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):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
The status of in vivo autofluorescence spectroscopy and imaging for oral oncology

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

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

Summary

Autofluorescence spectroscopy and imaging have been studied for the early detection and classification of (pre)malignancies of the oral mucosa. In the present review we will give an overview of the literature on autofluorescence imaging and spectroscopy for various clinical questions. From the studies performed so far we hope to conclude whether autofluorescence spectroscopy and imaging are helpful in the diagnosis of lesions of the oral mucosa, and if this is the case: for which clinical questions they are suitable. A strong emphasis is put on in vivo human studies of the oral mucosa.

7.1 Introduction

The common procedure for detecting (pre)malignant lesions consists of visual inspection, followed by biopsy of any suspicious lesions found. However, benign lesions—which are very common and diverse (lichen planus, candida infections, inflammation, hyperkeratosis, ulcerations and so on)—may present very similar to early malignant or premalignant lesions, which makes it difficult to distinguish them even for experienced clinicians. Therefore, a technique that can distinguish between different lesion types in a reliable and non-invasive way would be very useful. Such a device providing in vivo lesion classification would be particularly useful for finding the optimal—i.e. most dysplastic—biopsy site, so that the risk of underdiagnosis and need for repeated biopsies is avoided. Another important clinical improvement would be to detect premalignant changes and lesions in an earlier stage, preferably before visual detection is possible. It has been claimed that autofluorescence spectroscopy and imaging can assist in oral oncology for the detection and classification of lesions. In this paper, we will give an overview of the literature on autofluorescence imaging and spectroscopy for various clinical questions. What we intend to conclude from the studies performed so far is whether autofluorescence spectroscopy and imaging are helpful in the diagnosis of lesions of the oral mucosa, and if this is the case: for which clinical questions they are suitable. Therefore, a strong emphasis is put on in vivo human studies.

A short history of fluorescence spectroscopy and imaging

Autofluorescence for the detection of malignant lesions emanated from photodynamic therapy, a technique for cancer treatment.[1,2] In this therapy, the light-sensitive drug (“photosensitizer”) is localized in a tumour, either through systemic or topical application or by administration of a precursor, such as protoporphyrin IX (PpIX). The photosensitizer produces singlet oxygen upon excitation with light of a certain wavelength, which damages vital cell organelles inducing death of cells in the direct environment. Since some of the sensitizers were believed to accumulate in malignant tissues, they could possibly serve for diagnostics as well. The use of exogenous fluorescence for tumour detection has been investigated for various organs.[3–8] For the oral cavity, some promising results have been obtained.[9–15] However, the use of exogenous fluorophores has some major drawbacks. A certain waiting time after application is necessary for the fluorophore to reach its optimal fluorescence intensity. Furthermore, the application of photosensitizers leaves the patient temporarily sensitive for light, which negatively affects his daily life. This makes the technique impractical, especially for use in regular screenings of high-risk patient groups. Finally, the specificity of the photosensitizers appeared to be less than expected.

In the late 1970s, it was discovered that autofluorescence (also called natural or endogenous fluorescence), which had until then been regarded only as a disturbing background signal in exoge-
nous fluorescence detection, could be used for cancer detection as well.\cite{16} In the following years, numerous study on autofluorescence have been performed. Interesting results have been obtained for autofluorescence images as well as for autofluorescence spectroscopy, in which the tissue fluorescence spectral shape is used for the classification of lesions.

**The biological origins of tissue autofluorescence**

Autofluorescence of tissues is produced by fluorophores that naturally occur in living cells after excitation with a suitable wavelength. The fluorophores can be located in the tissue matrix or in cells (e.g. collagen, elastin, keratin and NADH). The invoked intrinsic autofluorescence profile is altered by absorption and scattering events in the tissue before measurement. Absorption in tissue is mainly attributed to oxy- and deoxyhemoglobin, which have different absorption profiles. Scattering is due to inhomogeneities of refraction index caused by cell nuclei and cell organelles. The presence of disease changes the concentration of the fluorophores as well as the light scattering and absorption properties of the tissue, due to changes in a.o. blood concentration, nuclear size distribution, collagen content and epithelial thickness. For example, the epithelial layer shields the strongly fluorescent collagen layer and therefore the recorded fluorescence signal will be lower in the case of hyperplasia. Conversely, excessive keratin production by lesions may produce an increase in autofluorescence intensity. Cell metabolism may increase with malignant changes, which changes the balance between the fluorescent NADH (increase) and non-fluorescent NAD+ (decrease).

Although autofluorescence spectra contain only broad features in overlapping spectral regions - which makes it virtually impossible to extract quantitative knowledge of fluorophores, scatterers, blood concentration and oxygenation in the tissue - there is sufficient proof that tissue alterations are reflected in autofluorescence spectral shape and intensity.

**7.2 Autofluorescence imaging for oral oncology**

**Development of autofluorescence imaging**

Autofluorescence imaging techniques are capable of sampling several square centimeters at a time. Tissue is illuminated with a light source, mostly in the near-UV to green range of the spectrum, and images of the fluorescence produced in the tissue and altered by absorption and scattering events are recorded using a camera. Imaging techniques typically reduce the information contained in spectral shape because only a scalar is available for each sample point (no wavelength dependence). However, imaging has the advantage of providing two-dimensional information, which allows spotting of lesion-specific features such as (in)homogeneities, while the recording of large areas makes the technique potentially useful for localizing new lesions. Much use of autofluorescence imaging for detection of cancer has been made in the lungs, where it could be performed easily by making some adjustments to a normal bronchoscope.\cite{17–22} Autofluorescence bronchoscopy has been shown to achieve a higher specificity and sensitivity than white light bronchoscopy. Nowadays, several commercially available autofluorescence bronchoscopes are used in the clinic. Studies of autofluorescence imaging of (pre)malignant lesions in other organs, including the bladder,\cite{23,24} the gastrointestinal tract,\cite{6,25–29} and colon,\cite{30–36} have also shown encouraging results for autofluorescence imaging.

