and autofluorescence images at 405 nm excitation were utilized. The entire mucosal surface of the resected tissue was imaged using the MDM.

The HRME system is described in detail elsewhere [119]. Briefly, the system can image cellular detail of the upper epithelium by placing a fiber optic image guide guide into direct contact with the tissue. Contrast is achieved by applying a topical solution of proflavine (0.01% in water) with a cotton swab immediately before imaging. Proflavine is a fluorescent agent that binds to DNA, yielding bright nuclei surrounded by dark cytoplasm. The HRME system uses an LED with a center wavelength of 455 nm as an excitation source; this is directed through an objective lens and coupled to the fiber optic image guide, which delivers it to the tissue surface. Fluorescence emission light travels back though the image guide, is magnified by the objective lens, and is imaged to a CCD camera, where it is digitally stored and sent to a PC. The HRME system has a circular field of view with a diameter of 750 microns, and collects image data at 4 frames per second.

Both clinically normal and suspicious areas were measured with both the MDM system and the HRME system. Several suspected non-neoplastic and neoplastic sites were imaged on each resected specimen. HRME images were qualitatively assessed by

the authors (TJM and DMR) for the presence of nuclear detail. Only HRME images which revealed nuclear detail were included in this study. For each interrogated tissue site, one or more HRME images were collected inside a small, approximately 2 millimeter diameter area.

Histopathology diagnosis was obtained from the sites imaged with the MDM and HRME. The locations of regions imaged with the HRME were recorded using digital photography, and these locations were tracked throughout histopathology processing. The resected tissue was sectioned and stained according to standard histopathology practice and diagnosis was obtained from prepared slides by an expert Head and Neck pathologist (M.D.W.).

6.2.2 Comput*er Aided Diagnostics*

Image data obtained with both the MDM and HRME were used to develop objective classifiers to identify neoplastic tissue. Images from each device were first processed and used to train a linear discriminant classifier based on a single imaging modality. MDM and HRME results were then combined using multiple hybrid consensus techniques to train multimodal classifiers. Tissue sites imaged with both the MDM and HRME and with a corresponding pathologic diagnosis were included in this analysis.

For analysis of MDM data, circular regions of interest (ROIs) with a 20 pixel diameter, which correspond to an approximate 1mm diameter circular region of tissue, were selected from MDM images at sites imaged by the HRME. In addition, a normalization ROI was identified in each specimen from a pathologically normal area,

distant from any lesion. Features obtained from the pixel values inside these regions were then extracted. Previous *in vivo* studies have shown that the mean ratio of red to green autofluorescence at 405 nm excitation, normalized to a known non-neoplastic region, provides a high degree of discrimination between non-neoplastic and neoplastic tissue [96,120]. This feature was extracted from the ROIs of MDM images collected in this study; these values were normalized by the mean red to green autofluorescence ratio from the normalization ROIs.

This feature was used as input to a 2-class linear classifier based on Bayesian parameter estimation in order to obtain a predicted diagnosis of either non-neoplastic or neoplastic. The neoplastic class includes ROIs with diagnoses of mild, moderate, and severe dysplasia as well as carcinoma in situ and invasive carcinoma. The non-neoplastic class included epithelial tissue which was histopathologically normal with or without hyperplasia, hyperkeratosis, or inflammation. Leave-one- patient-out cross-validation was used avoid overtraining the algorithm [63]. A posterior probability was obtained for each site which is the predicted probability that the site is in the neoplastic class. A receiver-operating-characteristic (ROC) curve was generated by changing the threshold of the posterior probability value used to determine if a measurement was in the first or second class, and then plotting 1-specificity versus sensitivity at each threshold value or operating point. The area under the curve (AUC) of the ROC plot was used as a figure-of-merit (FOM) to determine classification performance. Sensitivity, specificity, positive predictive value, and negative predictive value were also determined at the operating point which maximized the sum of sensitivity and specificity and reported.

A variety of features were extracted from the HRME images at each site and used as input to a linear classifier. Regions of interest were manually selected from each HRME image, corresponding to regions subjectively showing the most nuclear detail. In addition, this ROI selection was done to eliminate the boundaries of the images, which did not contain any information. Features were extracted from these rectangular ROIs. In the cases where there was more than one HRME image collected from a site, the mean of these feature values was used for classification. First order statistical features such as the mean, standard deviation, and entropy of pixel values were extracted from the ROIs. Textural features obtained from gray-level co-occurrence matrices with pixel offsets from 1 to 10 pixels were also obtained. Spatial frequency content features were extracted by performing a 2-D Fourier transform on the image data inside the ROIs. Further detail about these features is provided in [120].

In order to select the most diagnostically valuable features from the HRME data, a forward sequential feature selection algorithm was implemented [63]. This algorithm tested each feature separately and then in combination using the linear classifier to determine the two features which classified the data most accurately based on maximal area under the ROC curve. Two features were chosen because adding additional feature did not statistically improve classification performance. The HRME data were classified using the linear classifier with the two selected features using leave-one-out cross-validation. Again, a posterior probability was obtained for each site An ROC plot, AUC, sensitivity, specificity, positive predictive value, and negative predictive value were obtained.

6.2.3 Combinatorial Methods

Two separate strategies were implemented to combine data from both imaging modalities to yield a single diagnosis. The first approach used was to develop a classifier using input from both imaging modalities to obtain a single diagnosis. We refer to this technique as the Nondirected Combinatorial Technique. The second strategy, referred to as the Directed Combinatorial Technique, used data from the MDM to first identify potentially neoplastic sites. At these sites, data from both the MDM and HRME modalities were used to classify the site as neoplastic or non-neoplastic. The Directed Combinatorial Technique was implemented by first using data from the MDM to train and test a linear classifier using input from all of the measurement sites. The operating point for this classifier was selected such that a sensitivity of 90% was achieved. Sites classified as non-neoplastic by this algorithm were not analyzed further. Sites classified as potentially neoplastic by this algorithm were analyzed further. A second algorithm was developed using both MDM and HRME data from these sites.

For both the Directed and Nondirected Combinatrorial Techniques, we explored eleven different methods to combine the results from two devices. Eight of these methods rely on combining the posterior probabilities obtained from each device for each site to produce a new metric which can be used to discriminate the classes. Many of these methods are described by Bendiktsson et al. [121]. The other three methods explored involve utilizing one or more classifiers in a different way, as described below.

The posterior probabilities (pp) per site were combined from the two devices in the following eight ways:

1. The maximum of the two pp's was used,

2. The minimum of the two pp's was used,
3. The mean of the two pp's was used,
4. The weighted sum of the two pp's was used where the weights were determined heuristically by the relative performance of each device. We have used the AUC obtained from each device as the FOM to determine the weights. This method is called Linear Opinion Pool (LOP).
5. The weighted sum of the two pp's (LOP) was used where the weights were determined by a search of different weights. The weights of each device were between 0 and 1 and the weight from the second device w_2, was equal to 1- w_1, where w_1 is the weight for the first device. Weights were tested in .01 increments.
6. The product of the two pp's was used. This method is called Logarithmic Opinion Pool (LOGP).
7. The weighted product of the two pp's (LOGP) was used where the weights were determined by a search. The weights were applied as exponents for the pp's.
8. The weighted sum of the two pp's (LOP) was used where weights were determined both per device and per class. This method is described in reference[121].

In addition to these methods, three other methods were used to synthesize the data. The pp from the MDM was used as the prior probability for the HRME linear classifier; extracted features from both devices were used in a single linear classifier; and the pp from the MDM was used as an additional feature for the HRME linear classifier.

6.3 Results

In total, resected specimens were obtained from 14 patients. Six of the resected specimens primarily involved the tongue, three primarily involved the floor of mouth, two primarily involved the buccal mucosa, one primarily involved the retromolar trigone, one involved the tongue and floor of mouth and one involved the gingiva and floor of mouth. From these specimens, 69 unique measurement sites were matched to a corresponding pathology diagnosis. These sites included 26 sites with a diagnosis of non-neoplastic, 3 sites with mild dysplasia, 4 sites with moderate dysplasia, 13 sites with severe dysplasia or carcinoma in situ, and 23 sites with invasive squamous cell carcinoma. When divided into two classes: the non-neoplastic class contains 26 sites and the neoplastic class contains 43 sites.

A ROI from the MDM autofluorescence images was selected at the each of the 69 sites. A normalization ROI was chosen for each patient and corresponded to a histopathologically normal area. 126 HRME ROIs were obtained from the 69 sites; when multiple ROIs were available, the features were averaged together.

