ORIGINAL ARTICLE

FLUORESCENCE INVESTIGATIONS TO CLASSIFY MALIGNANT LARYNGEAL LESIONS IN VIVO

Roland Rydell, MD, PhD,1,2 Charlotta Eker, PhD,3 Stefan Andersson-Engels, PhD,3 Annelise Krogdahl, MD, PhD,4 Peter Wahlberg, MD, PhD,1 Katarina Svanberg, MD, PhD5

1 Department of Oto-Rhino-Laryngology, Lund University Hospital, S-221 85 Lund, Sweden. E-mail: roland.rydell@med.lu.se
2 Department of Logopedics and Phoniatrics, Lund University Hospital, Lund, Sweden
3 Department of Physics, Lund Institute of Technology, Lund, Sweden
4 Department of Pathology, Odense University Hospital, Odense, Denmark
5 Department of Oncology, Lund University Hospital, Lund, Sweden

Accepted 25 June 2007
Published online 2 January 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hed.20719

Abstract: Background. The extent of surgical resection for malignant laryngeal lesions influences voice quality. An instrument to estimate histopathologic grading of dysplasia in vivo may spare normal tissue without increasing the risk of local failure.

Methods. Laryngeal lesions (N = 39; 21 after administration of delta-aminolevulinic acid (ALA)) were investigated with laser-induced fluorescence, and the results correlated to histopathologic grading in 4 groups: non/mild dysplasia (I), moderate dysplasia (II), severe dysplasia/cancer in situ (III), and carcinoma.

Results. At 337-nm excitation, there were differences in the fluorescence ratio $I_{431}/I_{390}$ between groups I/III and carcinoma. Following 405-nm excitation, $I_{500}$ group I differed significantly from group III and cancer. The sensitivity for the fluorescence method was 89%, and the specificity was 100%.

Conclusions. There are differences in the in vivo tissue fluorescence between tissue with different gradings of dysplasia and carcinoma.

Keywords: fluorescence; delta-Aminolevulinic acid; larynx; malignancy; histopathology

To cure a patient with laryngeal cancer by surgery, all cancer cells have to be removed. The ultimate measurement of success in this effort is that the patient is alive without disease at follow-up after 5 years. The intermediate early end-point, radical excision, is basically an extrapolation from investigations on the tissue that has been taken away. More precise would be to directly peroperatively determine if any cancer cells are left after the surgical excision. In vivo fluorescence investigations may be a way to identify such cells. The fluorescence method could also assist in the diagnosis. The diagnostic process for premalignant lesions, carcinoma in situ, and early-stage cancer of the larynx may demand repeated biopsies before a diagnosis can be reached, with increased risk of affecting voice quality as well as delaying the treatment. Both the type and the true extension of a lesion are important in deciding the treatment.
Fluorescence has been used successfully in studies of malignant lesions in various organs. In those studies, both endogenous tissue autofluorescence\(^1\)–\(^7\) and fluorescence of a tumor-marking agent have been investigated.\(^8\),\(^9\) Several groups have also studied the fluorescence from lesions in the head and neck region.\(^10\)–\(^33\) The results presented in these reports are promising, suggesting further, more detailed studies. The predictive value of autofluorescence results has been evaluated by our group in an earlier paper, using multivariate analysis of fluorescence spectra.\(^15\) For autofluorescence spectra, logistic regression based on principal component analysis (PCA) or PLS, or PLS-DA all resulted in sensitivities and specificities around 90% for lesion versus normal. Using ALA and 405-nm excitation gave a sensitivity of 100% and a specificity of 69%. By selecting wavelengths of interest instead of measuring whole fluorescence spectra, the amount of data is reduced and the speed of the measuring process increases. This is important in the clinical setting. It also makes it possible to measure several data points at the same time, and thus, it is possible to construct an imaging system designed to evaluate an area rather than a point.

