Chapter 2

Autofluorescence characteristics of healthy oral mucosa at different anatomical sites

D.C.G. de Veld [1,2]
M. Skurichina [3]
M.J.H. Witjes [1]
R.P.W. Duin [3]
H.J.C.M. Sterenborg [2]
W.M. Star [2]
J.L.N. Roodenburg [1]

[1]*Department of Oral and Maxillofacial Surgery, University Hospital Groningen*
[2]*Photodynamic Therapy and Optical Spectroscopy Programme, Erasmus Medical Center, Rotterdam*
[3]*Pattern Recognition Group, Technical University of Delft*

Lasers Surg Med 2003;32(5):367-76

Summary

***Background and Objectives**: Autofluorescence spectroscopy is a promising tool for oral cancer detection. Its reliability might be improved by using a reference database of spectra from healthy mucosa. We investigated the influence of anatomical location on healthy mucosa autofluorescence.*
***Study Design/Materials and Methods**: Spectra were recorded from 97 volunteers using seven excitation wavelengths (350–450 nm), 455–867 nm emission. We studied intensity and applied principal component analysis (PCA) with classification algorithms. Class overlap estimates were calculated.*
***Results**: We observed differences in fluorescence intensity between locations. These were significant but small compared to standard deviations (SD). Normalized spectra looked similar for locations, except for the dorsal side of the tongue (DST) and the vermilion border (VB). Porphyrin-like fluorescence was observed frequently, especially at DST. PCA and classification confirmed VB and DST to be spectrally distinct. The remaining locations showed large class overlaps.*
***Conclusions**: No relevant systematic spectral differences have been observed between most locations, allowing the use of one large reference database. For DST and VB separate databases are required.*

2.1 Introduction

Autofluorescence spectroscopy is a non-invasive and easily applicable tool for the detection of alterations in the structural and chemical compositions of cells, which may indicate the presence of diseased tissue [1–4]. Autofluorescence of tissues is due to several endogenous fluorophores. These comprise fluorophores from tissue matrix molecules and intracellular molecules like collagen, elastin, and NADH. Early detection of pre-malignant lesions and malignant tumours may reduce patient morbidity and mortality, and therefore is of great clinical importance [5,6].

(Pre-)malignant lesions of the oral mucosa are clinically not easily distinguishable from benign lesions. Current clinical diagnosis procedure, therefore, includes a biopsy. Unfortunately, it is difficult to determine the optimal, i.e., most dysplastic, location for biopsy, leading to repeated biopsies and the risk of underdiagnosis. Autofluorescence spectroscopy can be useful in guiding the clinician to the optimal location for biopsy and has shown promising results for lesion screening as well [7–14].

In a previous study of autofluorescence spectroscopy for the detection of oral lesions performed by our group, we obtained a sensitivity of 86% and a specificity of 100% for the classification of visibly abnormal tissue. In that study, we divided lesion spectra by corresponding contralateral spectra measured in the same patient to correct for intra- and inter-individual variations [13]. However, there is no evidence for the assumption that contralateral tissue in patients with oral lesions is normal. The influence of carcinogens like tobacco smoke and alcoholics can cause long-term damage of the oral mucosa ("condemned mucosa") which can lead to "field cancerization" [15]. This process is reflected in the fluorescence characteristics of the upper aerodigestive tract [16]. Our earlier study showed that spectroscopic changes occur not only at the center and border of lesions, but also in the surroundings, where no abnormalities are visible [13]. This suggests that the distinction between healthy and diseased tissue within a patient is not always well defined. If classification algorithms become less dependent on a patient's own "healthy" oral mucosa, they will probably be more robust. Therefore, we are developing a reference database of autofluores-

cence spectra of healthy mucosa of a large group of volunteers.

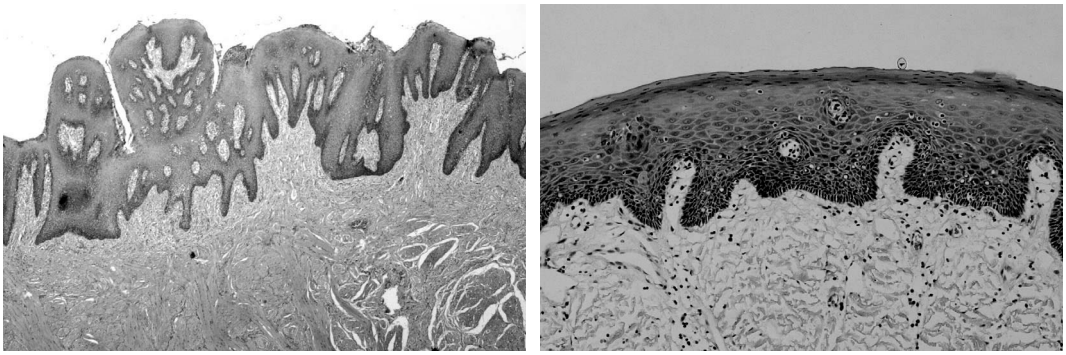

Figure 2.1. Sections of healthy oral mucosa of the dorsal side of the tongue (DST) (a) and of the palate (b). Note the difference in structure between the anatomical locations. Tongue tissue has a papillary structure whereas the palatal epithelium is flat.

However, unlike most other organs suitable for fluorescence detection of malignancies, the oral cavity is lined with a rich variety of mucosal types. Figure 2.1 shows sections of healthy oral mucosa of (a) the dorsal side of the tongue (DST) and (b) the palate, demonstrating the differences in structure between anatomical locations. These variations might translate into differences in autofluorescence spectral shape and intensity. For example, the cheeks, inner lip, soft palate, and floor of the mouth are lined with non-keratinized mucosa, while keratinized mucosa can be found on the hard palate, gingiva, and tongue. The presence of lingual papillae and taste buds makes the histological anatomy of the tongue unique. Furthermore, the palatal mucosa and gingiva are supported by bone, which might increase the reflection of incident and fluorescent light. Savage et al. examined five anatomical locations in the oral cavity and found statistically significant differences in emission wavelength intensity ratios [17]. If spectroscopic differences between anatomical locations are of the same order of magnitude and type as differences between healthy and diseased tissue, we will have to use separate location databases for future lesion comparison. In this way, the lesion classification algorithm will not be affected by variations that are irrelevant for the purpose of lesion diagnostics. In the present study, we investigated the influence of anatomical location on healthy autofluorescence characteristics to determine if a separation into different anatomical reference databases is necessary. We measured oral mucosa autofluorescence spectra at 13 well-defined anatomical locations in 97 healthy volunteers. The spectra were acquired in preparation of a large-scale clinical study in which we measure autofluorescence characteristics of oral lesions.