Since the discovery of the potential of autofluorescence imaging for the detection of oral cancer, many studies have been performed for the oral cavity and the rest of the upper aerodigestive tract. These studies have grossly addressed three clinical questions. The first one is whether autofluores-
cence imaging is capable of providing a higher contrast between a lesion and (surrounding) healthy tissue than white light inspection. The second is whether autofluorescence imaging is helpful in differentiating between different lesion types, in particular between benign, dysplastic and malignant lesions. The third research question is aimed at the detection of unknown lesions and unknown extensions of known lesions, which is useful for tumour demarcation. Since most studies address several questions at a time, we will discuss them collectively.

Studies on autofluorescence imaging in oral oncology

In the human oral cavity, the first investigation into in vivo autofluorescence imaging has been performed by Harris and Werkhaven.[37] They noticed endogenous autofluorescence around 630 nm in 188 tumours of the oral mucosa, but also in healthy oral mucosa. This fluorescence is associated with porphyrins, as has been established later on.[38,39] According to Harris and Werkhaven, the occurrence of red autofluorescence could explain false-positives when applying photosensitizers for tumour detection. They therefore thought of autofluorescence as a disturbing rather than as a contributing factor. When Yang et al. studied autofluorescence spectra from the oral cavity, they suggested the possibility of detecting cancer by means of autofluorescence itself. They proposed to detect tumours by the 630 and 690 nm peaks that are associated with porphyrins, which they believed to be localized and retained in malignant tumours.[40]

Interesting work on autofluorescence imaging of the human oral mucosa in vivo has since then been performed by Onizawa et al. Applying an instant photography camera with ultraviolet flash lamp and a 480 nm longpass filter, an orange fluorescence was observed in 14 of 16 malignant tumours and in only one of 16 benign lesions.[41] This porphyrin-like fluorescence was probably produced by microorganisms living on ulcerating or necrotic surfaces, which is consistent with the observation that the fluorescent materials could be wiped off.

In a study with a larger patient population, autofluorescence photographs were acquired with a similar set-up in 130 patients (79 malignant tumours and 51 benign lesions).[42] They found 91.1% sensitivity and 84.3% specificity for distinguishing malignant from benign lesions. When dysplastic lesions were regarded as precancerous, these numbers increased to 93.8% and 95.5%.

In view of these results, one could suggest that the appearance of orange autofluorescence accompanies the transformation from benign into dysplastic lesions, rather than from dysplastic into malignant lesions. In a following study, Onizawa et al. showed that autofluorescence of malignant lesions tends to shift from yellow–orange to red–orange with further degrees of malignancy.[38] However, large overlaps between values for squamous cell carcinoma, dorsum of tongue and dental plaque were observed.

In a study primarily devoted to in vitro demarcation of lesions with autofluorescence microscopy, Fryen et al. also investigated cancerous lesions of the upper aerodigestive tract in vivo using a blue light detection system.[43] Thin, normal epithelium was nearly invisible, while connective tissue was strongly fluorescing. Premalignant and cancerous lesions varied between irregular opaque light and darker reddish areas. Even small cancerous lesions that were hard to visualize were well detectable using autofluorescence imaging, which is a step towards invisible lesion detection. No benign lesions were included.

Kulapaditharom et al. studied 31 normal sites and 35 inflammations, 4 granulomas, 15 dysplastic and 13 neoplastic lesions in the head and neck region through an endoscope using 442 nm excitation.[44] Nineteen of these lesions were located in the oral cavity, of which eight were either dysplastic or neoplastic. They compared the results to white light endoscopy using receiver–operator characteristics curves and found that autofluorescence imaging raised sensitivity from 68% to
The status of in vivo autofluorescence spectroscopy and imaging for oral oncology

93%, and specificity from 70% to 79% for all head–neck locations combined. In the oral cavity, a 100% sensitivity and a 73% specificity were achieved. For these numbers, dysplasia and neoplasia were considered as positives, and normal mucosa, granuloma and inflammation were considered as negatives. Diagnosis was performed not only by looking for brown or brownish red areas, probably associated with porphyrins, but also by spotting different types of dark areas, of which a focal dark area was the strongest indication for malignancy. The authors note that autofluorescence imaging is especially useful for detection of lesions because of the low number of false-negatives. They state that due to its relatively low specificity, especially for granulomas and inflammation, autofluorescence imaging is to be used as a diagnostic rather than as a confirmatory test.

Betz et al. studied autofluorescence imaging in 30 patients with a tumor of the oral mucosa or oropharynx.[45] Using 375–440 nm excitation they found sufficient to excellent demarcation by lower fluorescence intensity in 20 tumors. However, 10 tumors were not distinguishable from their host tissues. These tumors were all located at the tongue, the soft palate or the tonsillar sinus. In general, flat epithelial lesions were found to be outlined subjectively better by autofluorescence imaging than large, exophytic tumors. The authors state that porphyrin-like fluorescence is not a good diagnostic indicator. According to them, the porphyrins are microbially synthesized and therefore limited to the necrotic surface of exulcerated tumors. Also, porphyrin-like fluorescence has been observed on the dorsums of tongues and gingival plaques. Furthermore, only 33% of the tumors were mostly covered by strongly red fluorescing material. The observed red spots did not seem to be spread homogeneously over the lesions’ surfaces, which makes porphyrin fluorescence unsuitable for the demarcation of oral cancer.