Tissue fluorescence arises as light is absorbed by certain molecules in the tissue, which act as fluorophores, meaning that they release some of the absorbed light energy by emission of fluorescence light. Since some of the absorbed energy is lost in nonradiative processes, the released fluorescence photons will have a lower energy, hence a longer wavelength than the incoming excitation light. Under UV or blue irradiation, several fluorophores in the tissue, eg, collagen, elastin, NADH, and porphyrins, will yield fluorescence in the visible wavelength region. The shape of the tissue fluorescence spectrum is a broad distribution (see for example Figure 1), the shape of which depends mainly on the excitation wavelength, presence of tissue fluorophores, and tissue optical properties. Fluorescence spectra can thus yield information about the biochemical, and indirectly the pathologic, properties of the tissue illuminated.

There are 2 basic clinical problems in the management of premalignant and malignant laryngeal lesions. First, small and thin lesions are difficult to evaluate by the histopathologic examination. Initial biopsies, mostly made at referring hospitals, are often not sufficient for a conclusive diagnosis. This effects the management at the definite surgery. Second, the margins of the specimens from laser excisions are difficult to evaluate due to tissue damage from the laser, leaving us in doubt whether the excision is radical or not.

The overall aim of our work with fluorescence investigations is to find a simple and quick in vivo fluorescence method to separate normal tissue from different histopathologically defined groups of premalignant and malignant laryngeal lesions online during the surgical procedure, in order to predict what type of lesion it is and to excise lesions with a safe but minimal margin.

In the present study, lesions in the larynx have been examined by autofluorescence emission spectroscopy and/or after administration of delta-aminolevulinic acid (ALA) orally, 3 hours prior to examination. The results have been correlated with corresponding measurements from normal mucosa and histopathologic examination of tissue samples from the investigated lesions. The aim of the study was to compare optical findings of different histopathologically defined groups of premalignant and malignant laryngeal lesions using autofluorescence and fluorescence following oral administration of ALA.

**MATERIALS AND METHODS**

**Subjects.** Thirty-four patients who were referred to the University Hospital due to laryngeal lesions with suspicion of dysplasia or malignancy were included in the study. The patients had a total of 47 laryngeal lesions: 41 on the vocal folds and 6 on the false vocal folds. The mean age of the patients was 65 years, with a range of 23 to 87 years. Twenty-six of the patients were men and 8 were women.

The local ethical committee has approved the study, and all patients were well informed about...
the investigations and gave their consent to participate in the study.

Fifteen of the patients, with 21 lesions, received ALA prior to the fluorescence examination, while the other 19 patients with 26 lesions were examined utilizing the tissue autofluorescence alone. ALA was dissolved in about 100 mL of orange juice immediately before use, in a dose corresponding to 5 or 7.5 mg/kg body weight of the patient. Orange juice was used as a solvent, as ALA is chemically stable in an acidic environment below pH 5.5. The ALA was administered 3 to 5 hours prior to the fluorescence examination.

**Measurements.** All fluorescence examinations were performed under general anesthesia at direct laryngoscopy, in combination with a biopsy for histopathologic examination. Most patients were examined sequentially using 2 different excitation wavelengths, 337 and 405 nm. The near ultraviolet light at 337-nm excited several endogenous tissue chromophores efficiently, providing tissue autofluorescence, while it did not excite the ALA-induced protoporphyrin IX (PpIX). The violet excitation wavelength at 405 nm was used as it well matches the PpIX fluorescence excitation peak (the Soret band). It also provided tissue autofluorescence, partially from other tissue chromophores than for the 337 nm excitation. Fluorescence spectra were recorded from the same tissue spots for both excitation wavelengths. The fluorescence emission from several tissue locations was recorded from each patient, both from the lesion/lesions and from normal tissue areas. Nondegraded tissue areas from the nonaffected vocal fold were used as the normal tissue areas. One biopsy sample was collected from each lesion, but not from normal tissue. Only measurements from locations that were biopsied were used. The site of measurement was documented with video and the subsequent biopsy was taken from the same site. The size of the biopsy was usually 2 to 3 mm, whereas the tip of the optical fiber was 0.6 mm.