We used elaborate statistical methods in order to establish any, potentially subtle, differences between spectra from different anatomical sites. After studying fluorescence intensity, we therefore applied k-means clustering, principal component analysis (PCA) using multiple classifiers, and we calculated class overlaps of different locations based on the classification errors of the different classifiers.

2.2 Materials and methods

Volunteer population

Autofluorescence spectra were collected from 97 volunteers with no clinically observable lesions of the oral mucosa, after they had given their informed consent. The population included volunteers from the Department of Oral and Maxillofacial Surgery of the University Hospital of Groningen, as well as patients who had been referred to the hospital for conditions not affecting their oral mucosa. This study was approved by the Institutional Review Board of the University Hospital of Groningen.

Experiments

Before we recorded the spectra, volunteers were asked to complete a questionnaire concerning their smoking and drinking habits, most recently consumed food and beverage and the use of any medication. A visual inspection of the oral cavity was performed by an experienced dental hygienist to be certain that no oral lesions were present at the time of measurement. If present, the volunteers were asked to remove their dentures. All volunteers then rinsed their mouth for 1 minute with a 0.9% saline solution in order to minimize the influence of consumed food and beverages.

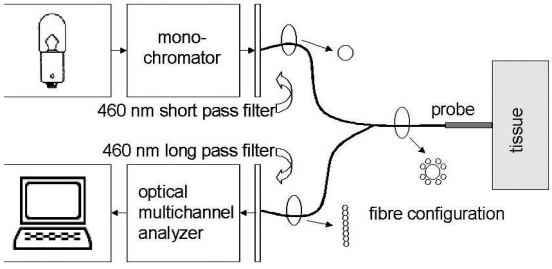

Figure 2.2. The experimental set-up.

The measurement set-up (Figure 2.2) consisted of a Xe-lamp with monochromator (Photomax 60100+ 77250, Oriel Instruments Corporation, Stratford, CT) for illumination and a 1,024 pixel optical multi-channel analyzer (OMA) (Instaspec-IV 77131 CCD-camera+Multispec 125 spectrometer, Oriel Instruments Corporation) for spectroscopy. The light source and OMA were connected to the separate illumination and fluorescence detection arms of a Y-fibre. A measurement probe, consisting of an imaging fibre bundle of 8 mm in diameter, was connected to the end of the third fibre arm where the illumination fibre and fluorescence fibre joined. Tissue excitation wavelengths were 350, 365, 385, 405, 420, 435, and 450 nm (bandwidth ≤15 nm full width half maximum). The total output at the tip of the probe ranged from 20 to 100 μW. The output varied with excitation wavelength due to the grating efficiency and spectral intensity dependence of the lamp output. A 460-nm shortpass filter was used to prevent any long wavelength light, due to higher order transmission of the monochromator, from reaching the mucosa. Fluorescence spectra were recorded in the 455–867 nm range. A 460-nm long pass filter removed any reflected excitation light. The short pass and long pass filters were a custom made filter set, designed to have minimum overlap (Omega Optical, Inc., Brattleboro, VT). Without this filter set the unintended transmission of stray light through the excitation monochromator that was reflected from the tissue surface influenced the shape of measured spectra. Using different filter sets for different

excitation wavelengths would have extended the emission range, but unfortunately for practical reasons this was not possible. However, since the emission spectra of the important tissue fluorophores are very broad, we expect to collect part of their information nonetheless [18].

The measurement probe was disinfected using 2% chlorhexidine digluconate in ethanol and covered with plastic film. The probe was placed in contact with the oral mucosa. The measurements were performed in a completely darkened room to prevent stray light from entering the OMA. In each subject, we measured excitation–emission maps at 13 anatomical positions covering different tissue types. The anatomical locations measured were the cheek, tongue (lateral border, dorsal and ventral side), transition of tongue to floor of mouth, floor of mouth, free mucosa of mandible and maxilla (i.e., mandibular and maxillary fold), gingiva, lower inner lip mucosa, vermilion border (VB) of the lip and soft and hard palate (Figure 2.3). Three experienced dental hygienists, that had been trained by a maxillofacial surgeon in placing the fluorescence probe at well-defined positions, performed the measurements.

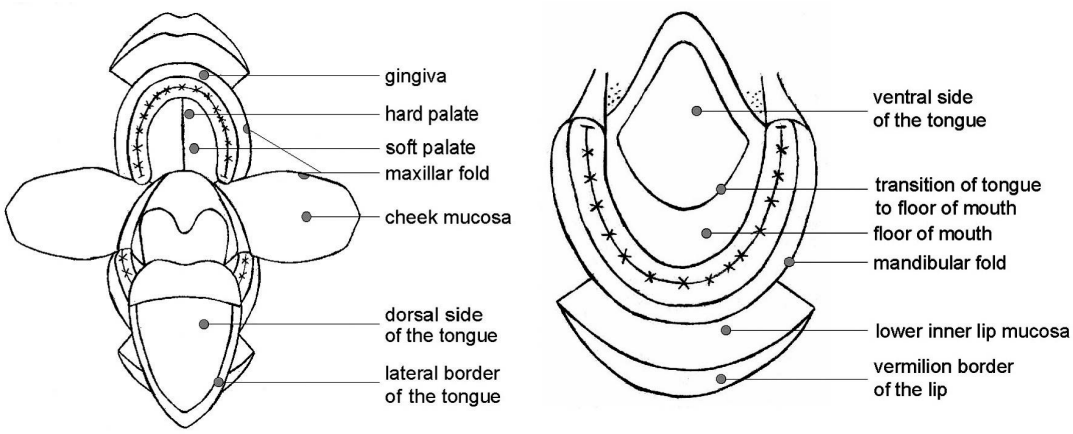

Figure 2.3. The 13 anatomical locations as measured in the volunteers.