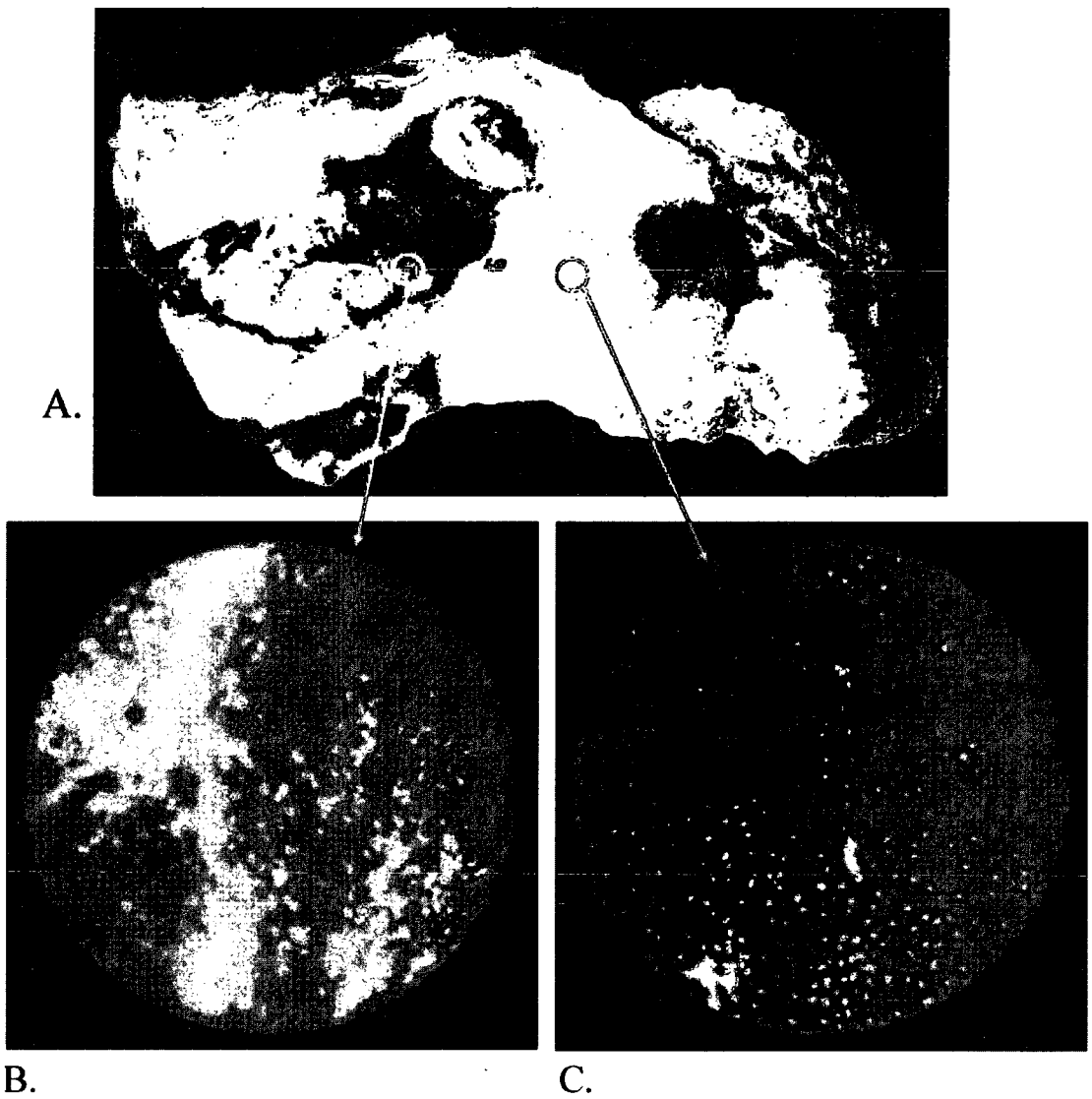
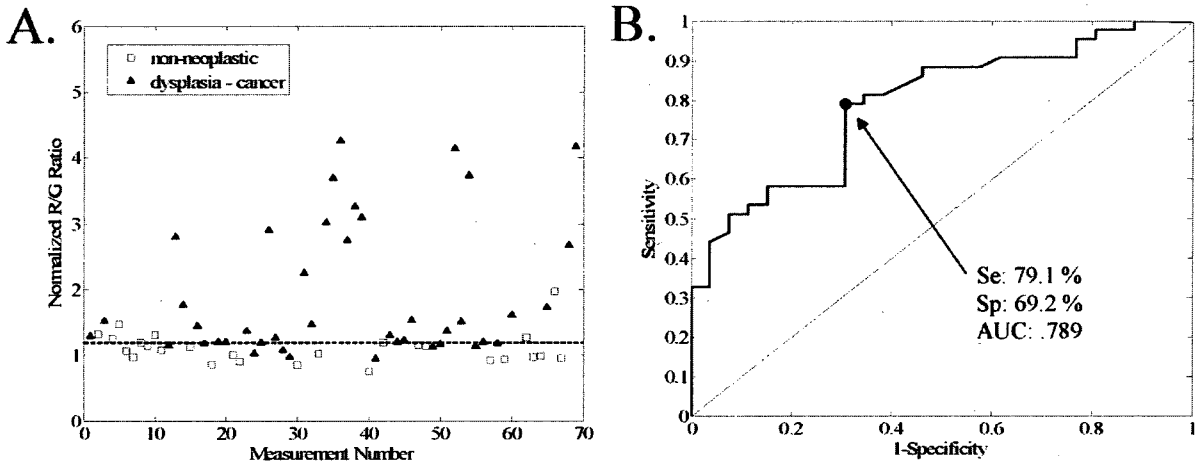


Figure 16. A. MDM Autofluorescence image at 405 nm excitation from resected tissue from the floor of mouth. MDM ROIs from a squamous cell carcinoma (SCC) site and a non-neopalstic site are indicated. B. HRME image collected from the SCC site. C. HRME image collected from the non-neoplastic site.

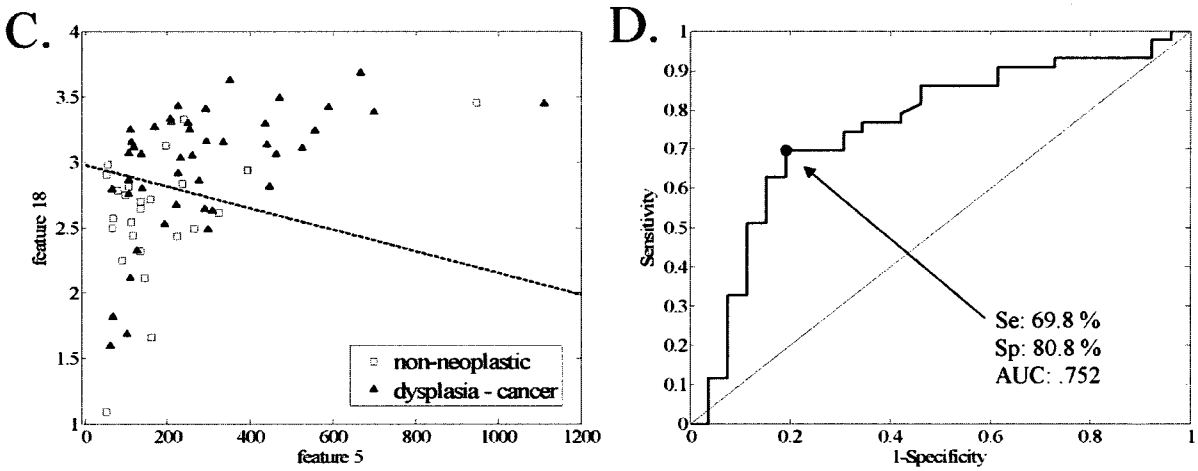
Table 9. This combinatorial methods and the AUC achieved. Both devices were used to measure all measurement sites. The percent increase in AUC values over each device alone is shown. Method 5 produced the highest AUC.

Combination Method	AUC	% increase over MDM	% increase over HRME
1. Maximum posterior probability.	.794	.68	5.6
2. Minimum posterior probability.	.824	4.4	9.6
3. Mean posterior probability.	.810	2.7	7.8
4. LOP, Weighted sum of posterior probabilities, weights determined from individual AUC for devices.	.772	-2.2	2.6
5. LOP, Weighted sum of posterior probabilities, weights determined from search. W_{MDM}=.89, W_{HRME}=.11	.875	10.9	16.4
6. LOGP, Product of posterior probabilities.	.811	2.8	7.9
7. LOGP, Weighted Product of posterior probabilities, weights determined from search. W_{MDM}=.87, W_{HRME}=.13	.819	3.9	9.0
8. Weights determined for device and class.	.789	-.06	4.9
9. MDM posterior probability used as HRME prior probability.	.817	3.5	8.6
10. MDM posterior probability used as feature in HRME classifier.	.784	-.68	4.2
11. One classifier using features from both devices.	.790	.11	5.1

MDM Alone



HRME Alone



Combination

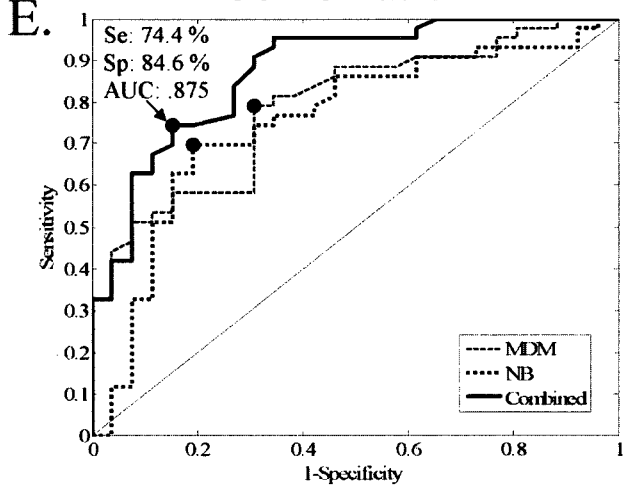


Figure 17. Results from the Nondirected Combinatorial Technique. Each device was used to measure all of the measurement sites. A. Scatterplot of normalized red to green ratio per measurement site calculated from MDM images. The indicated discrimination line yields a sensitivity of 79.1% and a specificity of 69.2%. B. ROC curve produced by the linear classifier using this feature. The indicated operating point corresponds to the discrimination line in A. C. Scatterplot of the two selected features from the HRME. The indicated discrimination line yields a sensitivity of 69.8% and a specificity of 80.8%. D ROC curve produced by the linear classifier using these features. The indicated operating point corresponds to the discrimination line in C. E. The ROC curves in B. and D. with the ROC curve produced by combination method 5 in table 9: the weighted linear opinion pool of posteriors probabilities from each device.