The laser-induced fluorescence (LIF) measurements were performed with a clinically adapted fluorosensor developed at the Lund Medical Laser Centre. The system is presented in detail elsewhere, and will only briefly be described here. As an excitation source a small nitrogen laser was used. This laser emits 3-nanosecond long pulses of near ultraviolet light at 337 nm, with a pulse energy of 100 μJ and a repetition rate of 15 Hz. The excitation light was focused into a 600-μm optical fiber, which was inserted through the laryngoscope and held in gentle contact with the tissue. A fraction of the excitation light guided to the tissue generated tissue fluorescence. Part of the induced fluorescence light was collected by the same optical fiber, guided back to the instrument, and coupled out from the excitation path through a dichroic mirror. The fluorescence was spectrally dispersed in a polychromator and detected by an image-intensified diode array detector.

**Evaluations.** In order to evaluate the potential of fluorescence to demarcate these lesions, all measured areas must be classified. The lesions were clinically classified and normal areas included by visual inspection and palpation during the examination. Routine histopathology of the collected biopsy samples was used to classify all lesions. In the evaluation of the fluorescence emission signals, the fluorescence intensities measured at several wavelengths regions (indicated, for example, by $I(635)$ for the fluorescence intensity at 635 nm) were used. These intensities were derived as the integrated intensities in 10-nm band passes from the spectra using a computer program especially developed for this purpose at the Lund Medical Laser Centre.

The results are presented as a ratio of the fluorescence intensity for a lesion divided by the fluorescence intensity from an adjacent normal region in the same patient, for example $I(390)_{\text{lesion}}/I(390)_{\text{normal}}$. By summarizing these lesion-to-normal ratios, the discrimination for each patient is compared rather than the absolute intensities of the fluorescence from the various tissue types. Thus, irrelevant patient-to-patient variation in the absolute fluorescence intensity was drastically reduced.

In a previous study we used multivariate analysis of fluorescence spectra, and now 1 step closer to the clinic, we use a simplified method with selected wavelengths. Twenty-five of 29 specimens used in the final material in the current report were also used in the previous publication. However, completely different evaluation methods were used (PLS, PCA, PLS-DA, and logistic regression techniques), and the aim of that study was not to compare optical findings of different histopathologically defined groups.

**Excitation at 337 nm.** As PpIX was poorly excited at this wavelength, these spectra could be considered as pure tissue autofluorescence even in the patients who were given ALA. The ratio $I(430)/I(390)$ was evaluated for all patients. The wave-
lengths chosen for this ratio are influenced to the same extent by the presence of blood, meaning that the ratio will be independent of the reabsorption of the fluorescence by blood. Also the absolute fluorescence intensity at 390 nm was evaluated.

**Excitation at 405 nm.** For the ALA patients, we examined the ratio \( I(635)/I(500) \) in which \( I(635) \) was the PpIX related signal with the superimposed tissue autofluorescence contribution subtracted, and \( I(500) \) was the tissue autofluorescence at 500 nm. The tissue autofluorescence contribution to the signal at 635 nm was estimated by fitting an exponential curve,

\[
l = a \exp(-b) + c
\]

to the tissue autofluorescence in the regions 530 to 580 nm and 730 to 750 nm. The non-ALA patients were evaluated using the overall intensity \( I(500) \).

**Overall Evaluation.** In an attempt to summarize the results, all 4 fluorescence intensity measures were used in a combined evaluation. If one of the dimensionless lesion-to-normal ratios were deviant, the fluorescence result was considered pathologic. The cutoff levels used were \( I(390) < 0.2; I(430)/I(390) > 1.2; I(500) < 0.2; \) and \( I(635)/I(500) > 8 \).

**Histopathologic Evaluation.** Following the LIF measurements, the lesions were biopsied at the most pathologic looking site and at the same spot the fiber was placed when measurements were performed. It would of course have been preferable to confirm normal sites with biopsies for histopathologic examination, but as it involved the vocal folds it would have affected voice outcome for the patients and could not be done. For each patient, 1 to 3 biopsies were taken from measured sites. The specimens were sent for routine histopathologic examination. Later, 39 of 47 specimens were reevaluated blindly by an independent pathologist at another hospital. The correlation coefficient was 0.76. In the final evaluation we used only the specimens that were classified in the same group by the 2 pathology departments. This gave a final number of 11 in group I, 0 in group II, 11 in group III, and 7 in the carcinoma group.