In a later study, Betz et al. compared normal inspection, 5-ALA-induced PpIX fluorescence imaging and autofluorescence imaging with the same set-up in 56 patients.[9] Autofluorescence imaging alone gave correct detection of malignancies in 63% of cases (compared with 91% for white light inspection). The demarcation showed considerable flaws. However, the early stage lesions were well detected and demarcated, probably due to their flat appearance. The authors therefore suggest that autofluorescence imaging is especially suitable for screening of premalignant and early malignant oral lesions.

Paczona et al. investigated the use of autofluorescence videomicroscopy for diagnosis of head and neck squamous cell carcinoma.[46] They measured 48 patients, of whom 8 had oropharyngeal and 3 had oral cancer, and compared the results with white light videomicroscopy. All oral and oropharyngeal lesions were visible by autofluorescence as well as white light videomicroscopy. Using autofluorescence, tumors presented as darker areas with a more or less accentuated reddish-blue color. However, in three patients with oropharyngeal cancer, the lesions turned out to be more extended using autofluorescence endoscopy. In two patients, an additional lesion was found by autofluorescence videomicroscopy (one severe dysplasia, one moderate dysplasia). A summary of the literature on autofluorescence imaging in the oral cavity is given in Table 7.1.
Table 7.1. Results from the literature for autofluorescence imaging of oral lesions.

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Excitation wavelength (nm)</th>
<th>Emission wavelength (nm)</th>
<th>Number of lesions (benign, dysplastic, malignant)</th>
<th>Method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Onizawa et al., 1996</td>
<td>UV flash lamp</td>
<td>&gt;480</td>
<td>34 (18, 0, 16)</td>
<td>Porphyrin-like fluorescence</td>
<td>Observed in 14 (16) malignant and 1 (16) benign</td>
</tr>
<tr>
<td>Onizawa et al., 1999</td>
<td>UV flash lamp</td>
<td>&gt;480</td>
<td>130 (79, 7, 44)</td>
<td>Porphyrin-like fluorescence</td>
<td>94% sens. and 96% spec. for malignant vs. benign + dysplastic lesions</td>
</tr>
<tr>
<td>Fryen et al., 1997</td>
<td>Unclear (halogen source)</td>
<td>Unclear</td>
<td>Unknown</td>
<td>Irregular opaque light and darker reddish areas</td>
<td>Even not clearly visible tumours detectable using autofluorescence</td>
</tr>
<tr>
<td>Betz et al., 1999</td>
<td>375-440</td>
<td>&gt;515</td>
<td>30 (0, 0, 30)</td>
<td>Dark areas, porphyrin-like fluorescence</td>
<td>Sens. 67%, especially flat lesions. Porphyrin fluorescence not useful.</td>
</tr>
<tr>
<td>Betz et al., 2002</td>
<td>375-440</td>
<td>&gt;515</td>
<td>56 (0, 0, 56)</td>
<td>Dark areas</td>
<td>Sens. 63% as compared to 91% for WLE. Flat early-stage lesions well detectable.</td>
</tr>
<tr>
<td>Kulapaditharom et al., 2001</td>
<td>442</td>
<td>480-520 + &gt;630</td>
<td>67 (39, 15, 13) (different head-neck regions combined)</td>
<td>Reddish, brown fluorescence + dark areas</td>
<td>Compared to WLE: sens. raised from 68 to 93%, spec. from 70 to 79% (combined locations)</td>
</tr>
<tr>
<td>Paczona et al., 2003</td>
<td>375-440</td>
<td>&gt;515</td>
<td>11 (0, 0, 11)</td>
<td>Diminished green, shift to reddish-blue, violet autofluorescence</td>
<td>Sens. 100% for autofluorescence and WLE. Two additional dysplasias, 3 lesion extensions spotted</td>
</tr>
</tbody>
</table>

7.3 Autofluorescence spectroscopy for oral oncology

Generally, autofluorescence spectroscopy systems consist of a light source - usually in the near-UV to visible wavelength range - that excites the tissue through a fiber. The fluorescence that is produced in the tissue is analyzed by a spectrograph, while the reflected excitation light is filtered out. The recorded fluorescence spectra can be saved to a computer, which allows mathematical
spectral analysis of many types. Much use has been made of multivariate techniques like Principal Components Analysis (PCA), but also of emission wavelength ratios and artificial neural networks. Many studies have been performed to distinguish oral lesions (cancerous, dysplastic or benign) from healthy oral mucosa using autofluorescence spectroscopy. Fewer studies have tried to distinguish between benign and (pre)malignant lesions. In this section, we will give an overview of studies on autofluorescence spectroscopy on the oral mucosa for both questions separately.

**Distinguishing lesions from healthy oral mucosa**

In 1995, Kolli et al. studied the use of autofluorescence spectra for the detection of neoplastic upper aerodigestive mucosa.\[47\] They studied 31 patients with 27 malignant tumours and 4 potentially premalignant lesions, and found significant differences in fluorescence intensity ratios between healthy mucosa and the lesions. Two emission scans (excitation 300 and 340 nm) and two excitation scans (emission 380 and 450 nm) were applied. For all scans, wavelength ratios were calculated. Three of the four scans showed significant differences between lesions and healthy tissue, one excitation scan did not. In another study, Kolli et al. found that changes in autofluorescence emission and excitation characteristics correlated with epithelial thickness.\[48\] Both studies were only aimed at finding significant differences in autofluorescence spectra between tumours and healthy mucosa, and did not attempt to classify yet.