Figure 16A shows a representative MDM autofluorescence image at 405 nm excitation collected from resected tissue from the floor of mouth. MDM ROIs are indicated from a site with squamous cell carcinoma (red circle) and a site which is non-neoplastic (blue circle). Note the relative decrease in autofluorescence at the carcinoma site. Figure 16B shows an HRME image collected at the carcinoma site. Figure 16C shows an HRME image for the non-neoplastic site. Note the nuclear crowding and disorganized architecture in 16B compared to 16C.

Figure 17 shows results from the Nondirected Combinatorial Technique where data from both devices was used from all measurement sites. Figure 17A is a scatterplot of normalized red to green ratio per measurement site calculated from MDM images. The indicated discrimination line yields a sensitivity of 79.1% and a specificity of 69.2%. Figure 17B is the ROC curve produced by the linear classifier using this feature alone. The indicated operating point corresponds to the discrimination line in Figure 17A. and is located at the Qpoint, the point on the ROC curve which is the minimum distance from the upper left hand corner of the plot. Figure 17C is the scatter plot of the two selected feature values from the HRME. The two features chosen by the sequential feature selection algorithm were pixel grayscale variance (a first-order feature) and the pixel pair correlation value calculated from the gray level co-occurrence matrix at an offset of 1. The indicated discrimination line yields a sensitivity of 69.8% and a specificity of 80.8%. Figure 17D is the ROC curve produced by the linear classifier using these features. The indicated operating point corresponds to the discrimination line in 17C and is also at the Qpoint.

Table 9 indicates the AUC achieved using each of the 11 methods to combine data from the two modalities. Table 9 also summarizes the change in AUC for the combined modalities relative to the AUC achieved using each device by itself. Method 5 produces the highest AUC of all of the combination methods. This method highly favors the MDM posterior probabilities by a factor of 8. Figure 17E shows the ROC curve produced using Method 5, the weighted linear opinion pool of posteriors probabilities from each device. This combined method produced an AUC of .875, a sensitivity of 74.4% and a specificity of 87.5%. This AUC is a 10.9% increase of the MDM alone and a 16.4% increase over the HRME alone. The second highest AUC (.824) was produced by method 2 where the lower posterior probability was chosen from the devices for each measurement site. The third highest AUC (.819) was produced by the weighted product of posterior probabilities where the MDM was again weighted highly favorably.

Table 10. The combinatorial methods and the performance achieved. The MDM was used to identify measurement sites with a greater than 90.7% probability of being neoplastic and the HRME was used to measure only these sites. The percent increase in AUC, sensitivity, and specificity values over the MDM alone are shown. The sensitivity of the MDM by itself was 90.7%, the specificity was 38.5%, and the AUC was .789. Method 5 produced the highest increase in AUC.

Combination Method	AUC	% inc. over MDM	Se	% inc over MDM	Sp	% inc. over MDM
1. Maximum posterior probability.	.830	+5.16	76.7	-14.7	84.6	+119.8
2. Minimum posterior probability.	.828	+4.99	76.7	-14.7	88.5	+129.8
3. Mean posterior probability.	.832	+5.44	76.7	-14.7	84.6	+119.8
4. LOP. Weighted sum of posterior probabilities, weights determined from individual AUC for devices.	.832	+5.50	76.7	-14.7	84.6	+119.8
5. LOP. Weighted sum of posterior probabilities, weights determined from search. W_{MDM}=.89, W_{HRME}=.11	.847	+7.37	86.1	-4.4	80.8	+109.8
6. LOGP. Product of posterior probabilities.	.830	+5.22	76.7	-14.7	84.6	+119.8
7. LOGP. Weighted Product of posterior probabilities, weights determined from search. W_{MDM}=.87, W_{HRME}=.13	.833	+5.56	83.7	-6.98	80.8	+109.8
8. Weights determined for device and class.	.827	+4.82	76.7	-14.7	84.6	+119.8
9. MDM posterior probability used as HRME prior probability.	.827	+4.88	76.7	-14.7	84.6	+119.8
10. MDM posterior probability used as feature in HRME classifier.	.809	+2.58	84.6	-5.98	75.0	+94.8
11. One classifier using features from both devices.	.818	+3.74	74.4	-17.3	84.6	+119.8

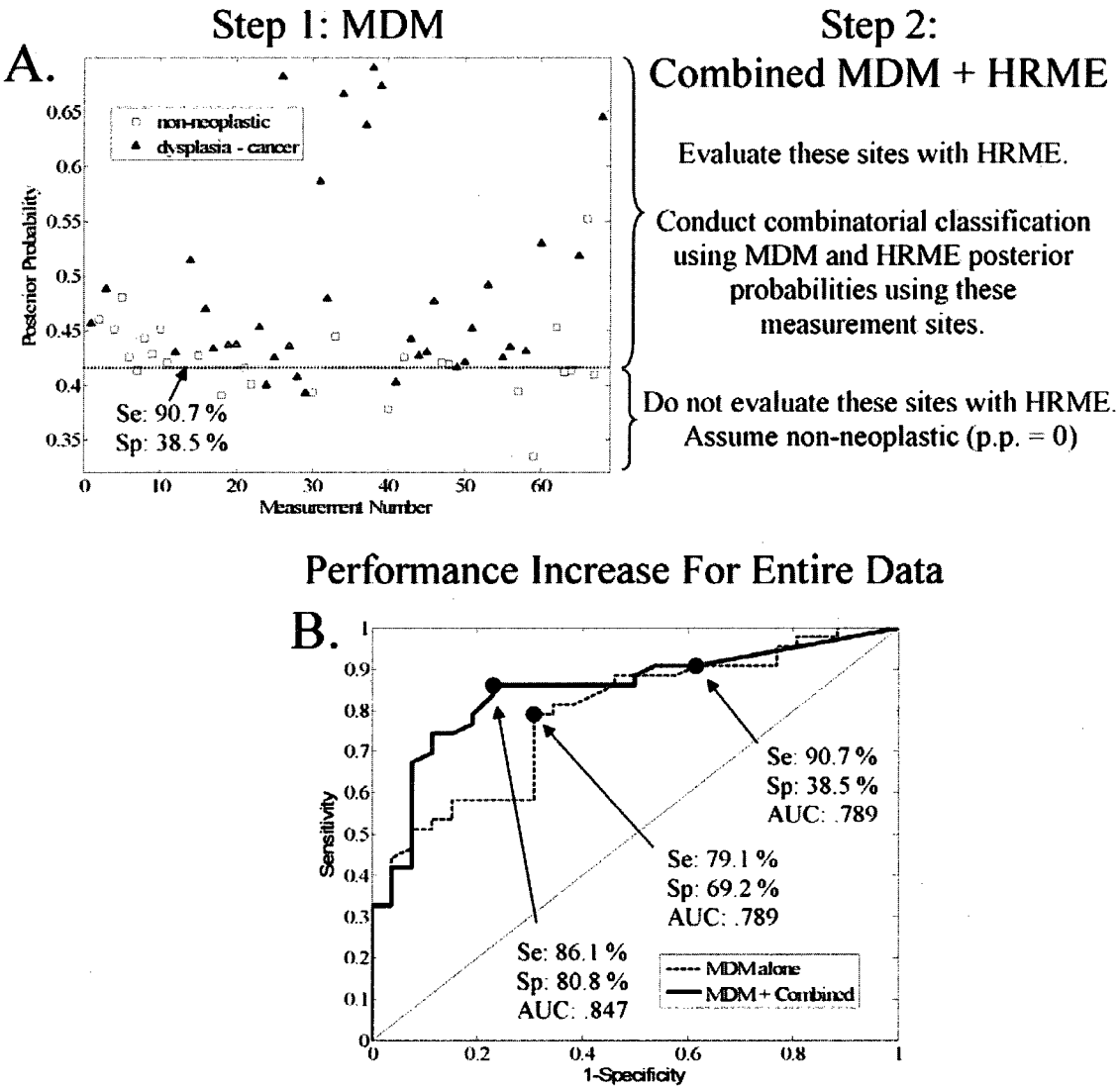


Figure 18. Results from the Directed Combinatorial Technique: when the MDM is first used to identify measurement sites with a sensitivity of 90%, and then the HRME is used to aid diagnosis of these measurement sites. A. Scatterplot of posterior probabilities per measurement site produced from data only from the MDM. The indicated discrimination line yields a sensitivity of 90.7% and a specificity of 38.5%. Measurements above this line are tested using the HRME. Measurements below this line are assumed non-

neoplastic and assigned a posterior probability of 0. B. ROC curves produced by the MDM alone and from the 2 step method described. The AUC increases from .789 to .847 using this two step method. The qpoint is indicated on both ROC curves. The operating point yielding 90.7% sensitivity is also shown on the ROC produced from the MDM alone.