Typical autofluorescence spectra following 337-nm excitation from the normal vocal fold and an adjacent squamous cell cancer lesion are shown in Figure 1. As can be seen in the figure, the overall fluorescence intensity is drastically lower for the carcinoma than for the normal vocal fold. In addition, the maximum is redshifted from 390 nm in the normal tissue to approximately 450 nm in the carcinoma. In the evaluation of the spectra, the integrated fluorescence intensity in a 10-nm interval around the central wavelengths of 390 and 431 nm was used.

Following excitation with 337 nm, \( I(390) \) and \( I(431)/I(390) \) were evaluated. The mean fluorescence intensities were first tested for all 4 histopathologic groups using ANOVA. If the test showed statistically significant difference, tests between the different histopathologic groups were performed with \( t \) test. Correlation between the 2 pathology departments was calculated from the 39 specimens examined at both departments.

**RESULTS**

Results from the routine histopathologic examination of the 47 laryngeal lesions are shown in Table 1.

Of the 47 biopsies, 39 were evaluated blindly by an independent pathologist at another hospital (8 specimens were lost, see earlier). The correlation between the 2 histopathologic examinations is shown in Table 1. A total of 29/39 were classified in the same group whereas 10/39 in different. The correlation coefficient was 0.76. In the final evaluation we used only the specimens that were classified in the same group by the 2 pathology departments. This gave a final number of 11 in group I, 0 in group II, 11 in group III, and 7 in the carcinoma group.

<table>
<thead>
<tr>
<th>Dept. B</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>Ca</th>
<th>Total B</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>11</td>
<td>–</td>
<td>2</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>II</td>
<td>1</td>
<td>–</td>
<td>5</td>
<td>–</td>
<td>6</td>
</tr>
<tr>
<td>III</td>
<td>1</td>
<td>–</td>
<td>11</td>
<td>–</td>
<td>12</td>
</tr>
<tr>
<td>Ca</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Missing</td>
<td>4</td>
<td>–</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total A</td>
<td>17</td>
<td>0</td>
<td>21</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.** Results from routine histopathologic examination of the collected biopsies (A, \( N = 47 \)) correlated to examination by an independent pathologist at another hospital (B, \( N = 39 \)).
cence intensities and standard deviations of the lesion-to-normal discrimination ratios were calculated. A ratio of 1 means that the function on an average has the same value in the lesion as in the adjacent normal tissue. The average intensity ratios at 390 nm were $I_{I} = 0.60$; $I_{III} = 0.37$; and for carcinomas $I_{Ca} = 0.20$. These differences were not significant (analysis of variance [ANOVA] = 0.082).

Results from the different fluorescence intensities were first tested for all 3 histopathologic groups using ANOVA. If the test showed statistically significant differences, tests between the different histopathologic groups were performed with $t$ test.

The ratio $I(431)/I(390)$ after excitation with 337 nm showed an average ratio close to 1 except for the carcinoma group (1.80, Figure 2).

Differences between the groups were found significant according to ANOVA test. The $t$ tests showed significant differences between I/cancer ($p = .039$) and III/cancer ($p = .04$).

Following 405-nm excitation, autofluorescence $I(500)$ and ALA effect $I(635)/I(500)$ were evaluated. The absolute intensities at 500 nm showed in all lesion types a lesion-to-normal ratio considerably lower than 1, and the ratio decreased with increasing grade of dysplasia (Figure 3). The result from the ANOVA for the absolute fluorescence intensity $I(500)$ was significant ($p = .002$). Further $t$ test showed that group I differed from group III ($p = .019$) and carcinoma ($p = .003$), but groups III and carcinoma did not differ ($p = .016$).

With excitation at 405 nm, the fluorescence intensity at 500 nm ($I(500)$) is due to tissue autofluorescence only. That makes it possible to evaluate $I(500)$ for both the ALA patients and the non-ALA patients. The contribution of ALA is seen only in the ratio $I(635)/I(500)$.