For each anatomical location and excitation wavelength, three sequential measurements of 1-second integration time were recorded. This allowed us to remove occasional spectra containing extremely high values for discrete pixels due to electronic noise. On each measurement day, a set of background spectra was recorded for each excitation wavelength by placing the probe in a light-absorbing black cylinder box, thus enabling us to subtract any fluorescence generated by the equipment. The output of a spectral calibration lamp (Hg(Ar) 6035, Oriel Instruments Corporation) and reference spectra of a stable fluorescing standard at different excitation wavelengths were measured to correct for the spectral variation in lamp output as well as for any other varying experimental circumstances.

Data processing

Data preprocessing. Detection wavelength calibration was performed for all spectra using the mercury–argon calibration lamp. Background spectra for the appropriate excitation wavelength were subtracted. The spectral regions below 455 nm and above 867 nm were omitted since they contained no fluorescence. Spectra were averaged per four CCD pixels and eventually consisted of 199 sample points. Averaging the data in this way was justified because the spectral resolution of the

spectrograph was less than 2 nm (867–455 nm/199 points). To be certain that pixel averaging did not influence the analysis, part of it was performed on spectra without averaging measurement points. This did not influence the results significantly.

Data analysis. First the k-means clustering algorithm was applied to all non-normalized autofluorescence spectra as well as to spectra normalized by their mean intensity [19]. We applied the k-means algorithm for 2, 3, and 4 clusters to see if any clear clustering was apparent. As the algorithm was applied to non-normalized as well as to normalized spectra for all seven excitation wavelengths, this resulted in 3x2x7=42 k-means clustering tests. After this, we applied PCA to establish any differences between different anatomical locations. We used three principal components (PCs) of spectra that were normalized by the area under the 498–520 nm part of the spectrum. This was done in order to study spectral shape instead of intensity. PCA was applied for each of the seven excitation wavelengths separately. We examined the spectral shape of the PCs. PC scores for different locations were analyzed using one-way analysis of variance, followed by a multiple comparison test of the means of single PC scores for all locations. The mean value of a PC score for a certain location was considered to be significantly different from that of another location only if the 95% confidence intervals (CI) of the mean values did not overlap.

We then applied leave-one-out classification methods on the first ten PC scores for autofluorescence spectra using three different classifiers. Our purpose is not to classify spectra into anatomical sites, since this is clinically irrelevant. However, a good classification score, i.e., low classification errors, may indicate the presence of spectral differences between anatomical locations. For this reason classification errors give information about the dependence of spectral characteristics on anatomical site. We performed classification methods for the 405 nm excitation wavelength spectra only because a trial analysis showed no significantly different results for different excitation wavelengths. Furthermore, 405 nm excitation is commonly used in autofluorescence spectroscopy [4,20–23]. The spectra were normalized individually by dividing the intensity at each of the 199 spectral samples by the total fluorescence intensity (=the area under the curve) of the spectrum. We chose this method for normalization because in a small trial, it revealed more differences between anatomical sites than non-normalized spectra or spectra that were normalized by the fluorescence intensity around 500 nm. The total fluorescence intensity was added to the data as a 200^{th} feature.

The classifiers used were the Karhunen-Loeve Linear Classifier (KLLC), which is also known as regularized linear classifier assuming normal distributions, the quadratic classifier (QC) assuming normal distributions and a neural network (NN) with one hidden layer, consisting of ten neurons [24–26].

Finally, we assessed the class overlap of autofluorescence spectra of the remaining anatomical location classes that did not seem separable. Assuming normal distribution of locations, we could estimate the class overlap by a classification error obtained by the QC applied to a sequence of two-class problems. In these two-class problems, we compared each location separately to all locations combined.

2.3 Results

General description of the data

We measured 49 men and 47 women with a mean age of 50 years (range 18–85, standard deviation (SD) 16 years). Of the 96 volunteers, six were measured again on another occasion and one subject was measured on three different occasions to study measurement reproducibility. In five volunteers, a total of six locations could not be measured for various reasons, for example retching reflexes and extreme mandibular resorption. Thus, a total of 9,295 autofluorescence spectra was collected. The mean healthy autofluorescence spectra of different anatomical locations were approximately of the same spectral shape for each excitation wavelength. Figure 2.4 shows spectra recorded at 405-nm excitation for four representative locations. Spectra recorded at other wavelengths are not shown because they look very similar. All spectra show dips around 540 and 577 nm that are caused by oxyhemoglobin absorption. A Student's t-test showed significant differences ($P<0.05$) in total fluorescence intensity between almost all different anatomical locations for most of the excitation wavelengths. Averaged over all volunteers, the total fluorescence intensity varied greatly between different locations, up to a factor 2.3. However, the distributions of fluorescence intensities were overlapping.

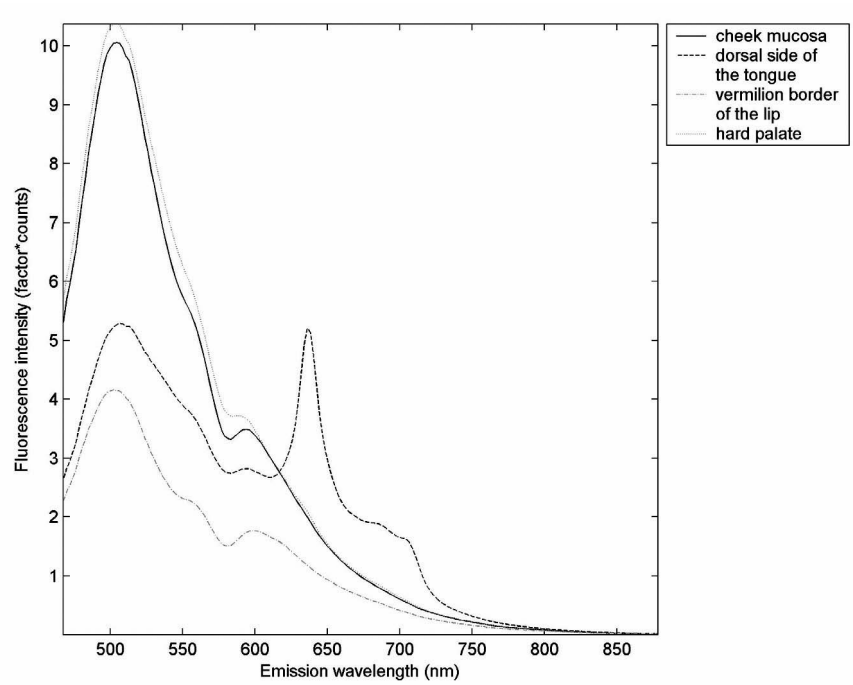

Figure 2.4. Mean autofluorescence spectra of healthy oral mucosa obtained at four representative locations, excitation wavelength 405 nm.