Figure 18 and Table 10 show results from the Directed Combinatorial Technique. The MDM was first used to identify measurement sites with a sensitivity of 90.0%. For this analysis however, the closest operating point available from the linear classifier trained on the MDM data was at 90.7% sensitivity. The HRME data was then used to aid in diagnosis of only these measurement sites. Figure 18A is a scatter plot of posterior probabilities per measurement site produced from data only from the MDM. The indicated discrimination line yields a sensitivity of 90.7% and a specificity of 38.5%. Measurements below this line are assumed non-neoplastic. There are 14 measurement sites below the threshold, 10 sites are non-neoplastic and 4 are neoplastic confirmed by histopathology. HRME data is utilized in the combination methods with measurements above this line. ROC curves and performance metrics were computed from the combination of posterior probabilities from all measurement sites.

The highest AUC from the Directed Combinatorial Technique was achieved using the weighted linear opinion pool of posteriors probabilities from each device (Table 10). The same weights used for both the Directed and Nondirected Combinatorial Techniques. Figure 18B shows the ROC curves produced by the MDM alone and from the Directed

Combinatorial Technique. The AUC increases from .789 to .847 using this method. The sensitivity drops from 90.7% to 86.1%. The specificity increases from 36.5% to 80.8%. The qpoint is indicated on both ROC curves. The operating point yielding 90.7% sensitivity is also shown on the ROC produced from the MDM alone.

The highest AUC produced from the Nondirected Combinatorial Technique was .875 whereas the highest AUC achieved from the Directed Combinatorial Technique was .847. Data from 14 (20.3%) fewer measurements sites were required from the HRME for the Directed Combinatorial Technique.

6.4 Discussion

In this study, image data from two optical imaging devices were combined to make a quantitative, objective classification algorithm to distinguish between neoplastic and non-neoplastic oral mucosa. Both of the combination strategies improved classification performance over either device alone. The non-directed combinatorial technique yielded a 10.9% increase in AUC over the MDM alone and a 16.4% increase over the HRME alone. In addition to the combination of physically different measurement data provided by two different imaging modalities, independent imaging systems allow for the development of classification algorithms tailored to provide a desired sensitivity or specificity value. The Directed Combinatorial Technique described in this paper used the wide field imaging data alone, with a decision line set at 90% sensitivity as an initial screen. The abundant false positives remaining after this screen were reduced in a second step using combined wide field and high resolution imaging data. The selection of 90% sensitivity during the initial screen was made so that most of

the true positives would be captured while the number of measurements sites to be measured by the HRME is reduced. While the final result following the second screen did not achieve a 90% sensitivity (impossible unless the second method was able to classify with a 100% sensitivity at a nonzero specificity), the combined result still yields a significant improvement, 7.4% increase in AUC, over the MDM alone. Additionally, the final sensitivity at the Qpoint was higher and more reflective of the initial sensitivity threshold for the Directed Combinatorial Technique compared to the Nondirected (86.1% sensitivity for the Directed compared to 74.4% for the Nondirected). These results compared well to the weighted LOP classification results from the Nondirected Combinatorial Technique, but allows the user to better control the desired prediction efficiency, and in this case required measurement 20.3% fewer sites with the HRME.

In a previous *in vivo* study using the MDM a sensitivity of 100% and specificity of 91.4% was achieved for discriminating neoplastic from non-neoplastic tissue [111]. The MDM performance was significantly worse in this study which yielded a sensitivity of 79.1% and a specificity of 69.2%. These differences may be due to the changes in resected specimens compared to *in vivo* tissue. Qualitative comparisons of the same tissue *in vivo* and *ex vivo* show differences in autofluorescence intensity which may correlate with the cessation of blood flow and other biochemical changes.

Quantitative screening tools, such as the combinatorial wide field and high-resolution imaging algorithm presented in the text, offers clinicians an objective means of examining the oral cavity. Since this technique requires minimal input from the user and a carries a very low per-use cost, higher volume and more frequent screening could be implemented, with the goal of discovering neoplastic lesions of the oral cavity at an

earlier, more treatable stage. Such early diagnosis would enable more conservative, tissue-sparing treatment options, with a better overall outcome for the patient.

CHAPTER 7: CONCLUSION

7.1 Summary and Research Contributions

Widefield optical imaging is an attractive technology for improving oral cancer detection and delineation. This method allows the for relatively fast inspection of the entire oral mucosal surface and can be implemented inexpensively compared to other biomedical imaging screening and diagnostic technologies such as MRI, CT, and mammography. Optical imaging has the potential to exploit anatomical and biochemical changes which occur during malignant progression in order to detect disease. Examples of these changes include increased vessel growth due to angiogenesis and the remodeling of the basement membrane. In order to successfully implement widefield optical imaging as a diagnostic tool for oral neoplasia, it is necessary to investigate and quantify the performance of this method as well as determine an ideal instrumentation design. These issues were the primary objectives of this dissertation work.

The first part of this work, described in Chapter 3, was to design and construct a scientific grade widefield optical microscope with the capability of narrowband, autofluorescence, and polarized imaging modalities at a variety of exploratory illumination and collections wavelengths. This device is referred to as the Multipsectral Digital Microscope (MDM). The MDM can capture digital images at four narrowband illumination wavelengths, four autofluorescence excitation wavelengths, and with orthogonal and parallel polarization. The field of view is adjustable and can varied between approximately 1 cm to 7 cm. The spatial resolution is sufficient to capture vascular patterns common on the surface of oral mucosa. The MDM was constructed so

as to be portable and robust enough for *in vivo* imaging of patients. A custom mercury arc lamp illumination system was designed to provide sufficient illumination power for practical *in vivo* autofluorescence imaging with minimal motion artifacts, with exposure times of approximately .5 seconds.

Qualitative observations were made from image data from normal volunteers and cancer patients. Under narrowband illumination conditions, differences in visible vascular patterns were apparent in normal volunteers. At blue illumination, fine vascular patterns were visible on anatomic regions such as the hard palate, soft palate, inner lip, and floor of mouth. Vessels appeared dark compared to surrounding tissue. As illumination went from blue to green to yellow to red, finer vasculature increasingly became less visible and while larger, deeper vessels became more visible. These observations were likely a due to a combination of increased penetration depth at longer wavelengths and differences in hemoglobin absorption at the illumination wavelengths. Orthogonal, or cross polarization imaging had the general effect of revealing increased vascularity. Autofluorescence imaging revealed differences in fluorescence intensity at different anatomical sites but relatively uniform intensity within each site. Highly keratinized sites such as the hard palate generally had a higher autofluorescence intensity. Vessels were often visible under autofluorescence mode.

Multispectral imaging of oral lesions from patients revealed several qualitative observations which increased their visibility compared to white light. Relatively flat lesions without exophytic tumor growth commonly appeared to have negative contrast when compared to surrounding normal tissue using narrowband, autofluorescence, and cross polarized imaging. In narrowband imaging mode the negative contrast was more

easily observed using green and yellow illuminations and appeared to correlate with increased vasculature and vessel networks. In cross polarized, white light mode, these lesions appeared to be a darker red than surrounding tissue, also likely a result of vasculature networks below the mucosal surface. In autofluorescence mode, almost all lesions, including exophytic tumors, either appeared dark compared to surrounding normal tissue, or had an increased red fluorescence. These two observations have been previously reported in the literature. The decreased autofluorescence signal is most commonly attributed to the breakdown of collagen and elastin crosslinks in the extracellular matrix of the mucosal tissue. These crosslinks are brightly fluorescent and as tissue is remodeled during malignant progression, these crosslinks are thought to be degraded and restructured to allow new growth. A recent publication by our group helps to support this theory [6]. The decrease in autofluorescence appeared to be the most consistent observation which discriminated oral lesions with a diagnosis of dysplasia or cancer, from non-neoplastic tissue. The red fluorescence is commonly attributed to porphyrins, although it is unclear if the origins are endogenous or from bacterial growth.

Based on these qualitative observations, image processing and computer aided diagnostics were implemented to determine an objective diagnosis. This part of the work is detailed in Chapters 4 and 5. This work began with an analysis of the autofluorescence images from 56 patients with previously determined oral lesions and 11 normal volunteers. Regions of interest (ROIs) of lesions areas and clinically normal areas were selected by a physician blinded to the autofluorescence images. The mean red to green ratio of pixel intensities of these regions, normalized to ROIs with a non-neoplastic diagnosis, were compared and it was found that using this single feature, dysplastic and

cancerous lesions could be discriminated from non-neopalstic regions with a sensitivity of 100% and a specificity of 91.4%.

Probability maps of disease constructed from the classification results are also presented in Chapter 4. For the small number of samples tested, these maps qualitatively agree with histopathology and present an additional analysis tool for detection and/or delineation of the lesion extent.