Chen et al. performed an ex vivo study of oral mucosa and observed significant differences (p < 0.001) between healthy oral mucosa on the one hand, and malignant plus premalignant lesions on the other hand.\[49\] The best results were obtained for the 470/330 nm emission ratio at 300 nm excitation. This technique was able to classify abnormal tissue with a positive predictive value (PPV) of 94%, and normal oral mucosa with a PPV of 93%. However, the classification of premalignant vs. malignant lesions was not possible. Next to that, they studied ex vivo autofluorescence in normal and malignant oral tissues.\[50\] After excitation with 330 nm, they found significant differences in fluorescence intensity around 380 and 460 nm between the 15 normal and 15 malignant oral tissue samples from 15 oncology patients. However, standard deviations in fluorescence intensities were approximately five times larger than the differences between the two groups. Therefore, reliable classification on this basis would not be possible. Since the studies by Chen et al. have been performed ex vivo, they were not included in Table 7.2.

Gillenwater et al. recorded oral mucosa autofluorescence spectra from 8 healthy volunteers and 15 patients with premalignant or malignant lesions at 337, 365 and 410 nm excitation in vivo.\[51\] For lesions, they noticed a decreased intensity in the blue spectral regions, and an increased fluorescence around 635 nm (porphyrin-like fluorescence). Based on the ratio between these values, they achieved a sensitivity of 82% and a specificity of 100% for distinguishing lesions from healthy mucosa. Large, unexplained variations in fluorescence intensity were observed between individuals.

Schantz et al. measured autofluorescence at a fixed detection wavelength (450 nm) that they believe to represent mainly NADH fluorescence, while scanning the excitation wavelength from 290 to 415 nm in 35 oral cancer patients.\[52\] They noticed that a loss of tumour differentiation is associated with a lower maximum excitation wavelength. Also, they found a correlation between the excitation maximum of cancerous and contralateral healthy oral mucosa in the same patient. No classification was attempted and values for the excitation maxima showed overlaps between the different tumour groups.

Betz et al.\[45\] performed autofluorescence spectroscopy under broad blue excitation in 36 patients with a cancerous tumour of the oral cavity or oropharynx, and in one healthy volunteer.
In 34 of the 36 patients, neoplastic tissue showed lower overall and peak autofluorescence intensities than surrounding tissue, while spectral shapes were similar. Large inter- and intraindividual variances were found in autofluorescence intensities for healthy as well as diseased mucosa. The dorsal side of the tongue showed a relatively high intensity in the red region.

Inaguma et al. investigated porphyrin-like fluorescence in oral cancer.[53] Of 78 oral carcinomas studied in vivo under 410 nm excitation, 66 carcinomas (85%) produced red fluorescence around 630 nm, accompanied by a smaller peak at a slightly higher wavelength. From the 43 benign lesions that were measured as well, porphyrin-like fluorescence was observed in only three. These were two leukoplakias and one acute necrotizing ulcerative gingivitis. A capillary electrophoretic method was used to analyze the fluorescing substances in three of the tumours. Inaguma et al. conclude that several porphyrin compounds are responsible for the fluorescence, among which PpIX.

Heintzelman et al. tried to establish the optimal excitation wavelength for the detection of oral lesions.[54] They measured excitation–emission maps using 330–500 nm excitation wavelengths in 10 nm increments. The training set consisted of nine healthy volunteers, in whom 41 clinically normal sites were measured, and 17 patients with a known or suspected premalignant or malignant oral lesion, in whom 47 sites (10 lesion sites) were measured. The validation set consisted of 53 healthy volunteers, in whom 274 clinically normal sites were measured, and 3 patients in which 7 sites (4 lesion sites) were measured. All spectra suitable for analysis turned out to correspond to lesions that ranged from mildly dysplastic to cancerous, so no benign lesions were included in the analysis. Therefore, effectively the distinction between lesions and healthy oral mucosa has been made. Using different combinations of concatenated vectors normalized by the peak intensity, the authors achieved 100% sensitivity and 90% specificity for the training set and similar results for the full data set. When spectra were normalized before concatenation, the results were less satisfactory. The addition of porphyrin-like fluorescence information to the algorithm was not useful. The authors note that this is explained by its low specificity, since red fluorescence is observed at some normal, some dysplastic and some cancerous sites.

Majumder et al. studied human oral cavity cancers and found a sensitivity of 86% and a specificity of 63% for distinguishing mostly advanced tumours from visually uninvolved sites.[55] They used 337 nm excitation and investigated 25 patients with oral cancer. Discrimination was attempted using Principal Components Analysis (PCA). In a later study, Majumder et al. compared PCA and non-linear algorithms for classification of cancerous and normal tissue. This time, 16 patients with cancer of the oral cavity as well as 13 normal volunteers were included. The non-linear algorithm provided a sensitivity of 93% and a specificity of 96% for the training set, and 95% and 96% for the validation set. For PCA, these values were lower.[56] We noticed that the preprocessing of the data introduces some foreknowledge of the lesion type, because after the mean healthy spectrum has been subtracted from all spectra, the resultants are divided by the standard deviation of the spectral category from which these originate. The authors remark that patient selection may influence algorithm performance, but explain that the main object of this paper was to compare the different algorithms. A study with more than two classes, including benign lesions, in a less biased patient population will be performed at a later stage.

Muller et al. excited 91 tissue sites (normal, dysplastic and cancerous) in 15 patients and 8 healthy volunteers in vivo with 11 excitation wavelengths between 337 and 610 nm.[57] Benign lesions were measured but left out of the analysis because of the small size of these categories. Autofluorescence spectra recorded at 337 and 358 nm excitation were analyzed by fitting a linear combination of collagen and NADH fluorescence to the intrinsic autofluorescence as obtained by
a model. The fluorescent spectral components were extracted from the intrinsic autofluorescence by means of multivariate curve resolution. In combination with diffuse reflectance spectroscopy and light scattering spectroscopy, they achieved a sensitivity of 96% and a specificity of 96% for distinguishing cancerous and dysplastic lesions from normal tissue. The distinction between cancerous and dysplastic tissue was made with a sensitivity of 64% and a specificity of 90%.