The lowest total intensities were recorded at the VB of the lip and the lateral border of the tongue. Highest intensities were observed for the palatal locations and at the mucosa of the cheek. The mean SD of the total fluorescence intensity of different locations within individuals was

29–49% for separate excitation wavelengths. Besides the variation in fluorescence intensity between different anatomical locations, we observed an even larger variation between the individual volunteers. The average SD of total fluorescence intensities within separate anatomical locations for the group of volunteers was 55–63% for separate excitation wavelengths. The overall SD of total fluorescence intensities at all locations in all volunteers was approximately 86% for each excitation wavelength.

Some anatomical locations produced distinct autofluorescence spectra. For example, in 94% of all healthy volunteers the DST showed a fluorescence peak at 636 nm for at least one of the seven-excitation wavelengths. The highest intensity for this peak was produced at 405 and 420 nm excitation. The 636-nm peak could be as large as or even larger than the autofluorescence intensity maximum around 520 nm. For relatively high 636 nm peak intensities, an accompanying smaller peak centered around 705 nm was observed. The 636/705 peak combination strongly resembled the in vivo emission of Protoporphyrin IX, a photosensitizer commonly used in photodynamic therapy [27]. A peak of the same shape appeared at the lateral border of the tongue in 51% of the volunteers for at least one of the excitation wavelengths. At this location, the peak was generally much smaller than at the DST. In some volunteers, porphyrin-like peaks occasionally appeared at other locations as well. Percentages of volunteers showing these peaks ranged from 1 to 14% for different locations for at least one of the excitation wavelengths. In other locations than the DST, the peaks were generally small. However, in some cases they were as large as at the DST. Locations frequently showing porphyrin-like peaks (>10% of volunteers for at least one of the excitation wavelengths) were free mucosa of the mandible, gingiva, soft and hard palate.

The reproducibility of our measurements was studied using multiple sessions in our volunteers. We found the SD of total autofluorescence intensity between different measurements to be 38% in the combination group of anatomical sites showing the least interlocational intensity variations (floor of mouth, lateral border of tongue, free mucosa of the maxilla, and free mucosa of the mandible).

Statistical results

K-means clustering

The k-means clustering plots showed no obvious clustering with regard to anatomical location. Therefore, we evaluated the cluster labels the algorithm had attached to spectra of the 13 anatomical locations. We calculated the percentage of the dominant label that a location group had received. We then subtracted the percentage it would have received when no information was contained in the spectra, so that all labels would have been attached randomly (i.e., 25% for four clusters). We considered only the cases in which this difference 'D' was larger than 20%. The localizations showing most separate clustering, with values of D over 60%, were the dorsum of tongue and the VB of the lip. In general, larger D arose from normalized spectra than from nonnormalized spectra. In general, excitation wavelengths above 385 nm showed more clustering. For 350 nm excitation, some clustering was found for the hard palate, with four of the six clustering methods for this excitation wavelength showing D between 21 and 26%.

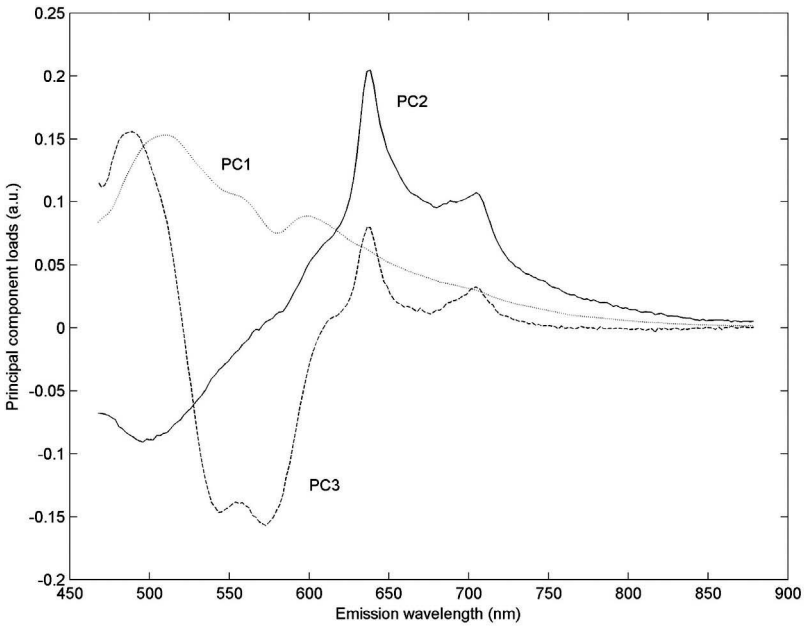

Figure 2.5. The first three principal components (PCs) for autofluorescence spectra of all 13 healthy anatomical locations, excitation wavelength 365 nm. Spectra were normalized by the area under the 498–520 nm part of the spectrum.