Previous studies have reported high sensitivity and specificity of oral lesion detection using autofluorescence based on qualitative observation and discrimination [21,55]. This work is the first in which the instrument performance was determined objectively using computer aided diagnosis. The accuracy of detection based only from the ratio of mean red to green feature suggests that this feature is highly conserved among premalignant and malignant lesions in the oral cavity.

Chapter 5 significantly extends the computer aided diagnosis by including images from the narrowband and polarized imaging modalities. A plethora of features were extracted from these images based in part on the observations made in Chapter 3. Textural and frequency domain features were extracted under the hypothesis that these features might quantify the vascular pattern observations seen from lesions described in Chapter 3. Additional statistical and colorimetric features were also included in the analysis. Ultimately, features from the other modalities caused no increase in performance over the single feature of red to green ratio from the 405 nm excitation autofluorescence image, and no other features performed as well as this single feature. This finding suggests that features produced by vascular pattern increases observed in some oral lesions, and abnormal textural appearance of tumors, do not translate into a

quantitative features which are universal enough to provide an accurate diagnosis for a variety of oral lesions with a spectrum of diagnoses. This finding also supports the dominance of the decreased autofluorescence and the feature of red to green ratio for discrimination of non-neoplastic and neoplastic tissue.

The analysis in Chapter 5 also included a comparison of image contrast computed from the different imaging modalities from lesions with different diagnostic grades, as well an attempt to classify lesions into one of three classes: non-neoplastic, dysplasia, or cancer. The largest majority of lesions had an increased contrast compared to white light at 405 nm excitation autofluorescence mode. Importantly, while this increase was statistically significant between non-neoplastic tissue and neoplastic tissue, there was no statistical difference between diagnostic grades within the neoplastic class. Similarly, 3-class classification by computer aided diagnostics was largely unsuccessful. These results suggest that the tested imaging modalities, and the red to green ratio from 405 nm autofluorescence images in particular, are capable of discriminating non-neoplastic tissue from neoplastic tissue, but are not capable of making a more specific diagnosis. This finding may have significant implications concerning the clinical use of this technology.

In Chapter 6, a high resolution microendoscope (HRME), capable of imaging cellular detail when used with a nuclear fluorescent dye, was used simultaneously with the MDM in a study of *ex vivo* oral resections presented. This work was conducted to determine if these two different optical devices could improve diagnostic performance over either alone. The quantities measured by the two devices are fundamentally different. The difference in the red to green ratio feature extracted from MDM images in normal and abnormal tissue was likely based on changes in the basement membrane of

tissue. The HRME system, conversely, images only the upper layers of the epithelium and the information content in largely based on cell-to-cell spacing and nuclear-to-cytoplasmic ratio. Additionally, the MDM can be used to scan the entire oral mucosal surface quickly where as the HRME has a small field of view suggesting that it may be used for confirming or denying suspicious areas identified by the MDM.

A pilot study of ex vivo tissue from 14 patients was conducted using image data collected from both the MDM and HRME. Multiple consensus hybrid techniques, which are methods of combining data from both devices to make a single diagnosis, were tested to explore the benefits of using the two devices together. An 11% increase in performance over the MDM alone and a 16% increase over the HRME alone was achieved using one combination strategy. While these are substantial increases, the final sensitivity and specificity achieved were only 74.4% and 84.6% respectively suggesting that the MDM analysis methods developed for *in vivo* tissue may not be as effective in *ex vivo* tissue. This performance difference could also conceivably be due to the small sample size in this study. A second combination strategy achieved similar results but required approximately 20% fewer measurements sites with the HRME.

7.2 Future Research Directions

There are several future research directions which would behoove the further development of widefield optical imaging for detection and delineation of oral neoplasia. Firstly, the MDM or another analogous portable system with autofluorescence capability at 405 nm excitation should be tested on a population with a high incidence of focal inflammation and other common benign conditions which are likely to occur in a

screening population. The patient population tested in the clinical studies in Chapters 3-6 were from a tertiary care center where malignant and premalignant lesions had already been verified by biopsy. While the non-neoplastic regions selected for analysis almost always contained either inflammation, hyperkeratosis, hyperplasia, or a combination of these conditions, it was not tested if other common benign conditions would produce false positives. Furthermore, testing this technology on a larger patient population and in different geographic locations will help to confirm its effectiveness.

Secondly, further analysis should be conducted into the selection of the clinically normal area used to normalize the red to green ratio. It was observed, but not quantified during the analysis conducted in this work, that the classification algorithm accuracy was dependent on the choice of location for this selected clinically normal region. While there was an attempt to select a clinically normal region contralateral to the lesions or distal, but on the same anatomic site as the lesion of interest, this was not always possible because of the captured field of view available in the collected images. The classification was particularly sensitive to selected region when it was chosen near the dorsal tongue, which often had increased red fluorescence even in normal volunteers.

Thirdly, the probability mapping analysis described in Chapter 4 should be expanded to reveal any greater potential. The probability mapping procedure was only demonstrated on a few patients, largely because of the practical limitations of obtaining histopathology for entire resected specimens. Histopathology may be easier to correlate to images of resected tissue and this may provide a viable pathway to further this analysis. Additionally, work should be conducted to help segment the lesion and define the lesion border based on the probability mapping. K-means clustering and other

unsupervised learning techniques may aid in this problem. The threshold for the lesion border can be altered and compared to full resection pathology maps to find appropriate values. Furthering this area has the potential of extending the use of widefield imaging to margin detection in addition to screening.

Fourth, the instrument hybrid analysis described in Chapter 6 should be extended to other secondary instruments to help answer of the limitations of the MDM. Results from Chapter 5 indicate that the red to green ratio can discriminate non-neoplastic tissue from neoplastic tissue well, but likely cannot make a more specific diagnosis with the tested configuration. Clinically, more diagnostic information would be beneficial to guide treatment. Additionally, a confirmation of the MDM diagnosis by a device which measures a fundamentally different morphological or biochemical feature would be beneficial. Depth sensitive point probe fluorescence and reflectance spectroscopy is a well tested technology which could be combined with the widefield approach. In a complementary fashion, the widefield imaging may help to answer some of the limitations of point probe devices by identifying suspicious lesions after screening the entire oral mucosa, a procedure which is impractical with any device with a small field of view or measurement area.

Finally, closer investigation into the biological origins of the most important imaging features observed in this work would help in understanding why this method works well in a tertiary care patient population, and if it should generalize to larger, more diverse populations with more benign conditions. Additionally, understanding the origins of these observations may help to explain why the natural discrimination line when using the red to green ratio appears to be between non-neopalstic tissue and lesions with any

grade of dysplasia or cancer, including mild dysplasia. In order to supplement the work already done by Pavlova et al. in this area [6] I would recommend conducting a tissue microarray study to compare regions with decreased autofluorescence to regions with normal autofluorescence. Immunohistochemical staining of suspected causes such as MMPs, known to breakdown the basement membrane, and collagen and elastin fibers could shine light on the underlying biological functions being measured with the autofluorescence imaging.

REFERENCES

1 Society, A.C., *Cancer Facts & Figures 2008*. 2008, American Cancer Society: Atlanta.

2 Jemal, A., M.J. Thun, L.A. Ries, H.L. Howe, H.K. Weir, M.M. Center, et al., *Annual report to the nation on the status of cancer, 1975-2005, featuring trends in lung cancer, tobacco use, and tobacco control.* J Natl Cancer Inst 100(23): p. 1672-94, (2008).

3 American Cancer Society, "Cancer Facts and Figures: 2005". (2005).

4 Silverman, S., *Oral Cancer*. Atlas of Clinical Oncology, ed. Fifth. 2003, Hamilton, London: American Cancer Society.

5 Center for Disease Control, "Morbidity and Mortality Weekly Report: Oral Cancer". 47(No. RR-14), (1998).

6 Pavlova, I., M. Williams, A. El-Naggar, R. Richards-Kortum, and A. Gillenwater, *Understanding the biological basis of autofluorescence imaging for oral cancer detection: high-resolution fluorescence microscopy in viable tissue.* Clin Cancer Res 14(8): p. 2396-404, (2008).

7 NCI, "SEER Program: Oral Cancer". p. 96-99, (2003).

8 Tsai, T., H.M. Chen, C.Y. Wang, J.C. Tsai, C.T. Chen, and C.P. Chiang, *In vivo autofluorescence spectroscopy of oral premalignant and malignant lesions: distortion of fluorescence intensity by submucous fibrosis.* Lasers Surg Med 33(1): p. 40-7, (2003).

9 Muller, M.G., T.A. Valdez, I. Georgakoudi, V. Backman, C. Fuentes, S. Kabani, et al., *Spectroscopic detection and evaluation of morphologic and biochemical changes in early human oral carcinoma.* Cancer 97(7): p. 1681-92, (2003).

10 Pavlova, I., K. Sokolov, R. Drezek, A. Malpica, M. Follen, and R. Richards-Kortum, *Microanatomical and biochemical origins of normal and precancerous cervical autofluorescence using laser-scanning fluorescence confocal microscopy.* Photochem Photobiol 77(5): p. 550-5, (2003).

