Autofluorescence spectroscopy for the classification of oral lesions

Veld, Diana Cornelia Gerarda de

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2005

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Chapter 3

Effects of individual characteristics on healthy oral mucosa autofluorescence spectra

D.C.G. de Veld 1,2
H.J.C.M. Sterenborg 2
J.L.N. Roodenburg 1
M.J.H. Witjes 1

1Department of Oral and Maxillofacial Surgery, University Hospital Groningen
2Photodynamic Therapy and Optical Spectroscopy Programme, Erasmus Medical Center, Rotterdam

Oral Oncol 2004 Sep;40(8):815-23
Chapter 3

Summary

Background and Objectives: Autofluorescence spectroscopy is a tool for detecting tissue alterations in vivo. In our previous study, we found spectral differences between clinically normal mucosa of different patient groups. These are possibly caused by associated patient characteristics. In the present study, we explore the influences of volunteer characteristics on healthy oral mucosa autofluorescence.

Study Design/Materials and Methods: Autofluorescence spectra were recorded in 96 volunteers with no clinically observable oral lesions. We applied principal components analysis to extract the relevant information. We used multivariate linear regression techniques to estimate the effect of volunteer characteristics on principal component scores.

Results: Statistically significant differences were found for all factors but age. Skin color strongly affected autofluorescence intensity. Gender differences were found in blood absorption. Alcohol consumption was associated with porphyrin-like peaks. However, all differences but those associated with skin color were of the same order of magnitude as standard deviations within categories.

Conclusions: The effects of volunteer characteristics on autofluorescence spectra of the oral mucosa are measurable. Only the effects of skin color were large. Therefore, in lesion classification, skin color should be taken into account.

3.1 Introduction

Autofluorescence spectroscopy is a non-invasive tool for the detection of alterations in the structural and chemical composition of cells[1-4]. Autofluorescence of tissues is produced by several endogenous fluorophores. These include fluorophores from tissue matrix molecules and intracellular molecules like collagen, elastin, keratin and NADH. The presence of disease changes the concentration of these fluorophores as well as the light scattering and absorption properties of the tissue, due to changes in blood concentration, nuclear size distribution, epithelial thickness and collagen content. It is therefore expected that the presence of disease will be identifiably reflected in autofluorescence spectral shape and intensity.

In our previous study on lesion classification, we observed statistically significant spectral differences between clinically normal mucosa of oral cancer patients and clinically normal mucosa of patients with benign lesions[5, chapter 4 of this thesis]. This suggests that autofluorescence spectroscopy is not only sensitive to disease, but also to other tissue alterations. Our hypothesis is that the non-pathological tissue alterations that cause these differences in autofluorescence characteristics correspond to characteristics of the individual, such as tobacco and alcohol consumption, age and wearing of dentures.

The inflammatory effects of nicotine on oral mucosa have been demonstrated in vitro[6]. Epithelial atypia of the oral mucosa is correlated with smoking habits, while epithelial differentiation due to nicotine has been demonstrated in vitro[7;8]. Pigmentation effects of smoking were observed, as well as correlations between smoking behavior and leukoedema[9–11]. Keratinization indexes of the oral mucosa change due to smoking[12] and an increased cell proliferation index is associated with smoking behavior[13].

Alcohol has been found to increase the permeability of the oral mucosa in animals and humans[14–16]. Reduction in total cell area, mean cytoplasmic and mean nuclear area were observed in humans with a history of chronic alcohol use[17;18]. Effects of chronic alcohol use on the
oral mucosa in rodents include, amongst others, a thickening of the basal cell layer and basal cell pleomorphism, enlarged cell nuclei and a tendency towards epithelial dysplasia[19–21].

Denture wearing is associated with stomatitis, angular cheilitis, traumatic ulcers, denture irritation hyperplasia and flabby ridges[22]. The wearing of dentures is associated with changes in keratinisation and epithelial thickness, inflammatory changes, and the replacement of collagen fibers with oxytalan fibers[23–25]. In skin, the autofluorescence characteristics alter with age and possibly this could happen with the oral mucosa as well[26].

To understand the differences in healthy oral mucosa autofluorescence spectra between different patient groups, we searched for correlations between characteristics of the individual and spectral properties in our healthy volunteers reference database. This was done by applying multivariate linear regression techniques on principal components (PC) scores, which contain the most relevant information from the recorded autofluorescence spectra. The multivariate linear regression technique was chosen because it reduces the risk of attributing the observed spectral differences to variables that are merely correlated to the true causal factors.