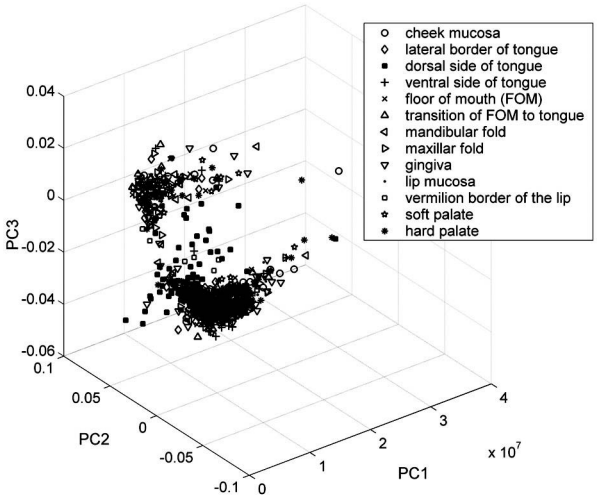

Figure 2.6. The first three PC scores plotted in three-dimensional space for the dataset of all 13 healthy anatomical locations, recorded at excitation wavelength 405 nm. Spectra were normalized by the integrated area under the spectrum, which was added as an extra feature to the data. There is no apparent clustering with regard to anatomical location. PCA can cause symmetrical clouds of points with respect to the coordinate origin; please note that these are not data clusters.

Principal component analysis

Figure 2.5 shows an example of the loadings of the first three PCs for all locations. The PCs shown correspond to an excitation wavelength of 365 nm. PC1 corresponds to the mean autofluorescence spectrum for all locations. PC2 looks similar to the porphyrin-like peak at 636 nm. PC3 mainly seems to show the influence of blood absorption, with absorption maxima around 540 and 575 nm. There appear to be some common features for PC1 and PC2, as well as for PC2 and PC3. The latter similarity is even more pronounced for higher excitation wavelengths. The mean scores for PC2 and PC3 were only 3 and 0.2% of the mean score for PC1, respectively. The three PCs together accounted for 98.8–99.7% of the variance in spectra for the different excitation wavelengths. This means that the data are well described by these first three PCs. Figure 2.6 shows the first three PCs for the dataset at 405-nm excitation. PCA produces symmetrical mappings because the sign of coefficients can be positive or negative. This can cause symmetrical clouds of points with respect to the coordinate origin; however, these are not data clusters. It is clear that the general dataset does not consist of separate groups of objects. This also explains the ineffectiveness of the k-means-clustering algorithm for this dataset.

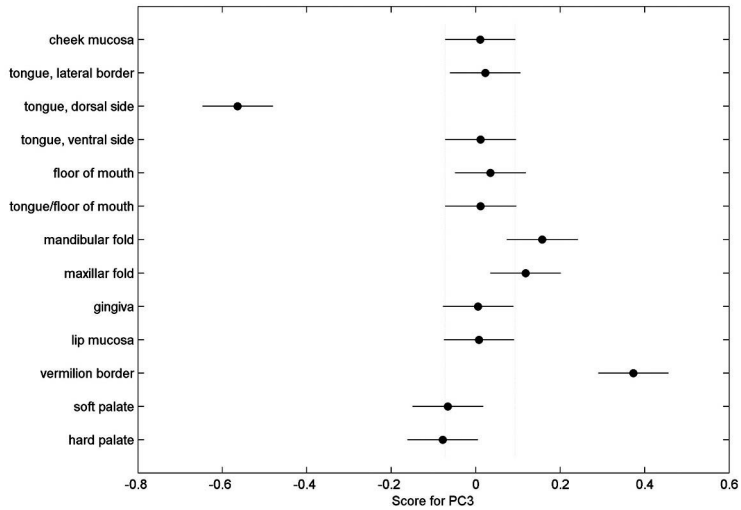

Figure 2.7. Representative example of a multicomparison of PC scores for normalized autofluorescence spectra obtained at 13 sites. Mean PC score values are plotted with 95% confidence intervals (CI). Only non-overlapping CI show significant differences between locations. This example shows scores of the third PC for 435-nm excitation.

We performed a multicomparison analysis of the mean values and 95% CI of the first three PC scores for different anatomical locations. This was done for all seven excitation wavelengths separately. In 86% of all 21 multicomparisons (=three PCs for seven excitation wavelengths), the DST and/or the VB of the lip were the only clear outliers. For all other locations, the 95% CI showed overlap, meaning that the first three PC scores are not statistically different. A representative example is shown in Figure 2.7. To study the differences between autofluorescence spectra obtained at different anatomical sites in more detail, we performed classification of the locations using the first ten PCs. These turned out to preserve 99.99% of the total variance in the data, meaning

that the data are almost completely covered. We studied the classification errors to see whether the autofluorescence spectra of different healthy locations are different. The lower the classification errors are, the more spectral differences must exist between anatomical locations. The results are shown in Table 2.1. In order to know whether the obtained class classification errors show significant differences between data distributions, they have to be compared to the random classification error. If the classification error does not fall within two SD intervals from this random error, the data distributions of PC scores for separate locations are significantly different. In this study, the predicted random classification error can be calculated to be 0.923 (=12/13) with a SD of 0.0302. The 2-STD interval for random classification, therefore, is [0.863; 0.984]. Obtained errors falling outside this interval are marked with (*) in Table 2.1.

Location	KLLC	QC	NN (10 neurons)
Cheek mucosa	0.7051*	0.9103	0.7179*
Lateral border of tongue	0.8462*	0.8846	0.6538*
Dorsal side of tongue	0.2949*	0.3462*	0.1795*
Ventral side of tongue	0.7013*	0.9870	0.4675*
Floor of mouth	0.8974	0.6024*	0.7308*
Transition of tongue to floor of mouth	0.7564*	0.8333*	0.9231
Mandibular buccal fold	0.8333*	0.8974	0.8205*
Maxillar buccal fold	0.8205*	0.9487	0.9487
Gingiva	0.9744	0.9487	0.9744
Lip mucosa	0.8205*	0.4487*	0.8974
Vermilion border	0.4744*	0.5385*	0.1282*
Soft palate	0.8442*	0.8442*	0.7662*
Hard palate	0.5769*	0.7949*	0.3462*

Table 2.1. Leave-One-Out classification errors per class (location) under 405 nm excitation, obtained by the Karhunen-Loeve Linear Classifier (KLLC), the Quadratic Classifier (QC), and the Neural Network (NN) with one hidden layer, consisting of ten neurons. Classification errors with (*) show that there exist significant differences in distributions of the data classes (locations).

The results show that all locations except the gingiva have significantly different distributions in the 10-dimensional subspace spanned by the first ten PCs. All three considered classifiers agree that the DST and the VB of the lip are most easily separated from all other locations. The remaining 11 locations have different distributions for the first ten PCs, which are probably heavily overlapping. To check this, we performed a class overlap estimate.