11 Cotran, *Pathologic Basis of Disease*, ed. sixth. 1999, Philadelphia: W.B. Saunders Company.

12 Drezek, R., M. Guillaud, T. Collier, I. Boiko, A. Malpica, C. Macaulay, et al., *Light scattering from cervical cells throughout neoplastic progression: influence*

of nuclear morphology, DNA content, and chromatin texture. J Biomed Opt 8(1): p. 7-16, (2003).

13 Drezek, R., K. Sokolov, U. Utzinger, I. Boiko, A. Malpica, M. Follen, et al., *Understanding the contributions of NADH and collagen to cervical tissue fluorescence spectra: modeling, measurements, and implications.* J Biomed Opt 6(4): p. 385-96, (2001).

14 Svistun, E., R. Alizadeh-Naderi, A. El-Naggar, R. Jacob, A. Gillenwater, and R. Richards-Kortum, *Vision enhancement system for detection of oral cavity neoplasia based on autofluorescence.* Head Neck 26(3): p. 205-15, (2004).

15 Shin, D.M., I.B. Gimenez, J.S. Lee, K. Nishioka, M.J. Wargovich, S. Thacher, et al., *Expression of epidermal growth factor receptor, polyamine levels, ornithine decarboxylase activity, micronuclei, and transglutaminase I in a 7,12-dimethylbenz(a)anthracene-induced hamster buccal pouch carcinogenesis model.* Cancer Res 50(8): p. 2505-10, (1990).

16 Todd, R. and D.T. Wong, *Epidermal growth factor receptor (EGFR) biology and human oral cancer.* Histol Histopathol 14(2): p. 491-500, (1999).

17 Rosenthal, E.L. and L.M. Matrisian, *Matrix metalloproteases in head and neck cancer.* Head Neck 28(7): p. 639-48, (2006).

18 Patel, B.P., S.V. Shah, S.N. Shukla, P.M. Shah, and P.S. Patel, *Clinical significance of MMP-2 and MMP-9 in patients with oral cancer.* Head Neck 29(6): p. 564-72, (2007).

19 Richards-Kortum, R. and E. Sevick-Muraca, *Quantitative optical spectroscopy for tissue diagnosis.* Annu Rev Phys Chem 47: p. 555-606, (1996).

20 Drezek, R., C. Brookner, I. Pavlova, I. Boiko, A. Malpica, R. Lotan, et al., *Autofluorescence microscopy of fresh cervical-tissue sections reveals alterations in tissue biochemistry with dysplasia.* Photochemistry and Photobiology 73(6): p. 636-641, (2001).

21 Lane, P.M., T. Gilhuly, P. Whitehead, H. Zeng, C.F. Poh, S. Ng, et al., *Simple device for the direct visualization of oral-cavity tissue fluorescence.* J Biomed Opt 11(2): p. 024006, (2006).

22 Downer, M.C., D.R. Moles, S. Palmer, and P.M. Speight, *A systematic review of test performance in screening for oral cancer and precancer.* Oral Oncol 40(3): p. 264-73, (2004).

23 Sankaranarayanan, R., B. Mathew, B.J. Jacob, G. Thomas, T. Somanathan, P. Pisani, et al., *Early findings from a community-based, cluster-randomized,*

controlled oral cancer screening trial in Kerala, India. The Trivandrum Oral Cancer Screening Study Group. Cancer 88(3): p. 664-73, (2000).

24 Ramadas, K., R. Sankaranarayanan, B.J. Jacob, G. Thomas, T. Somanathan, C. Mahe, et al., *Interim results from a cluster randomized controlled oral cancer screening trial in Kerala, India.* Oral Oncol 39(6): p. 580-8, (2003).

25 Nagao, T., N. Ikeda, H. Fukano, H. Miyazaki, M. Yano, and S. Warnakulasuriya, *Outcome following a population screening programme for oral cancer and precancer in Japan.* Oral Oncol 36(4): p. 340-6, (2000).

26 Gillenwater, A., R. Jacob, R. Ganeshappa, B. Kemp, A.K. El-Naggar, J.L. Palmer, et al., *Noninvasive diagnosis of oral neoplasia based on fluorescence spectroscopy and native tissue autofluorescence.* Arch Otolaryngol Head Neck Surg 124(11): p. 1251-8, (1998).

27 Epstein, J.B., C. Scully, and J. Spinelli, *Toluidine blue and Lugol's iodine application in the assessment of oral malignant disease and lesions at risk of malignancy.* J Oral Pathol Med 21(4): p. 160-3, (1992).

28 Ramanujam, N., M.F. Mitchell, A. Mahadevan, S. Thomsen, E. Silva, and R. Richards-Kortum, *Fluorescence spectroscopy: a diagnostic tool for cervical intraepithelial neoplasia (CIN).* Gynecol Oncol 52(1): p. 31-8, (1994).

29 Tunnell, J.W., A.E. Desjardins, L. Galindo, I. Georgakoudi, S.A. McGee, J. Mirkovic, et al., *Instrumentation for multi-modal spectroscopic diagnosis of epithelial dysplasia.* Technol Cancer Res Treat 2(6): p. 505-14, (2003).

30 Wang, C.Y., H.K. Chiang, C.T. Chen, C.P. Chiang, Y.S. Kuo, and S.N. Chow, *Diagnosis of oral cancer by light-induced autofluorescence spectroscopy using double excitation wavelengths.* Oral Oncol 35(2): p. 144-50, (1999).

31 Mitchell, M.F., S.B. Cantor, C. Brookner, U. Utzinger, D. Schottenfeld, and R. Richards-Kortum, *Screening for squamous intraepithelial lesions with fluorescence spectroscopy.* Obstet Gynecol 94(5 Pt 2): p. 889-96, (1999).

32 Dhingra, J.K., X. Zhang, K. McMillan, S. Kabani, R. Manoharan, I. Itzkan, et al., *Diagnosis of head and neck precancerous lesions in an animal model using fluorescence spectroscopy.* Laryngoscope 108(4 Pt 1): p. 471-5, (1998).

33 Nath, A., K. Rivoire, S. Chang, L. West, S.B. Cantor, K. Basen-Engquist, et al., *A pilot study for a screening trial of cervical fluorescence spectroscopy.* Int J Gynecol Cancer 14(6): p. 1097-107, (2004).

34 Clark, A.L., A.M. Gillenwater, T.G. Collier, R. Alizadeh-Naderi, A.K. El-Naggar, and R.R. Richards-Kortum, *Confocal microscopy for real-time detection of oral cavity neoplasia.* Clin Cancer Res 9(13): p. 4714-21, (2003).

35 Clark, A.L., A. Gillenwater, R. Alizadeh-Naderi, A.K. El-Naggar, and R. Richards-Kortum, *Detection and diagnosis of oral neoplasia with an optical coherence microscope.* J Biomed Opt 9(6): p. 1271-80, (2004).

36 Kluftinger, A.M., N.L. Davis, N.F. Quenville, S. Lam, J. Hung, and B. Palcic, *Detection of squamous cell cancer and pre-cancerous lesions by imaging of tissue autofluorescence in the hamster cheek pouch model.* Surg Oncol 1(2): p. 183-8, (1992).

37 Ntziachristos, V., C. Bremer, and R. Weissleder, *Fluorescence imaging with near-infrared light: new technological advances that enable in vivo molecular imaging.* Eur Radiol 13(1): p. 195-208, (2003).

38 Kulapaditharom, B. and V. Boonkitticharoen, *Laser-induced fluorescence imaging in localization of head and neck cancers.* Ann Otol Rhinol Laryngol 107(3): p. 241-6, (1998).

39 Betz, C.S., M. Mehlmann, K. Rick, H. Stepp, G. Grevers, R. Baumgartner, et al., *Autofluorescence imaging and spectroscopy of normal and malignant mucosa in patients with head and neck cancer.* Lasers Surg Med 25(4): p. 323-34, (1999).

40 Lieber, C., Demos, S., *Multimodal near infrared spectral imaging as an exploratory tool for dysplastic esophogeal lesion identification.* Optics Express 14(6), (2006).

41 Zeng, H., MacAulay,C, *Real-time endoscopic fluorescence imaging for early cancer detection in the gastrointestinal tract.* Bioimaging 6, (1998).

42 Andersson-Engels, S., *Medical diagnostic system based on simultaneous multispectral fluorescence imaging.* Appl Opt 33(34), (1994).

43 Park, S.Y., T. Collier, J. Aaron, M.K. Markey, R. Richards-Kortum, K. Sokolov, et al., *Multispectral digital microscopy for in vivo monitoring of oral neoplasia in the hamster cheek pouch model of carcinogenesis.* Optics Express 13(3): p. 749-762, (2005).

44 Onizawa, K., H. Saginoya, Y. Furuya, and H. Yoshida, *Fluorescence photography as a diagnostic method for oral cancer.* Cancer Lett 108(1): p. 61-6, (1996).

45 Andersson-Engels, S., C. Klinteberg, K. Svanberg, and S. Svanberg, *In vivo fluorescence imaging for tissue diagnostics.* Phys Med Biol 42(5): p. 815-24, (1997).