The literature on distinguishing lesions from healthy oral mucosa in vivo using autofluorescence spectroscopy has been summarized in Table 7.2.
Table 7.2. Results from the literature for in vivo autofluorescence spectroscopy of oral lesions: distinction between lesions and healthy mucosa.

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Excitation wavelength(s) (nm)</th>
<th>Emission wavelength(s) (nm)</th>
<th>Number of lesions (benign, dysplastic, malignant)</th>
<th>Method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolli et al., 1995</td>
<td>Different combinations (200-430)</td>
<td>Different combinations (300-660)</td>
<td>31 (unknown, unknown, 27)</td>
<td>Emission and excitation wavelength ratios</td>
<td>Significant differences between healthy and lesions</td>
</tr>
<tr>
<td>Kolli et al., 1995</td>
<td>Different combinations (200-430)</td>
<td>Different combinations (300-660)</td>
<td>19 (unknown, unknown, 18)</td>
<td>Emission wavelength ratios</td>
<td>Correlation found with epithelial thickness</td>
</tr>
<tr>
<td>Gillenwater et al., 1998</td>
<td>337, 365, 410</td>
<td>380-680</td>
<td>33 sites (&gt;0, &gt;0, &gt;0)</td>
<td>635 / (445-490) ratio</td>
<td>Sens. 82%, spec. 100%</td>
</tr>
<tr>
<td>Schantz et al., 1998</td>
<td>290-415</td>
<td>450</td>
<td>35 (0, 0, 35)</td>
<td>Excitation maximum</td>
<td>Loss of tumour differentiation -&gt; lower excitation maximum</td>
</tr>
<tr>
<td>Betz et al., 1999</td>
<td>375-440</td>
<td>470-600</td>
<td>36 (0, 0, 36)</td>
<td>511 nm intensity</td>
<td>34 tumours showed lower intensity</td>
</tr>
<tr>
<td>Inaguma et al., 1999</td>
<td>410</td>
<td>605-700</td>
<td>121 (43, 0, 78)</td>
<td>Porphyrin-like peak</td>
<td>66 of 78 tumours and 3 of 43 benign lesions: porphyrin-like fluorescence</td>
</tr>
<tr>
<td>Heintzelman et al., 2000</td>
<td>330 –500 in 10 nm increments (Exc+10) to 700</td>
<td>(Exc+10) to 700</td>
<td>14 (0, 6, 8)</td>
<td>PCA on concatenated spectra</td>
<td>100% sens., 90% spec. using 350, 380 and 400 nm excitation spectra in concatenation</td>
</tr>
<tr>
<td>Van Staveren et al., 2000</td>
<td>420</td>
<td>465-650</td>
<td>22 (9, 13, 0)</td>
<td>Artificial neural network</td>
<td>86% sens., 100% spec.</td>
</tr>
<tr>
<td>Majumder et al., 2000</td>
<td>337</td>
<td>350-700</td>
<td>25 (0, 0, 25)</td>
<td>PCA + logistic discrimination</td>
<td>86% sens., 63% spec.</td>
</tr>
<tr>
<td>Majumder et al., 2003</td>
<td>337</td>
<td>375-700</td>
<td>16 (0, 0, 16)</td>
<td>Nonlinear maximum presentation and discrimination feature</td>
<td>95% sens., 96% spec.</td>
</tr>
<tr>
<td>De Veld et al., 2003</td>
<td>6 λ. between 365-450</td>
<td>467-867</td>
<td>134 (106, 12, 17)</td>
<td>PCA with various classifiers, artificial neural networks, emission wavelength ratios</td>
<td>ROC-area 0.88 for lesions vs. healthy, 0.97 for malignant tumours vs. healthy mucosa (ROC-area 1 = perfect)</td>
</tr>
<tr>
<td>Muller et al., 2003</td>
<td>11 λ. between 337 and 610</td>
<td>350-700</td>
<td>27 sites in 15 patients (0, 19, 12)</td>
<td>Logistic regression</td>
<td>96% sens., 96% spec.</td>
</tr>
</tbody>
</table>

Table 7.2. Results from the literature for in vivo autofluorescence spectroscopy of oral lesions: distinction between lesions and healthy mucosa.
Classification of different lesion types

In this section, we will discuss several studies that have attempted to classify oral lesions using autofluorescence spectroscopy. Dhingra et al. studied healthy oral mucosa of 10 volunteers and a combination of 19 oral and oropharyngeal lesions in vivo by means of 370 and 410 nm excitation.[58] The lesions comprised hyperplasia and/or hyperkeratosis (n = 8), lichen planus (n = 1), dysplasia (n = 4), carcinoma in situ (n = 1) and invasive squamous cell carcinoma (n = 5). For 410 nm excitation, the average ratio of the fluorescence intensity at the peak intensity centered around 490 nm of contralateral normal mucosa to that of malignant lesions was 6, while for benign lesions this ratio was less than 2. Also, they noticed a porphyrin-like peak at 640 nm under excitation with 410 nm for the neoplastic lesions, while this peak did not occur in healthy mucosa or benign lesions. Using these two characteristics in a scatter plot and drawing a straight line between neoplastic lesions on the one hand and benign and healthy mucosa on the other, they achieved fairly good separation. Of the 19 lesions, 17 were diagnosed correctly and two were false-positives. The separation line was drawn such that only the information in the blue spectral part was used for classification, and not the porphyrin-like peak. An interesting detail is that Dhingra et al. chose to use a relatively high pressure of the probe onto the tissue (300g/cm²), because this gave the highest fluorescence intensity. This is probably explained by the pushing away of blood by the fiber, which is consistent with their autofluorescence spectra showing no blood absorption dips. Their results therefore do not rely on blood absorption information, but only on fluorophore concentrations, tissue architecture and the distribution of scatterers. The authors believe that the decreased blue fluorescence intensity probably relies on epithelial thickening and cellular proliferation, which shelters the highly fluorescent collagen layer and thus reduces the measured fluorescence intensity. The good results as achieved by Dhingra et al. are in our opinion not very reliable. There was no validation set but only a small dataset that has been classified afterwards, which makes their results very sensitive to overfitting.