Class overlap estimate

A multicomparison of the first three PCs and classification trials for the first ten PCs have shown that autofluorescence spectra obtained from the DST and the VB of the lip are clearly different from those obtained from other anatomical locations. The remaining 11 locations seem to be distributed similarly but not identically, since they still are distributed significantly different. The first

three PCs show that these 11 locations are heavily overlapping. In order to test whether they also overlap in the first ten PCs space, the class overlap between these locations is considered for a set of 2-class problems: each of the 11 anatomical locations separately against all 11 locations combined. The results are shown in Table 2.2.

Location class (compared to group of 11 locations)	Class overlap (%)
Cheek mucosa	66
Lateral border of tongue	96
Ventral side of tongue	94
Floor of mouth	84
Transition of tongue to floor of mouth	80
Mandibular fold	87
Maxillar fold	97
Gingiva	96
Lip mucosa	73
Soft palate	79
Hard palate	64

Table 2.2. Leave-One-Out classification error of the QC for the first ten principal components (PCs) of 405 nm excitation autofluorescence spectra. An overlap of more than 18% means that a location overlaps with other locations. A 100% class overlap would indicate complete overlap of the data.

For a 2-class problem the maximum classification error (50%) is obtained when data classes are completely overlapping (100% class overlap). Consequently, the class overlap for a 2-class problem is defined as two times the classification error. If the chosen location overlaps only with itself (represented in the large data class consisting of all 11 locations), then the obtained classification error is equal to 1/11=0.0909 and the class overlap is equal to 2x1/11=0.1818. Therefore, if the obtained class overlap exceeds the value of 0.1818, it means that the selected location overlaps with other locations. Table 2.2 shows that indeed the remaining 11 locations are heavily overlapping (65–97% overlap).

2.4 Discussion

This study shows that most mucosal linings in the oral cavity can spectroscopically be regarded as nearly identical. The fluorescence intensity has its maximum between 500 and 510 nm and is comparable to values found in other studies for excitation at approximately the same wavelength, *in vivo* as well as *in vitro* [10,13,28–30]. A prominent dip in fluorescence emission was seen in most spectra around 575 nm, as well as a less pronounced dip around 540 nm, as has been observed before [10,13,28,30]. From 13 anatomical locations in the oral cavity, 11 were not clearly separable by k-means clustering or by PCA. Only the VB of the lips and the dorsum of the tongue were reliably classified as spectroscopically different. The results need to be carefully interpreted because small differences exist among the anatomical locations. Comparing the first three PCs of the autofluorescence spectra,we found that the DST and the VB of the lip are clear outliers for

almost all excitation wavelengths and PCs. For the remaining 11 locations, the 95% CI for the mean values of the first three PCs overlapped, meaning that they are not statistically different. A refined PCA using the first ten PCs on the 405-nm excitation spectra using two classifiers and a NN showed that spectral shape is statistically different for all anatomical locations except for the gingiva. The DST and the VB of the lip were most easily classified, showing again that they are very different from the other anatomical locations in the oral cavity. Classification error rates of the remaining 11 locations were statistically different from the random classification error. However, class overlap estimates made clear that the 11 locations heavily overlapped (64–97%). This is a result of the fact that the spread of PC scores of spectra recorded within an anatomical location is much larger than the differences between the means of these groups. Although we do not know at this point what the differences in PC scores between benign and (pre-)malignant lesions are, we can conclude that differences in means are irrelevant if the spread within an anatomical location is much larger. This convinces us that these 11 locations can be combined in a reference database for future lesion diagnostics.

For the DST and the VB of the lip, there is no a priori basis to conclude that the differences in mean PC scores between locations are irrelevant. Therefore, initially two separate reference databases will have to be created for these locations.

Previous attempts have been made to identify cancerous lesions in different organs by comparing fluorescence intensities between normal and diseased tissue [1,31,32]. Often intensities at a selected emission wavelength are used or the ratio of two emission wavelength (ranges) is calculated [33–35]. We found that total fluorescence intensities differed significantly between almost all locations for most excitation wavelengths. However, the SD of the total fluorescence intensities within locations were much larger than the differences between locations, so that the intensity distributions of anatomical locations heavily overlapped. This means that the differences in total autofluorescence intensity between different anatomical locations are irrelevant for our purpose. We found that the variation in total fluorescence intensity is large between volunteers for one specific anatomical site (55–63%). Anatomical locations within one specific volunteer showed high fluorescence intensity SD (29–49%) in spite of the fact that these anatomical locations are spectroscopically comparable. This tells us that total fluorescence intensities are highly inconsistent, both intra- and inter-individually. The use of ratios of emission peaks at a single wavelength was not investigated in this study. However, we anticipate that ratio-techniques and statistical methods applied to normalized spectra will be more accurate than techniques that rely on the absolute intensity for their diagnostics. These techniques will reduce interpatient variability as well as irreproducibility, since spectral shape seems more consistent than fluorescence intensity.

There are several possible biological explanations for the large intra- and inter-subject variability in total fluorescence intensity. Fluorescence intensity can be influenced by intersubject variability in the amount of blood, with absorption leading to a wavelength-dependent decrease in fluorescence intensity. Also, variability in the presence of porphyrin-producing microorganisms can influence the total fluorescence intensity by providing an additional amount of fluorescence at 636 nm. Besides the biological variation, varying experimental circumstances can influence the total fluorescence intensity. Although it is our intention to keep the fluorescence probe in the same position for all measurements, variations in positioning will probably account for some part of the variation in fluorescence intensity.