46 Gandjbakhche, A.H., V. Chernomordik, D. Hattery, M. Hassan, and I. Gannot, *Tissue characterization by quantitative optical imaging methods.* Technol Cancer Res Treat 2(6): p. 537-51, (2003).

47 Arens, C., T. Dreyer, H. Glanz, and K. Malzahn, *Indirect autofluorescence laryngoscopy in the diagnosis of laryngeal cancer and its precursor lesions.* Eur Arch Otorhinolaryngol 261(2): p. 71-6, (2004).

48 Palasz, Z., A. Grobelny, E. Pawlik, M. Fraczek, M. Zalesska-Krecicka, A. Klimczak, et al., *Investigation of normal and malignant laryngeal tissue by autofluorescence imaging technique.* Auris Nasus Larynx 30(4): p. 385-9, (2003).

49 Nakayama, A., A.C. Bianco, C.Y. Zhang, B.B. Lowell, and J.V. Frangioni, *Quantitation of brown adipose tissue perfusion in transgenic mice using near-infrared fluorescence imaging.* Mol Imaging 2(1): p. 37-49, (2003).

50 Drezek, R.A., R. Richards-Kortum, M.A. Brewer, M.S. Feld, C. Pitris, A. Ferenczy, et al., *Optical imaging of the cervix.* Cancer 98(9 Suppl): p. 2015-27, (2003).

51 Utzinger, U., M. Bueeler, S. Oh, D.L. Heintzelman, E.S. Svistun, M. Abd-El-Barr, et al., *Optimal visual perception and detection of oral cavity neoplasia.* IEEE Trans Biomed Eng 50(3): p. 396-9, (2003).

52 Coghlan, L., U. Utzinger, R. Drezek, D. Heintzelman, A. Zuluaga, C. Brookner, et al., *Optimal fluorescence excitation wavelengths for detection of squamous intra-epithelial neoplasia: results from an animal model.* Optics Express 7(12): p. 436-446, (2000).

53 Heintzelman, D.L., U. Utzinger, H. Fuchs, A. Zuluaga, K. Gossage, A.M. Gillenwater, et al., *Optimal excitation wavelengths for in vivo detection of oral neoplasia using fluorescence spectroscopy.* Photochem Photobiol 72(1): p. 103-13, (2000).

54 Ingrams, D.R., J.K. Dhingra, K. Roy, D.F. Perrault, Jr., I.D. Bottrill, S. Kabani, et al., *Autofluorescence characteristics of oral mucosa.* Head Neck 19(1): p. 27-32, (1997).

55 De Veld, D.C., M.J. Witjes, H.J. Sterenborg, and J.L. Roodenburg, *The status of in vivo autofluorescence spectroscopy and imaging for oral oncology.* Oral Oncol 41(2): p. 117-31, (2005).

56 Ji, Q., J. Engel, and E. Craine, *Texture analysis for classification of cervix lesions.* IEEE Trans Med Imaging 19(11): p. 1144-9, (2000).

57 Gossage, K.W., T.S. Tkaczyk, J.J. Rodriguez, and J.K. Barton, *Texture analysis of optical coherence tomography images: feasibility for tissue classification.* J Biomed Opt 8(3): p. 570-5, (2003).

58 Hadjiiski, L., B. Sahiner, and H.P. Chan, *Advances in computer-aided diagnosis for breast cancer.* Curr Opin Obstet Gynecol 18(1): p. 64-70, (2006).

59 Duda, R.O., P.E. Hart, and D.G. Stork, *Pattern Classification.* Second ed. 2001, New York: John wiley & Sons, Inc.

60 Jain, A.K., R.P.W. Duin, and J.C. Mao, *Statistical pattern recognition: A review.* Ieee Transactions on Pattern Analysis and Machine Intelligence 22(1): p. 4-37, (2000).

61 Johnson, R.A., D.W. Wichern, *Applied Multivariate Statistical Analysis.* Sixth ed. 2007, Upper Saddle Rive, NJ: Pearson Prentice Hall.

62 Breiman L, F.J., Olshen RA, Stone CJ, *Classification and Regression Trees.* The Wadsworth Statistics/Probability Series, ed. C.W. Bickel PJ, Dudley RM. 1984, Belmont, California: Wadsworth International Group.

63 Albert, A. and E.K. Harris, *Multivariate interpretation of clinical laboratory data.* 1987, New York: Marcel Dekker inc.

64 Liu, H. and L. Yu, *Toward integrating feature selection algorithms for classification and clustering.* Ieee Transactions on Knowledge and Data Engineering 17(4): p. 491-502, (2005).

65 Parkin, D.M., F. Bray, J. Ferlay, and P. Pisani, *Global cancer statistics, 2002.* CA Cancer J Clin 55(2): p. 74-108, (2005).

66 Jemal, A., T. Murray, E. Ward, A. Samuels, R.C. Tiwari, A. Ghafoor, et al., *Cancer statistics, 2005.* CA Cancer J Clin 55(1): p. 10-30, (2005).

67 Shibuya, H., S. Hisamitsu, S. Shioiri, J. Horiuchi, and S. Suzuki, *Multiple primary cancer risk in patients with squamous cell carcinoma of the oral cavity.* Cancer 60(12): p. 3083-6, (1987).

68 Slaughter, D.P., H.W. Southwick, and W. Smejkal, *Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin.* Cancer 6(5): p. 963-8, (1953).

69 Lam, S., *Localization of bronchial intraepithelial neoplastic lesions by fluorescence bronchoscopy.* Chest: p. 696-702, (1998).

70 Inaguma, M. and K. Hashimoto, *Porphyrin-like fluorescence in oral cancer: In vivo fluorescence spectral characterization of lesions by use of a near-ultraviolet excited autofluorescence diagnosis system and separation of fluorescent extracts by capillary electrophoresis.* Cancer 86(11): p. 2201-11, (1999).

71 Onizawa, K., H. Saginoya, Y. Furuya, H. Yoshida, and H. Fukuda, *Usefulness of fluorescence photography for diagnosis of oral cancer.* Int J Oral Maxillofac Surg 28(3): p. 206-10, (1999).

72 Paczona, R., S. Temam, F. Janot, P. Marandas, and B. Luboinski, *Autofluorescence videoendoscopy for photodiagnosis of head and neck squamous cell carcinoma.* Eur Arch Otorhinolaryngol 260(10): p. 544-8, (2003).

73 Pazouki, S., D.M. Chisholm, M.M. Adi, G. Carmichael, M. Farquharson, G.R. Ogden, et al., *The association between tumour progression and vascularity in the oral mucosa.* J Pathol 183(1): p. 39-43, (1997).

74 Subhash, N., J.R. Mallia, S.S. Thomas, A. Mathews, P. Sebastian, and J. Madhavan, *Oral cancer detection using diffuse reflectance spectral ratio R540/R575 of oxygenated hemoglobin bands.* J Biomed Opt 11(1): p. 014018, (2006).

75 Jacques, S.L., J.C. Ramella-Roman, and K. Lee, *Imaging skin pathology with polarized light.* J Biomed Opt 7(3): p. 329-40, (2002).

76 Lindeboom, J.A., K.R. Mathura, and C. Ince, *Orthogonal polarization spectral (OPS) imaging and topographical characteristics of oral squamous cell carcinoma.* Oral Oncol 42(6): p. 581-5, (2006).

77 Utzinger, U., M. Bueeler, S. Oh, D.L. Heintzelman, E.S. Svistun, M. Abd-El-Barr, et al., *Optimal visual perception and detection of oral cavity neoplasia.* Ieee Transactions on Biomedical Engineering 50(3): p. 396-399, (2003).

78 ACGIH, *American Conference of Governmental Industrial Hygienists Threshold Limit Values and Biological Exposure Indices for Chemical and Physical Agents.* 2004, Cincinnati: Signature.

79 Betz, C.S., H. Stepp, P. Janda, S. Arbogast, G. Grevers, R. Baumgartner, et al., *A comparative study of normal inspection, autofluorescence and 5-ALA-induced PPIX fluorescence for oral cancer diagnosis.* Int J Cancer 97(2): p. 245-52, (2002).

80 Poh, C.F., S.P. Ng, P.M. Williams, L. Zhang, D.M. Laronde, P. Lane, et al., *Direct fluorescence visualization of clinically occult high-risk oral premalignant disease using a simple hand-held device.* Head Neck 29(1): p. 71-6, (2007).

81 De Felice, C., S. Parrini, A. Barducci, G. Chitano, G. Tonni, and G. Latini, *Abnormal oral mucosal light reflectance in bronchopulmonary dysplasia.* Early Hum Dev 82(4): p. 273-8, (2006).

82 Lippman, S.M. and W.K. Hong, *Second malignant tumors in head and neck squamous cell carcinoma: the overshadowing threat for patients with early-stage disease.* Int J Radiat Oncol Biol Phys 17(3): p. 691-4, (1989).

83 Day, G.L. and W.J. Blot, *Second primary tumors in patients with oral cancer.* Cancer 70(1): p. 14-9, (1992).

84 Lingen, M.W., J.R. Kalmar, T. Karrison, and P.M. Speight, *Critical evaluation of diagnostic aids for the detection of oral cancer.* Oral Oncol 44(1): p. 10-22, (2008).