3.2 Materials and methods

Volunteer and patient population

Autofluorescence spectra were collected from 96 healthy volunteers with no clinically observable lesions of the oral mucosa as described in our previous study[27, chapter 2 of this thesis]. 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 same department because of conditions that did not affect the oral mucosa. This study was approved by the Institutional Review Board of the University Hospital of Groningen.

Experiments

Before recording the spectra, volunteers were asked to complete a questionnaire concerning their tobacco and alcohol consumption and most recently consumed food and beverage. A visual inspection of the oral cavity was performed by an experienced dental hygienist to ensure that no oral lesions were present at the time of measurement. The presence of any dentures was listed, as well as the volunteers’ date of birth, skin color and medical history relating to the oral cavity. If present, the volunteers and patients were asked to remove their dentures. All patients and volunteers rinsed their mouth during 1 minute with a 0.9% saline solution in order to minimize the influence of consumed food and beverages.

The measurement set-up (Figure 2.1), as described in detail in Chapter 2, consisted of a Xe lamp with monochromator, a spectrograph and a custom-made set of 460 nm longpass and shortpass filters[27]. Tissue excitation wavelengths were 365, 385, 405, 420, 435 and 450 nm. 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 expected to collect at least part of the relevant information[28;29].

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 spectrograph. A dental hygienist performed the measurements.
For each measured 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 calibration measurements was performed.

**Data processing**

Data preprocessing was performed as described in our previous study[27]. Preprocessed spectra consisted of 199 data points, covering the 467–867 nm emission range. We used two different normalization methods, which comprised normalization by the area under the curve and by the fluorescence intensity in the 502–523 nm range. The 502–523 nm spectral region corresponds to the peak autofluorescence intensity of the spectra when porphyrin-like peaks around 635 nm are not considered. We performed a principal components analysis (PCA) for each excitation wavelength separately. PCA searches for the orthogonal components of the spectra that account for the maximum of variance within the complete dataset. The first four principal components covered >99% of the variance, and therefore these were saved as spectrum representatives for each spectrum.

**Statistical analysis**

Because of their distinct spectral characteristics, we excluded spectra recorded from the vermilion border of the lip and the dorsal side of the tongue from all analyses[27]. We applied multivariate linear regression techniques for evaluating the contributions of different volunteer variables to PC scores. Tobacco and alcohol consumption, gender, denture wearing and skin color were considered as nominal values, while age was included as a continuous scale variable. To be less sensitive to outliers, the PC scores were truncated to exclude the upper and lower 5% of values. Since we study the prevalence of general trends, this data truncation is not disturbing. Differences with p<0.01 were considered as statistically significant.

**3.3 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 16 years). Ninety-one volunteers were Caucasian, 1 was Asian, 1 was Latin-American and 3 were Negroid. The medication used was diverse, most common were analgesics and antidepressants. These medications were not expected to have a direct influence on autofluorescence characteristics.

Of the volunteers, 3 were frequent users of inhalers for asthmatic diseases, and 7 volunteers used inhalers occasionally. Of the 96 volunteers, 5 were measured again on another occasion and one subject was measured on three different occasions.

The recently consumed food categories were too diverse to allow for a reliable statistical analysis and therefore were not considered. In five volunteers, a total of six locations could not be measured. In total, 9295 autofluorescence spectra were collected. The volunteer characteristics are described in Table 3.1. The general shape and intensity of the spectra for the group as a whole have been described in Chapter 2 of this thesis.
Effects of individual characteristics on healthy oral mucosa autofluorescence spectra

Statistical analysis
Normalization by the 502–523 nm intensity gave rise to 17% more statistically significant differences between PC scores. Therefore, we decided to summarize the results for 502–523 nm normalization only.

Since we used six excitation wavelengths and four PCs, a maximum of $6 \times 4 = 24$ statistically significant differences could be found for each comparison between two category groups. We expressed the number of statistically significant differences for each comparison as a percentage of 24, so that 0% means that no significant differences were found, and 100% means that all PC scores for all excitation wavelengths were significantly different.