Ingrams et al. performed an in vitro study on 10 dysplastic or malignant and 12 healthy or benign human biopsy specimens from clinically suspicious lesions and healthy oral mucosa using various excitation wavelengths.[59] The exact numbers of dysplastic, malignant, benign and healthy specimens are not given. The most prominent difference found between dysplastic or malignant and healthy or benign mucosa was the porphyrin-like peak centered at 635 nm and accompanied by a smaller peak at 690 nm. Ingrams et al. established 90% sensitivity and 91% specificity based on this peak. They note that if the observed fluorescence corresponds to Protoporphyrin IX, they have in effect equated the presence of a certain concentration of this compound in cells with the presence of dysplasia or malignancy. They also note that they cannot be sure whether these porphyrin-like substances correspond to dysplasia or malignancy because of accumulation of PpIX due to a lack of ferrochelatase in tumour cells, or because of bacteria that produce porphyrins. The latter would make the method less reliable, since bacteria can live at benign lesions as well. Since this study was performed in vitro, it has not been included in Table 7.3.

Van Staveren et al. investigated 22 visible lesions (9 benign, 13 dysplastic) in 21 patients suffering from oral leukoplakia, and from 12 locations in two healthy volunteers.[60] Autofluorescence spectra were recorded under 420 nm excitation and analyzed using an artificial neural network. The separation of abnormal from normal tissue resulted in 86% sensitivity and 100% specificity and has been added to Table 7.1, but classification according to the grade of dysplasia was not feasible. Surprisingly, the border of lesions seemed to contain more relevant information than the center of the lesion.

Wang et al. have performed various studies on autofluorescence spectroscopy for oral cancer
They performed studies in hamsters as well as in humans, both \textit{in vivo} and \textit{ex vivo}. In their \textit{in vivo} human study, they measured 15 healthy volunteers, 30 patients with oral submucous fibrosis, 39 patients with oral leukoplakia (of which 26 histologically diagnosed as epithelial hyperkeratosis and 13 as epithelial dysplasia) and 13 patients with oral squamous cell carcinoma.[64] For their distinction of premalignant and malignant from benign lesions under 330 nm excitation, they established a sensitivity of 81\% and a specificity of 96\% using a partial least squares and artificial neural network (PLS-ANN) classification algorithm. The test results were calculated on the basis of leave-one-out, so that the risks of overfitting were strongly reduced. All lesions measured by Wang et al. were located at the oral buccal mucosa, and all were induced by areca quid chewing and smoking habits. Their lesion set thus was quite homogeneous, which may have facilitated the classification. For use in the clinic, the algorithm would also have to be applicable for different lesion types and at other anatomical locations.

Tsai et al. studied the influence of oral submucous fibrosis (OSF) on autofluorescence spectra.[66] They measured 15 healthy volunteers and 149 patients with clinically suspected oral submucous fibrosis, epithelial hyperkeratosis, epithelial dysplasia and squamous cell carcinoma. All lesions were located at the buccal mucosa and induced by areca quid chewing. Using 330 nm excitation, they found that the 460/380 nm ratio increased from fibrosis to healthy mucosa to hyperkeratosis to dysplasia and finally to cancer (p < 0.01 for all comparisons). However, the presence of OSF made hyperkeratosis and dysplasia indistinguishable from healthy mucosa. Furthermore, the standard deviations for the ratio were largely overlapping for hyperkeratosis and dysplasia in both OSF and non-OSF patients (non-OSF: 0.97 ± 0.17 and 1.01 ± 0.19, respectively; OSF: 0.77 ± 0.26 and 0.68 ± 0.22, respectively). This means that dysplastic and hyperkeratotic lesions are indistinguishable.

In a study that we performed ourselves, we recorded autofluorescence spectra from 97 healthy volunteers and from 155 patients with 172 benign, dysplastic and malignant oral lesions.[67] We excluded all measurements from the dorsal side of the tongue and the vermillion border of the lip from the analysis because of the differences established between healthy locations.[68] These differences, that are irrelevant to our clinical question, could possibly disturb lesion classification. This left 134 lesions to be included in the spectral analysis. Applying various algorithms including PCA and emission wavelength ratios to spectra recorded under six different excitation wavelengths, we found excellent separation between healthy mucosa and oral cancer (area under the ROC-curve up to 0.97, with ROC-area = 0.5 corresponding to random classification and ROC-area = 1 to flawless classification). The separation between healthy mucosa and lesions of any kind was good (area under the ROC-curve up to 0.88). These results are included in Table 7.2. We found no strong dependency of the results on the excitation wavelength. Separation between benign lesions on the one hand and dysplastic and cancerous on the other was not feasible (ROC areas <0.65 for all methods and excitation wavelengths). We have explained this by noting the enormous variations in autofluorescence spectra, which overlap between the different classes and especially between benign and dysplastic lesions. In Figure 7.1 the median spectra for different categories are plotted, both non-normalized and normalized to the blue peak intensity. It is clear that in both cases benign and dysplastic lesion spectra are intertwined. The literature on classification of oral lesions \textit{in vivo} using autofluorescence spectroscopy has been summarized in Table 7.3.
The status of in vivo autofluorescence spectroscopy and imaging for oral oncology

Figure 7.1. (a) Median autofluorescence spectra of lesions of the oral mucosa of different types, excitation wavelength 405 nm and (b) normalized spectra.