ALA was administered to 15 patients with a total of 21 lesions, but after exclusion of the specimens in which the histopathology gradings were different, there were only 14 lesions left, 5 in group I, 7 in group III, and only 2 in the carcinoma group (Table 2). All 29 lesions were evaluated for autofluorescence $I(500)$.

Group III and carcinoma group showed high mean lesion-to-normal ratios at $I(635)/I(500)$, but there were few lesions in the groups and thus large standard deviations (Table 2).

The fluorescence after administration of ALA did not add any information to the information already gained by the results from the autofluorescence measurements.

Overall evaluation with all lesion-to-normal ratios used, all 11 group I lesions were defined as normal whereas 16/18 group III lesions and carcinomas were considered pathologic. Calculating with group I as normal and group III and carcinoma as pathologic, ie, lesions demanding treatment, the sensitivity for the fluorescence method was 89% and specificity 100%.

---

**Table 2.** $I(635)/I(500)$ after excitation with wavelength 405 nm.

<table>
<thead>
<tr>
<th>Group</th>
<th>$N$</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>III</td>
<td>7</td>
<td>36</td>
<td>43</td>
</tr>
<tr>
<td>Ca</td>
<td>2</td>
<td>18</td>
<td>21</td>
</tr>
</tbody>
</table>

As the results for $I(635)/I(500)$ did not show any significant difference within the material, as tested with ANOVA ($p = .260$), $t$ tests were not performed for this ratio.
DISCUSSION

Clinical Issues. In the present study, we found that we could separate carcinoma from group I and group III lesions with 1 measurement (ratio $I_{430}/I_{390}$), and group I from group III/carcinoma, with another ($I_{500}$). Combining all measurements made to separate group I lesions from group III lesions/carcinoma, the sensitivity and specificity was found to be high. Group III lesions differed significantly from the carcinoma group as measured by the ratio $I_{430}/I_{390}$ but not with $I_{500}$. The difficulties in separating group III lesions and carcinoma have been discussed by Arens et al., who were not able to differentiate between various higher grades of epithelial dysplasia and between carcinoma in situ and invasive cancer. As both group III lesions and carcinoma are lesions demanding treatment, these results are, however, clinically useful.

In an effort to avoid excision of normal tissue during laser surgery for a cancer in a functional site such as the glottis, a noninvasive online diagnostic tool to separate normal and abnormal tissue and categorize premalignant and malignant lesion is of course highly interesting. This is not a new concept. In vivo diagnostics of malignancies in laryngeal mucosa has been used earlier by toluidine blue staining. The sensitivity of the staining test in detection of malignant lesions was 91% and the specificity 52%. Fluorescence techniques have been shown to accurately differentiate normal from abnormal oral tissues in vitro, normal mucosa from oral cavity neoplasia, and mucosal neoplasias of the oral cavity and pharynx. Early studies of fluorescence in laryngeal lesions showed that the technique could be useful in a clinical setting, and recent studies as well as the present study support this.

Technical Aspects. The present study presents an analysis of fluorescence spectra recorded from laryngeal lesions, indicating that it is possible to obtain diagnostic information of clinical impact using a few detection wavelength bands only, a technique that can be fully implemented in an imaging instrumentation. The diagnostic potential of the fluorescence technique is based on variations in concentration and distribution of the fluorescent molecules for various tissue types. Examples of tissue substances with strong fluorescence are collagen, elastin, NADH, vitamin derivatives, and porphyrins. Following the excitation, the molecules lose some of the excess energy gained by absorbing the photon through thermal processes. The fluorescence is therefore altered toward a longer wavelength as compared with the excitation light (Stokes shift). The fluorescence spectrum is characteristic for each molecule, but as they interact strongly with other molecules in the surrounding tissue, spectra from different molecules are becoming more alike. However, as we have shown in the present study, the resulting fluorescence spectra from different histopathologically defined groups of premalignant and malignant laryngeal lesions differ enough for tissue diagnostic purposes. By changing the excitation wavelength, one will alter the efficiency with which the various fluorescent tissue molecules will be excited and thereby the resulting fluorescence spectrum. By choosing excitation wavelengths, it is possible to obtain specific diagnostic information. The reason for choosing 337 nm as the excitation wavelength is that excitation light at 337 nm comes from an accessible and reliable light source that excite collagen, elastin, NADH, and porphyrins. At 390 and 431 nm, the damping from blood is identical; using the 431/390 ratio thus gives results unaffected by the presence of blood. The excitation wavelength at 405 nm was chosen as it well matches the PpIX fluorescence excitation peak.