Influence of smoking
The results of the multivariate analysis of the PC scores are summarized in Table 3.2. The most significant differences were found between non-smokers and smokers (Table 3.2), especially for those consuming 10–20 cigarettes a day (21%) and heavy smokers (29%). The most differences were found for excitation wavelengths 435 and 450 nm. We examined the spectral shape of the PC loadings that gave rise to the observed differences. In most cases, these PC loadings turned out to correspond to the average autofluorescence spectrum (45% of significantly different PC scores) or porphyrin-like peaks (35%). In Figure 3.1, the average normalized autofluorescence spectra at 405 nm excitation are plotted for the different smoking behavior categories. Please note that in this non-multivariate figure the volunteers may differ not only in their tobacco consumption habits, so that we possibly look at associated factors as well. However, because of the relatively large numbers of volunteers in all smoking groups, we assume these to average out to a large extent. We can see from Figure 3.1 that an increase in tobacco consumption is correlated with a relative increase in red autofluorescence. We could observe no correlation between smoking behavior and absolute fluorescence intensity. Hardly any differences were observed between non-smokers and ex-smokers. Since the sample size of ex-smokers was larger than that of the separate smoking groups, for which we did find significant differences, there truly are few differences between non-smokers and ex-smokers in our spectra. Apparently, the effects of smoking on autofluorescence spectra recorded from the oral mucosa are of a temporary nature. We also observed

<table>
<thead>
<tr>
<th>Smoking habits</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-smokers</td>
<td>39</td>
</tr>
<tr>
<td>1-10 cig./day</td>
<td>14</td>
</tr>
<tr>
<td>10-20 cig./day</td>
<td>13</td>
</tr>
<tr>
<td>&gt; 20 cig./day</td>
<td>11</td>
</tr>
<tr>
<td>Former smokers</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alcohol consumption</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 5 alcoholic units/day</td>
<td>84</td>
</tr>
<tr>
<td>&gt; 5 alcoholic units/day</td>
<td>6</td>
</tr>
<tr>
<td>Former alcoholics</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dental prosthesis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No dentures</td>
<td>48</td>
</tr>
<tr>
<td>Lower dentures</td>
<td>3</td>
</tr>
<tr>
<td>Upper dentures</td>
<td>7</td>
</tr>
<tr>
<td>Full dentures</td>
<td>38</td>
</tr>
</tbody>
</table>

Table 3.1. Volunteer characteristics in terms of the categories that were included in the multivariate regression analysis.
a reasonable number of significant differences between heavy smokers and ex-smokers, which confirms the observation that ex-smokers’ autofluorescence characteristics are similar to those of non-smokers.

<table>
<thead>
<tr>
<th></th>
<th>Non-smokers</th>
<th>&lt;10 cig./day</th>
<th>10-20 cig./day</th>
<th>&gt;20 cig./day</th>
<th>Ex-smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-smokers</td>
<td>-</td>
<td>0</td>
<td>21</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>&lt;10 cig./day</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10-20 cig./day</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>&gt;20 cig./day</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>Ex-smokers</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3.2. Percentage of significantly different PC scores between smoking habits categories.

Although many statistically significant differences were observed between the smoking behavior categories, these differences can still be small compared to variations within the categories. To test this hypothesis, we calculated standard deviations and average values of PC scores for different categories. We found that for the comparisons giving the most significant differences (non-smokers vs. intermediate and heavy smokers), the standard deviations of the PC scores within the categories were of the same order of magnitude as the differences between these scores. For the categories showing fewer significant differences, standard deviations within categories were approximately 4–7 times larger than differences between categories. These findings make it clear that it will not be possible to perform classification of an individual’s smoking behavior by recording autofluorescence spectra.

**Influence of alcohol consumption**

In Figure 3.2, the average normalized autofluorescence spectra at 405 nm excitation are plotted for the three alcohol consumption behavior categories. Again, please note this univariate plot might display the influence of other, associated factors, like tobacco consumption, as well. The
results of the multivariate analysis of PC scores are summarized in Table 3.3. Most information about alcohol consumption was contained in the PCs resembling the porphyrin-like peak (50% of significantly different PC scores). The most relevant excitation wavelength was 405 nm, which is the peak excitation wavelength for porphyrin. For the comparisons showing the most significant differences, the standard deviations of the PC scores within the categories were about 0.65 times the differences between these scores. Again, reliable classification of alcohol consumption of individuals will not be possible.