Table 7.3. Results from the literature for in vivo autofluorescence spectroscopy of oral lesions: distinction between malignant and dysplastic lesions on the one hand, and benign lesions on the other.

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Excitation wavelength(s) (nm)</th>
<th>Emission wavelength(s) (nm)</th>
<th>Number of lesions (benign, dysplastic, malignant)</th>
<th>Method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dhingra et al., 1996</td>
<td>370 and 410</td>
<td>380-700</td>
<td>19 (9, 4, 6)</td>
<td>Intensity ratio lesion to contralateral</td>
<td>100% sens., 80% spec.</td>
</tr>
<tr>
<td>Tsai et al., 2003</td>
<td>330</td>
<td>340-601</td>
<td>104 (57, 27, 20)</td>
<td>480/360 nm emission ratio</td>
<td>81% sens., 87% spec.</td>
</tr>
<tr>
<td>Van Staveren et al., 2000</td>
<td>420</td>
<td>465-650</td>
<td>22 (9, 13, 0)</td>
<td>Artificial neural network</td>
<td>Classification not feasible</td>
</tr>
<tr>
<td>Wang et al., 2003</td>
<td>330</td>
<td>340-601</td>
<td>82 (56, 13, 13)</td>
<td>Partial least squares + artificial neural network</td>
<td>81% sens., 96% spec.</td>
</tr>
<tr>
<td>De Veld et al., 2003</td>
<td>6 between 365-450</td>
<td>467-867</td>
<td>134 (106, 12, 17)</td>
<td>PCA with various classifiers, artificial neural networks, emission wavelength ratios</td>
<td>Classification not feasible</td>
</tr>
</tbody>
</table>
7.4 Discussion

Autofluorescence imaging

The experimental approaches reported for autofluorescence imaging in the oral cavity are quite similar. The images are usually recorded by means of a CCD camera and occasionally by photography. Most of the times, the tissue is excited with light in the 375–440 nm range. This is due to the commercially available systems for fluorescence imaging that can be used for the lungs as well. Therefore, no studies have been performed to decide on the most favorable excitation wavelengths for the imaging of oral malignancies. However, this information can potentially be deduced from spectroscopy, in which excitation wavelength studies have been performed. The criteria used for labeling areas as lesions were diminished fluorescence intensity (in the blue or green) or increased red fluorescence that is probably associated with porphyrins. Although high sensitivities and specificities have been obtained using porphyrin-like fluorescence,[41,42] others studies have claimed that red fluorescence is not specific for malignancies.[45] This discrepancy can possibly be explained by the differences in patient population. Porphyrins have been associated mainly with ulcerating tumours, because of the bacteria that grow in these tissues. Therefore, a higher fraction of advanced tumours may increase the sensitivity and specificity of porphyrin-based tumour detection.

The first clinical question that we have asked was whether autofluorescence imaging can improve the contrast between lesions. For the answer, a comparison has to be made between white light inspection and autofluorescence imaging. A few studies have made this comparison and gave varying results. Kulapaditharom et al. found increased sensitivity and specificity as compared to white light inspection in a large and varied patient population.[44] On the other hand, Betz et al. found a lower sensitivity for autofluorescence imaging than for white light inspection in a large patient population consisting of only malignant tumours.[9,45] Therefore, we think that more research on the contrast between lesions and healthy oral mucosa should be performed that take special care in relating the results to the patient population. From the studies performed so far, we can only conclude that autofluorescence imaging of the oral mucosa possibly improves lesion contrast and that this is certainly the case for flat, early lesions.[9,45]

The second question we asked was whether autofluorescence imaging is helpful in differentiating between different lesion types. This question has been addressed with good results by means of porphyrin-like fluorescence.[41,42] However, this fluorescence has been shown to be not very specific for malignant lesions.[45] Probably, this inconsistency is related to differences in the lesion dataset - for example, ulcerating tumours are known to very frequently produce porphyrin-like fluorescence. Using differences in autofluorescence intensity, Kulapaditharom et al. found a 100% sensitivity for distinguishing dysplastic and malignant from benign lesions in the oral cavity, but the specificity was only 73%.[44] However, this was still a higher value than for white light endoscopy. Overall, the specificity of autofluorescence imaging for distinguishing (pre)malignant from benign lesions does not seem to be very promising.

The third research question concerned the detection of previously unknown and invisible lesions and of unknown extensions of visually detected lesions. Indications have indeed been found that autofluorescence imaging is capable of detecting invisible lesions or invisible tumour extensions.[43,46] However, these results concern only small numbers and therefore the question requires further research. The detection of invisible tumour extensions is probably the most practical way to establish the potential of the method for detection of invisible tissue alterations. This approach does not require complete scanning of the oral cavity and the taking of validation biopsies on visually healthy-looking tissue. If the detection of invisible tumour extensions using auto-
fluorescence imaging gives promising results, the scanning of high-risk groups, e.g. patients who have suffered from oral cancer in the past, could be the next step.