We frequently observed 636 nm peaks that we believe to be caused by porphyrins, which are fluorescing substances produced by living cells. This porphyrin-like peak has been reported by other authors [11,28]. It seems generally accepted that it is caused by endogenous porphyrin pro-

duction. This may be either by cells of the body, or by microorganisms that find a natural habitat in papillary cavities like those at the DST. The porphyrin-like peaks were found most frequently at this location (94% of volunteers). The presence of a porphyrin-like peak at this location probably explains the low error rates in location classification for the DST. The variability in the appearance of porphyrin-like peaks at the lateral border of tongue may be due to small variations in the location of the measurement probe, because the lateral border of the tongue is in fact a transformation zone of papillary to normal flat mucosa. Other locations frequently showing porphyrin-like peaks were the mandibular fold, the gingiva, the soft and the hard palate. We assume the 636 nm peaks at the mandibular fold and the gingiva to be caused by porphyrins appearing in dental plaque [36]. The peaks appearing at the soft and hard palate may be explained by the frequent contact of these locations with the tongue. This is plausible because at least part of the porphyrin-like fluorescence producing substances is located on top of the surface of the tongue. This we have concluded from an experiment in which we tried to diminish the 636-nm peak intensity. Rinsing the mouth with saline had no effect on the height of the peak at the tongue, but rubbing the tongue could reduce the 636-nm peak to about 30–40% of its original height. If all of the fluorescence producing substances were located in the tongue tissue cells, it would have been impossible to diminish the porphyrin-like peak. Porphyrin fluorescence transfer has been observed previously by Harris et al., who found transfer from squamous cell carcinoma's in the hamster cheek pouch to surrounding tissue and a Q-tip [37]. Another explanation for porphyrin-like peaks appearing at other locations than the tongue might be found in bacterial infections that are not noticed clinically [38,39].

Although successes have been obtained in diagnosing malignancies searching for porphyrin-like peaks appearing in these lesions, we believe that this method is not reliable for the oral cavity [9,11]. The presence of porphyrin-like peaks in autofluorescence spectra of healthy oral mucosa, most frequently on the DST but also at all other sites, will lead to false positive classifications.

In three volunteers, we observed fluorescence peaks in the 670–680 nm emission wavelength range that were about as high as the fluorescence intensity at 520 nm. These 670–680 nm peaks appeared for excitation wavelengths 405–450 nm and occurred at the floor of mouth in one volunteer and at the mandibular fold in two volunteers. The 670–680 nm peaks might be caused by pheophorbide A and/or pheophytin A, degradation products of chlorophyll A that have been shown to cause the 674 nm fluorescence peak in mouse skin [40]. These substances might be present in the oral cavity due to recently consumed food, for example, leafy vegetables.

For future lesion diagnostics, we plan to investigate the use of our reference database for comparison. Different anatomical locations showed significantly different autofluorescence intensities. However, the interpatient variability was much larger. We presume that, because of its large variation between both anatomical locations and volunteers, the total fluorescence intensity can not easily be used as a method of diagnosing tumour tissue when using a reference database.

We have concluded that all anatomical locations, except the DST and the VB of the lip, are interchangeable for the purpose of autofluorescence spectroscopy for the detection of (pre)malignant lesions. For our lesion diagnostics algorithm development, we will compare the results for methods using reference spectra recorded at the contralateral anatomical position in the same patient and for methods in which we compare the lesion spectra to our reference database. For the first type of method, we can probably use not only spectra of contralateral tissue in the same patient, but spectra of any anatomical location that we have found to be interchan-

geable with it. For the second type of method,we will use separate reference databases for the DST and the VB of the lip. We will study whether the use of a reference database improves the sensitivity and specificity of early cancer detection.

2.5 Acknowledgements

We thank Tom Bakker Schut for allowing us to use his implementation of the k-means clustering algorithm. We are grateful to all volunteers for their valuable contribution and to Mirjam Wouda, Irènke de Jong-Orosz, and Ada Schokkenbroek for performing the many measurements.

2.6 References

1. D'Hallewin MA, Baert L, Vanherzeele H. In vivo fluorescence detection of human bladder carcinoma without sensitizing agents. J Am Paraplegia Soc 1994;4:161–164.
2. Georgakoudi I, Sheets EE, Muller MG, Backman V, Crum CP, Badizadegan K, Dasari RR, Feld MS. Trimodal spectroscopy for the detection and characterization of cervical precancers in vivo. Am J Obstet Gynecol 2002;3:374–382.
3. Gillenwater A, Jacob R, Richards-Kortum R. Fluorescence spectroscopy: A technique with potential to improve the early detection of aerodigestive tract neoplasia. Head Neck 1998;6:556–562.
4. Zellweger M, Grosjean P, Goujon D, Monnier P, van den BH, Wagnieres G. In vivo autofluorescence spectroscopy of human bronchial tissue to optimize the detection and imaging of early cancers. J Biomed Opt 2001;1:41–51.
5. Silverman S. Early diagnosis of oral cancer. Cancer 1988; 8(Suppl):1796–1799.
6. Hyde N, Hopper C. Oral cancer: The importance of early referral. Practitioner 1999;1603:753, 756–761.
7. Betz CS, Stepp H, Janda P, Arbogast S, Grevers G, Baumgartner R, Leunig A. A comparative study of normal inspection, autofluorescence, and 5-ALA-induced PPIX fluorescence for oral cancer diagnosis. Int J Cancer 2002;2:245–252.
8. Chen CT, Wang CY, Kuo YS, Chiang HH, Chow SN, Hsiao IY, Chiang CP. Light-induced fluorescence spectroscopy: A potential diagnostic tool for oral neoplasia. Proc Natl Sci Counc Repub China B 1996;4:123–130.
9. Gillenwater A, Jacob R, Ganeshappa R, Kemp B, El Naggar AK, Palmer JL, Clayman G, Mitchell MF, Richards-Kortum R. Noninvasive diagnosis of oral neoplasia based on fluorescence spectroscopy and native tissue autofluorescence. Arch Otolaryngol Head Neck Surg 1998;11:1251–1258.
10. Heintzelman DL, Utzinger U, Fuchs H, Zuluaga A, Gossage K, Gillenwater AM, Jacob R, Kemp B, Richards-Kortum RR. Optimal excitation wavelengths for in vivo detection of oral neoplasia using fluorescence spectroscopy. Photochem Photobiol 2000;1:103–113.
11. Inaguma M, Hashimoto K. 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 1999;11:2201–2211.
12. Onizawa K, Saginoya H, Furuya Y, Yoshida H. Fluorescence photography as a diagnostic method for oral cancer. Cancer Lett 1996;1:61–66.
13. van Staveren HJ, van Veen RL, Speelman OC, Witjes MJ, Star WM, Roodenburg JL. Classification of clinical autofluorescence spectra of oral leukoplakia using an artificial neural network: A pilot study. Oral Oncol 2000;3:286–293.
14. Wang CY, Chen CT, Chiang CP, Young ST, Chow SN, Chiang HK. A probability-based multivariate statistical algorithm for autofluorescence spectroscopic identification of oral carcinogenesis. Photochem Photobiol 1999;4:471–477.
15. Slaughter D, Southwick H, Smejkal W. Field cancerization in oral stratified squamous epithelium. Cancer 1953;6:963–968.
16. Kolli VR, Shaha AR, Savage HE, Sacks PG, Casale MA, Schantz SP. Native cellular fluorescence can identify changes in epithelial thickness in-vivo in the upper aerodigestive tract. Am J Surg 1995;5:495–498.
17. Savage HE, Kolli V, Ansley J, Chandawarkar RY, Alfano RR, Schantz SP. Innate tissue fluo-