![Figure 3.2](image)

*Figure 3.2. Mean autofluorescence spectra recorded at 405 nm excitation for the three alcohol consumption categories. All individual spectra were normalized by the 502-523 nm intensity. The right part of the spectrum rises with alcohol consumption. For the average individual consuming over five alcoholic beverages a day, a clear porphyrin-like peak is observed.*

<table>
<thead>
<tr>
<th></th>
<th>&lt;5 Alcohol consumptions/day</th>
<th>&gt;5 Alcohol consumptions/day</th>
<th>Ex-alcoholics</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;5 Alcohol consumptions/day</td>
<td>-</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>&gt;5 Alcohol consumptions/day</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Ex-alcoholics</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 3.3. Percentage of significantly different PC scores between alcohol consumption habits categories.*

**Influence of denture wearing**

We found statistically different PC scores between upper denture wearers and volunteers without dentures (42%). A large amount of differences was present between upper denture wearers and full denture wearers as well (29%). Differences seemed to be associated mainly to blood absorption (75% of significantly different PC scores). Most significant differences were observed for 365 and 435 nm excitation. To test the influence of anatomical location on the amount of differences, we repeated the calculations while omitting each of the 11 anatomical locations in turn. This resulted in almost equal amounts of significant differences, except for the mandibular and maxillary fold. For these two locations, the omission resulted in a 25% decrease of significant differences. The overall conclusion is that the information about denture wearing is not specifically contained within autofluorescence spectra recorded at any specific location. However, the influence of upper
dentures is slightly more prominent in the mandibular and maxillary folds.

**Influence of skin color**

Despite the small number of subjects in the dark skin color group, significant differences were very frequently found between Caucasian and dark skin color volunteers (67% of PC comparisons), as well as between lightly colored and dark skin volunteers (42%) (Figure 3.3). This is probably caused by mucosal pigmentation for dark skin type volunteers, which strongly influences the spectral shape by absorbing both excitation and emission light. The differences were mainly contained in the PC resembling the average autofluorescence spectrum. Most differences were found for 365, 385 and 450 nm excitation. For the significantly different PC scores, the standard deviations within dark and Caucasian skin color categories were about half the size of the differences between the categories.

![Image of mean autofluorescence spectra](image)

**Figure 3.3.** Mean autofluorescence spectra recorded at 405 nm excitation for the three skin color categories. All individual spectra were normalized by the 502–523 nm intensity. The right part of the spectrum is raised for pigmented skin. This is probably caused by relatively more absorption in the left part of the spectrum.

**Influence of gender and age**

The multivariate analysis led to 33% significant differences between male and female volunteers. These differences could be attributed mainly to blood absorption (67% of significantly different PC scores). There was no dependence on excitation wavelength. Hardly any dependence of PC scores on age could be observed.

3.4 Discussion

We studied autofluorescence spectral dependence on volunteer characteristics using a multivariate analysis on PC scores and found many significant differences for spectra normalized by their 502–523 nm intensity. However, standard deviations within groups were relatively large, which makes classification impossible. The first four PC loadings of the complete dataset for all excitation wavelengths resembled the average autofluorescence spectrum, blood absorption dips, porphyrin-like peaks and mixtures of these. This means that these features account for most of the variance within the volunteer population. Most significant differences observed between volunteer characteristics groups could be attributed to porphyrin-like peaks (40%), followed by PCs resembling the average autofluorescence spectrum (25%) and blood absorption (23%). The remaining significant differences corresponded to two PCs, both occurring at 450 nm excitation, that could
Effects of individual characteristics on healthy oral mucosa autofluorescence spectra

not be reduced to any known autofluorescence features.

In terms of excitation wavelength, the lowest excitation wavelength (365 nm) and the two highest ones (435 and 450 nm) were responsible for most of the observed differences. We found statistically significant differences in autofluorescence characteristics to be associated with skin color, tobacco and alcohol consumption, the wearing of upper dentures and gender. Because this was a volunteer study, we did not take any tissue samples from the volunteers, which makes it difficult to attribute the found differences to any histological changes. Autofluorescence spectra are inherently complex and cannot be reduced reliably to constituting fluorophore spectra, and therefore are not suitable for clear explanations of the observed significant differences. Therefore, only the shape of the PC loadings that are responsible for the differences can give us some information about the origins.

Tobacco consumption mainly had an effect on the score for the PC resembling the average autofluorescence spectrum. We could not interpret changes in this PC score in terms of histological alterations. The non-multivariately analyzed normalized spectra showed an increase in red autofluorescence with tobacco consumption. In lesions, a relatively large contribution of red autofluorescence is frequently observed as well[30;31]. In these cases, this is caused by a decrease in green autofluorescence, probably due to thickening of the epithelial layer. In our study, however, smoking consumption was not associated with a decrease in absolute green autofluorescence, since the spectra of smokers had a higher absolute intensity for all emission wavelengths. Apparently, the effect of smoking that is responsible for changes in autofluorescence is not epithelial thickening. To our surprise, we found that the effect of smoking on autofluorescence characteristics was of a temporary nature.