85 Gillenwater, A., V. Papadimitrakopoulou, and R. Richards-Kortum, *Oral premalignancy: new methods of detection and treatment.* Curr Oncol Rep 8(2): p. 146-54, (2006).

86 Oh, E.S. and D.M. Laskin, *Efficacy of the ViziLite system in the identification of oral lesions.* J Oral Maxillofac Surg 65(3): p. 424-6, (2007).

87 Rosin, M.P., C.F. Poh, M. Guillard, P.M. Williams, L. Zhang, and C. MacaUlay, *Visualization and other emerging technologies as change makers for oral cancer prevention.* Ann N Y Acad Sci 1098: p. 167-83, (2007).

88 Poh, C.F., L. Zhang, D.W. Anderson, J.S. Durham, P.M. Williams, R.W. Priddy, et al., *Fluorescence visualization detection of field alterations in tumor margins of oral cancer patients.* Clin Cancer Res 12(22): p. 6716-22, (2006).

89 Svistun, E., R. Alizadeh-Naderi, A. El-Naggar, R. Jacob, A. Gillenwater, and R. Richards-Kortum, *Vision enhancement system for detection of oral cavity neoplasia based on autofluorescence.* Head and Neck-Journal for the Sciences and Specialties of the Head and Neck 26(3): p. 205-215, (2004).

90 de Veld, D.C., M. Skurichina, M.J. Witjes, R.P. Duin, H.J. Sterenborg, and J.L. Roodenburg, *Clinical study for classification of benign, dysplastic, and malignant oral lesions using autofluorescence spectroscopy.* J Biomed Opt 9(5): p. 940-50, (2004).

91 Lam, S., C. MacAulay, and B. Palcic, *Detection and localization of early lung cancer by imaging techniques.* Chest 103(1 Suppl): p. 12S-14S, (1993).

92 Park, S.Y., M. Follen, A. Milbourne, H. Rhodes, A. Malpica, N. MacKinnon, et al., *Automated image analysis of digital colposcopy for the detection of cervical neoplasia.* J Biomed Opt 13(1): p. 014029, (2008).

93 Kara, M.A., F.P. Peters, P. Fockens, F.J. ten Kate, and J.J. Bergman, *Endoscopic video-autofluorescence imaging followed by narrow band imaging for detecting early neoplasia in Barrett's esophagus.* Gastrointest Endosc 64(2): p. 176-85, (2006).

94 Matsuda, T., Y. Saito, K.I. Fu, T. Uraoka, N. Kobayashi, T. Nakajima, et al., *Does Autofluorescence Imaging Videoendoscopy System Improve the Colonoscopic Polyp Detection Rate?-A Pilot Study.* Am J Gastroenterol, (2008).

95 Rahman, M., P. Chaturvedi, A. Gillenwater, and R. Richards-Kortum, *Low-cost, multimodal, portable screening system for early detection of oral cancer.* J. Biomed. Opt. 13(3), (2008).

96 Roblyer, D., R. Richards-Kortum, K. Sokolov, A.K. El-Naggar, M.D. Williams, C. Kurachi, et al., *Multispectral optical imaging device for in vivo detection of oral neoplasia.* J Biomed Opt 13(2): p. 024019, (2008).

97 Braakhuis, B.J., M.P. Tabor, J.A. Kummer, C.R. Leemans, and R.H. Brakenhoff, *A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications.* Cancer Res 63(8): p. 1727-30, (2003).

98 Zhou, X.-H., McClish, D. K., Obuchowski N. A., *Statistical Methods in Diagnostic Medicine.* 2002, New York: John Wiley and Sons.

99 Newcombe, R.G., *Two-sided confidence intervals for the single proportion: comparison of seven methods.* Stat Med 17(8): p. 857-72, (1998).

100 Kurie, J.M., J.S. Lee, R.C. Morice, G.L. Walsh, F.R. Khuri, A. Broxson, et al., *Autofluorescence bronchoscopy in the detection of squamous metaplasia and dysplasia in current and former smokers.* J Natl Cancer Inst 90(13): p. 991-5, (1998).

101 Hirsch, F.R., S.A. Prindiville, Y.E. Miller, W.A. Franklin, E.C. Dempsey, J.R. Murphy, et al., *Fluorescence versus white-light bronchoscopy for detection of preneoplastic lesions: a randomized study.* J Natl Cancer Inst 93(18): p. 1385-91, (2001).

102 Chen, J.K., R.V. Katz, and D.J. Krutchkoff, *Intraoral squamous cell carcinoma. Epidemiologic patterns in Connecticut from 1935 to 1985.* Cancer 66(6): p. 1288-96, (1990).

103 Mohamed, S.S. and M.A. Salama, *Prostate cancer spectral multifeature analysis using TRUS images.* IEEE Trans Med Imaging 27(4): p. 548-56, (2008).

104 Das, A., C.C. Nguyen, F. Li, and B. Li, *Digital image analysis of EUS images accurately differentiates pancreatic cancer from chronic pancreatitis and normal tissue.* Gastrointest Endosc 67(6): p. 861-7, (2008).

105 Ganeshan, B., K.A. Miles, R.C. Young, and C.R. Chatwin, *Texture analysis in non-contrast enhanced CT: Impact of malignancy on texture in apparently disease-free areas of the liver.* Eur J Radiol, (2008).

106 Lingley-Papadopoulos, C.A., M.H. Loew, M.J. Manyak, and J.M. Zara, *Computer recognition of cancer in the urinary bladder using optical coherence tomography and texture analysis.* J Biomed Opt 13(2): p. 024003, (2008).

107 Ericson, M.B., J. Uhre, C. Strandeberg, B. Stenquist, O. Larko, A.M. Wennberg, et al., *Bispectral fluorescence imaging combined with texture analysis and linear discrimination for correlation with histopathologic extent of basal cell carcinoma.* J Biomed Opt 10(3): p. 034009, (2005).

108 Srivastava, S., J.J. Rodriguez, A.R. Rouse, M.A. Brewer, and A.F. Gmitro, *Computer-aided identification of ovarian cancer in confocal microendoscope images.* J Biomed Opt 13(2): p. 024021, (2008).

109 Hamarneh, G., C. A., and T. Gustavsson, *Active Contour Models: Application to Oral Lesion detection in Color Images.* IEEE International Conference on Systems, Man, and Cybernetics 4: p. 2458-2463, (2000).

110 Mattsson, U., A. Chodorowski, T. Gustavsson, G. Heyden, M. Jontell, and F. Bergquist, *Use of computer-assisted image analysis for noninvasive evaluation of oral lichenoid reactions and oral leukoplakia.* Oral Surg Oral Med Oral Pathol Oral Radiol Endod 79(2): p. 199-206, (1995).

111 Roblyer, D., *Objective Detection and Delineation of Oral Neoplasia Using Autofluorescence Imaging.* Cancer Prevention Research, Accepted for Publication, (2009).

112 Peli, E., *Contrast in complex images.* J Opt Soc Am A 7(10): p. 2032-40, (1990).

113 Bex, P.J. and W. Makous, *Spatial frequency, phase, and the contrast of natural images.* J Opt Soc Am A Opt Image Sci Vis 19(6): p. 1096-106, (2002).

114 Argenti, F., L. Alparone, and G. Benelli, *Fast Algorithms for Texture Analysis Using Cooccurrence Matrices.* Iee Proceedings-F Radar and Signal Processing 137(6): p. 443-448, (1990).

115 Schwarz, R.A., D. Arifler, S.K. Chang, I. Pavlova, I.A. Hussain, V. Mack, et al., *Ball lens coupled fiber-optic probe for depth-resolved spectroscopy of epithelial tissue.* Opt Lett 30(10): p. 1159-61, (2005).

116 Carlson, K., M. Chidley, K.B. Sung, M. Descour, A. Gillenwater, M. Follen, et al., *In vivo fiber-optic confocal reflectance microscope with an injection-molded plastic miniature objective lens.* Appl Opt 44(10): p. 1792-7, (2005).

117 Rouse, A.R. and A.F. Gmitro, *Multispectral imaging with a confocal microendoscope.* Optics Letters 25(23): p. 1708-1710, (2000).

118 Tearney, G.J., M.E. Brezinski, B.E. Bouma, S.A. Boppart, C. Pitris, J.F. Southern, et al., *In Vivo Endoscopic Optical Biopsy with Optical Coherence Tomography.* Science 276: p. 2037-2039, (1997).

119 Muldoon, T.J., S. Anandasabapathy, D. Maru, and R. Richards-Kortum, *High-Resolution Imaging in Barrett's Esophagus: A Novel, Low- Cost Endoscopic Microscope* Gastrointestinal Endoscopy in press, (2008).

120 Roblyer, D., C. Kurachi, V. Stepanek, R.A. Schwarz, M. Williams, A. El-Naggar, et al., *Computer Aided Diagnosis of Oral Cancer and Precancer Using Multispectral Widefield Optical Imaging.* IEEE Trans Med Imaging, (2009 planned submission).

121 Benediktsson, J.A., J.R. Sveinsson, and P.H. Swain, *Hybrid consensus theoretic classification.* Ieee Transactions on Geoscience and Remote Sensing 35(4): p. 833-843, (1997).