rescence of the oral mucosa of controls and head and neck cancer patients. Proc SPIE 1995;2387:2–9.
18. Drezek R, Sokolov K, Utzinger U, Boiko I, Malpica A, Follen M, Richards-Kortum R. Understanding the contributions of NADH and collagen to cervical tissue fluorescence spectra: Modeling, measurements, and implications. J Biomed Opt 2001;4:385–396.
19. MacQueen J. Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, USA: University of Berkeley Press; 1967. pp 297–281.
20. Eker C, Montan S, Jaramillo E, Koizumi K, Rubio C, Andersson-Engels S, Svanberg K, Svanberg S, Slezak P. Clinical spectral characterisation of colonic mucosal lesions using autofluorescence and delta aminolevulinic acid sensitisation.Gut 1999;4:511–518.
21. Eker C, Rydell R, Svanberg K, Andersson-Engels S. Multivariate analysis of laryngeal fluorescence spectra recorded in vivo. Lasers Surg Med 2001;3:259–266.
22. Sterenborg HJ, Saarnak AE, Frank R, Motamedi M. Evaluation of spectral correction techniques for fluorescence measurements on pigmented lesions in vivo. J Photochem Photobiol B 1996;3:159–165.
23. Vengadesan N, Aruna P, Ganesan S. Characterization of native fluorescence fromDMBA-treated hamster cheek pouch buccal mucosa for measuring tissue transformation. Br J Cancer 1998;3:391–395.
24. Fukunaga K. Introduction to statistical pattern recognition. San Diego, USA: Academic Press; 1990.
25. Raudys S, Skurichina M. 1994. Small sample properties of ridge estimate of the covariance matrix in statistical and neural net classification. Proceedings of the 5th Tartu Conference. New trends in probability and statistics: Multivariate
statistics and matrices in statistics. pp 237–245.
26. Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Netw 1989; 2:359–366.
27. Robinson DJ, de Bruijn HS, de Wolf WJ, Sterenborg HJ, Star WM. Topical 5-aminolevulinic acid-photodynamic therapy of hairless mouse skin using two-fold illumination schemes: PpIX fluorescence kinetics, photobleaching, and biological effect. Photochem Photobiol 2000;6:794–802.
28. Betz CS, Mehlmann M, Rick K, Stepp H, Grevers G, Baumgartner R, Leunig A. Autofluorescence imaging and spectroscopy of normal and malignant mucosa in patients with head and neck cancer. Lasers Surg Med 1999;4:323–334.
29. Ingrams DR, Dhingra JK, Roy K, Perrault DF, Jr., Bottrill ID, Kabani S, Rebeiz EE, Pankratov MM, Shapshay SM, Manoharan R, Itzkan I, Feld MS. Autofluorescence characteristics of oral mucosa. Head Neck 1997;1:27–32.
30. Roy K, Bottrill ID, Ingrams DR, Pankratov MM, Rebeiz EE, Woo P, Kabani S, Shapshay S, Manoharan R, Itzkan I, Feld MS. Diagnostic fluorescence spectroscopy of oral mucosa. SPIE Proc 1996;2395:135–142.
31. Zargi M, Smid L, Fajdiga I, Bubnic B, Lenarcic J, Oblak P. Laser induced fluorescence in diagnostics of laryngeal cancer. Acta Otolaryngol Suppl Stockh 1997;527:125–127.
32. Lohmann W, Mussmann J, Lohmann C, Kunzel W. Native fluorescence of the cervix uteri as a marker for dysplasia and invasive carcinoma. Eur J Obstet Gynecol Reprod Biol 1989;3:249–253.
33. Bogaards A, Aalders MC, Zeyl CC, de Blok S, Dannecker C, Hillemanns P, Stepp H,

Sterenborg HJ. Localization and staging of cervical intraepithelial neoplasia using double ratio fluorescence imaging. J Biomed Opt 2002;2:215–220.
34. Wang CY, Lin JK, Chen BF, Chiang HK. Autofluorescence spectroscopic differentiation between normal and cancerous colorectal tissues by means of a two-peak ratio algorithm. J Formos Med Assoc 1999;12:837–843.
35. Wang CY, Chiang HK, Chen CT, Chiang CP, Kuo YS, Chow SN. Diagnosis of oral cancer by light-induced autofluorescence spectroscopy using double excitation wavelengths. Oral Oncol 1999;2:144–150.
36. Konig K, Schneckenburger H, Hibst R. Time-gated in vivo autofluorescence imaging of dental caries. Cell Mol Biol (Noisy-le-grand) 1999;2:233–239.
37. Harris DM, Werkhaven J. Endogenous porphyrin fluorescence in tumours. Lasers Surg Med 1987;6:467–472.
38. Konig K, Teschke M, Sigusch B, Glockmann E, Eick S, Pfister W. Red light kills bacteria via photodynamic action. Cell Mol Biol (Noisy-le-grand) 2000;7:1297–1303.
39. Brazier JS. Analysis of the porphyrin content of fluorescent pus by absorption spectrophotometry and high performance liquid chromatography. J Med Microbiol 1990;1:29–34.
40. Weagle G, Paterson PE, Kennedy J, Pottier R. The nature of the chromophore responsible for naturally occuring fluorescence in mouse skin. J Photochem Photobiol B 1988;2:313–320.