The effects of alcohol consumption habits were noticed when comparing <5 and >5 consumptions/day categories, and were associated with porphyrin-like peaks. Possibly, the regular presence of alcoholic beverages in the oral cavity shifts the oral flora balance in such a way that porphyrin-producing micro-organisms are more likely to prevail. Also, chronic alcohol abuse is associated with a diminished care for oral hygiene[32;33], which augments the number of micro-organisms.

Denture wearing affected autofluorescence characteristics, especially in the case of upper dentures with own lower teeth. Possibly, in this case masticatory forces are large and therefore induce measurable changes to the supportive soft tissue. We expected to find influences of keratinization, reflected in fluorescence intensity. The main correlation found, however, was with blood absorption. Possibly, the thickening of the mucosa by keratinization in the case of denture wearing reduces the spectral influence of blood absorption. The effects were not specific for the gingiva.

In skin, the alterations in fluorescence spectra with age are commonly attributed to photoaging due to UV exposure. For the oral mucosa, we found hardly any significant influence of age on autofluorescence spectra, which might be explained by the low amount of daylight reaching the oral mucosa.

Gender did have a significant influence on autofluorescence spectra. To our surprise, males’ autofluorescence spectra showed more blood absorption than females’. The average autofluorescence spectra for men and women show equal peak intensities, indicating that the higher blood absorption is not a normalization artifact. The observed differences can perhaps be explained by the differences in vascularization between men and women or by a difference in epithelial thickness. We could not find substantial evidence for these hypotheses in the literature. However, differences between male and female near-infrared spectra recorded from the skin of the forearm have been observed[34]. These differences could be attributed to a smaller peak at 1212 nm (lipid) and a larger absorption peak at 578 nm (blood absorption) in male volunteers. The 1212 nm peak falls
outside our recording range, but the observation of more blood absorption in males is equal to our own. This could mean that similar differences in blood absorption between males and females exist in skin as well as in oral mucosa. Because of the multivariate analysis applied, the established differences cannot be produced by any gender-correlated factor that was itself included in the analysis, such as smoking behavior. However, non-included and therefore irretrievable gender-correlated factors might be of influence.

The effects of skin color on autofluorescence spectroscopy were very prominent, even though the number of volunteers with a dark skin color was only 3. These differences are caused by mucosal pigmentation, which absorbs excitation and fluorescence light. When autofluorescence is applied for lesion diagnostics, the presence of mucosal pigmentation should be taken into account. In one volunteer with a dark skin color, the autofluorescence was almost completely absorbed due to oral pigmentation. For patients with strong oral pigmentation, autofluorescence spectroscopy may therefore not be applicable.

The differences resulting from volunteer characteristics help to explain the large variation observed in autofluorescence spectra of healthy oral mucosa, that complicates lesion diagnostics[3;27]. However, the observed differences are generally small—in the same order of magnitude as the standard deviations within categories. This means that it will not be possible to discriminate volunteer characteristics by means of autofluorescence spectroscopy.

Although the differences associated with all other factors are small, they can still provide interesting information about the effects of individual characteristics on oral mucosa autofluorescence for groups of people. We believe that the differences observed between clinically normal mucosa of different patient groups as observed in our previous study, is at least partially caused by differences in alcohol and tobacco consumption, gender, and wearing of dentures between these groups. However, we believe that it will be impossible to correct for interindividual variations when classifying lesions, since a very large part of the variation in the population was not associated with the factors studied. It seems improbable that other etiological factors can be found that can help to explain this variation. Possibly, the variations are related to different mucosa types for individuals, that may be defined genetically. The only parameter that explained enough of the variation to be taken into account in lesion classification, was skin color.

3.5 Acknowledgements

This work was supported by the Dutch Cancer Society (“Nederlandse Kanker Bestrijding”), grant RUG-99-1869. We are grateful to all volunteers and patients for their valuable contribution and to Mirjam Wouda, Irène de Jong-Orosz and Ada Schokkenbroek for performing the many measurements. Wilfried Graveland has supplied very valuable assistance in performing the multivariate linear regressions.
3.6 References