**Autofluorescence spectroscopy**

In autofluorescence spectroscopy of the oral cavity, many different approaches have been chosen. The excitation wavelengths used span a wide range from 200 to 610 nm, and often spectra recorded at several excitation wavelengths have been combined for lesion classification. Regarding the classification methods used, a similar variety is observed. Multivariate statistical techniques are popular, in particular PCA. Emission wavelength (red/green) ratios exploiting the relative decrease in green autofluorescence intensity in lesions have been applied, as well as fluorescence intensity, neural networks and porphyrin-like peaks. For distinguishing lesions from healthy mucosa, several studies have tried to establish the most successful excitation wavelengths, while others have adopted optimal excitation wavelengths from earlier studies ex vivo or in other organs. Those that searched for optimal wavelengths, found various results. The combination of 350, 380 and 440 nm as well as single wavelengths 337 nm, 410 nm and six wavelengths between 365 and 450 nm have been reported.\[51,54,58,67\] In view of these inconsistent results, we suggest that the exact wavelength used for excitation might not be so relevant for the method's effectiveness as is often believed. We believe that this is due to the broad excitation bands of the most important fluorophores, like collagen. The methods used for classification of spectra have also given varying results. For example, porphyrin-like fluorescence gave good results in some studies, but in others turned out to be worthless. Therefore, further research is required to validate different methods for distinguishing between healthy and diseased oral mucosa. For an honest comparison of methods, we believe that ROC curve areas should be calculated because these do not rely on an arbitrarily chosen threshold. Furthermore, we believe that it is very important to look at the characteristics of the lesion dataset. In general, more benign lesions and more variation in the dataset complicate the classification. However, the reality in oral pathology is just that: relatively many benign lesions and much variation in lesion type. A useful classification algorithm therefore must be able to function well in these complex circumstances.

From the literature, we can conclude that autofluorescence spectroscopy seems to be very accurate for distinguishing lesions from healthy oral mucosa (sensitivity 82–100%, specificity 63–100%). This is especially true for distinguishing malignant tumours from healthy mucosa, for which sensitivities and specificities >95% are no exception. However, the method seems to be less specific for classification of different lesion types. In our own study, we found that classification between different lesion types was not feasible. Other studies have claimed good results, but to our opinion these studies were not convincing, as discussed above. We therefore believe that at this moment, performing an optical biopsy using *in vivo* autofluorescence spectroscopy is not possible for the oral cavity.

We can think of various reasons for the problems arising in discriminating between benign lesions on the one hand and dysplastic and malignant lesions on the other hand. In general, tumours as well as benign and dysplastic lesions can be of various degrees of keratinization, hyperplasia and blood content. All these factors influence the shape and intensity of autofluorescence spectra, and since large variances are observed within different lesion categories, a mixing of categories by means of autofluorescence spectroscopy classification can easily result. Information provided by other fluorophores like NADH and FAD may be more specific for (pre)malignancy, but in our opinion it has not been convincingly shown that NADH and FAD fluorescence are measurable from
oral mucosa in vivo. We believe that it is not feasible to extract the exact intrinsic autofluorescence from the recorded signal, because all methods for doing this - e.g. using reflectance spectra - require prior knowledge of the tissue under investigation. If it were feasible, the question would still remain whether intrinsic fluorescence spectra would give better results. To our opinion, it has not been convincingly shown that scattering and blood absorption artifacts play a less important role for classification than the presence of certain fluorophores. Furthermore, many other factors than disease have been shown to influence autofluorescence spectra as well, such as gender, pigmentation, tobacco and alcohol consumption habits and the wearing of dentures. [69] These influences might be disturbing factors in lesion classification. However, we do not know of any reliable way to correct autofluorescence spectra for these individual factors. [69]

7.5 The place of autofluorescence diagnostics in oral oncology - conclusions

In summary, both autofluorescence imaging and spectroscopy give good results for the distinction of lesions from normal mucosa. In the case of imaging, this is very useful because it gives the clinician a tool to scan the oral cavity for new lesions, and possibly to assess invisible extensions of visually detected lesions. In fact, there are indications that autofluorescence imaging can be used to find lesions that are not or not easily noticed by visual inspection.

We believe that in contrast to imaging, autofluorescence spectroscopy is for practical reasons not suitable to detect new lesions or to demarcate lesions. In our opinion, it is not feasible to scan the complete oral cavity using the small sampling area that results from the use of an optical fiber. Beginning lesions are very small and therefore each and every part of the mucosa would have to be measured separately. We have performed an unpublished pilot follow-up study of patients with a history of oral cancer. These patients, who are at a high risk for developing second primary tumours, were scanned at anatomical locations that are prone to develop oral cancer, in particular the lateral border of the tongue. However, we soon found that it would be impossible to scan the complete oral cavity—not only because of the time-consuming nature of the measurements, but also because it would be impossible to mark which locations have already been measured.

Because scanning the complete oral cavity using point spectroscopy is impossible, the measurement probe will in practice only be placed at lesions that have already been found by visual inspection or another technique, which explains the excellent figures for sensitivity and specificity found in many studies. As a consequence, the relevant question will not be whether the tissue at this location is normal or not, because it has already been established to be abnormal. The relevant question will instead be whether the visible tissue alterations are of a benign or a (pre)malignant nature. We believe that this is the most relevant clinical question for classification of visible lesions, because it is crucial for the treatment planning. If possible, autofluorescence spectroscopy could be used to find the optimal, most dysplastic location for biopsy. Unfortunately, the literature shows that autofluorescence is not specific enough for this purpose. We therefore believe that autofluorescence imaging might be appropriate as an easy-to-use, sensitive and inexpensive method for lesion detection, although further research is still necessary.
The status of in vivo autofluorescence spectroscopy and imaging for oral oncology

7.6 References

15. Zheng W, Soo KC, Sivanandan R, Olivo M. Detection of squamous cell carcinomas and pre-cancerous lesions in the oral cavity by quantification of 5-aminolevulinic acid induced fluorescenc-
34. Eker C, Montan S, Jaramillo E, Koizumi K, Rubio C, Andersson-Engels S et al. Clinical spec-