One explanation for the difficulty to differentiate between group III lesions and carcinoma could be the superficial measurement of the fluorescence. The cell layers close to the basement membrane are probably not involved to the same extent as the more superficially located cells. The exact distance and direction the light has been travelling in the tissue (measurement geometry) is difficult to know in real clinical situations. The exact excitation light energy reaching the tissue might also be difficult to know with high accuracy, as there might be pulse-to-pulse fluctuations in the laser. One can minimize the influences of these 2 uncertainties in the analysis by evaluating the shape of the fluorescence spectrum, rather than the absolute intensity. Forming dimensionless ratios of 2 fluorescence intensities is 1 way to do exactly this.

In this study, the fluorescence lesion-to-normal ratios has been evaluated, rather than the mean values of the fluorescence functions per se. In this way, much of the irrelevant patient-to-patient fluctuation could be eliminated. It means, however, that it is not possible to measure a lesion in any patient and diagnose it. First it is necessary to measure a fluorescence spectrum from a normal
area to be able to calibrate the fluorescence scale for this particular patient. Sometimes patients have general mucosal changes making it difficult to find normal mucosa for comparison.

**Histopathology.** It would have been preferable to confirm normal sites with biopsies for histopathologic examination. However, the normal sites were in the glottis, and biopsies could not be performed as this would affect voice outcome for the patients. The local ethical committee had approved the study, but this approval did not include biopsy of a normal fold.

The tissue probed by fluorescence is often much smaller (diameter 0.6 mm, depth a few hundred micrometers) than that examined for histopathology (diameter 1–3 mm, depth more than a millimeter). The pathology report also concerns the worst region in this volume. If malignant involvement is present at any even so small region within the biopsy, it will be given the diagnosis of malignant lesion. This possible heterogeneity within the small biopsy sample may cause difficulties in the evaluation of the fluorescence signals, even though the positions of the fluorescence recording and biopsy was very well matched.

It is necessary to calibrate the fluorescence diagnostic system by correlating to a correct diagnosis. Histologic examination is the most important tool that will establish a definite diagnosis of a precancerous lesion or a carcinoma. Although it is the golden standard, the diagnostic diversity can be the effect of several independent factors, and even the highest imaginable degree of standardization of a scoring system will not fully eliminate observer bias. Attention is drawn to particular diagnostic problems concerning biopsy quality, absence of step sectioning and changes that were difficult to discern. In our study, no attempt to any standardization of the score system or classification was made between the 2 departments of pathology. The result has been a satisfactory correspondence between the result of first and second gradings ($\lambda = 0.76$). To minimize the risk of any systemic error in the study, we only used the specimens in which the 2 gradings were the same. With regard to the grading of dysplasia, several investigations have shown an interobserver bias of a size that makes it necessary to come up with a reliable diagnosis before the prognosis can be taken into account. Immunohistochemical and molecular methods proved particularly important new diagnostic tools to support the histologic diagnosis.

**CONCLUSION**

The results from the present study show that there are differences in autofluorescence emission spectroscopy between laryngeal lesions with non or mild dysplasia (I), severe dysplasia or cancer in situ (III) and carcinoma. Following excitation with 337 nm, the ratio $I(430)/I(390)$ was significantly different when groups I and III were compared to carcinoma. $I(500)$ after 405-nm excitation differed significantly between groups I and III, as well as between groups I and carcinoma. In this study, there was no benefit of using ALA. However, the low total number of specimens in the final evaluation in which ALA was administered may have affected this result. We are further examining the effect of ALA in an ongoing study. Using optimal cutoff levels for all evaluations made, 11 of 11 group I lesions were defined as normal, whereas 16 of 18 group III lesions and carcinomas were considered pathologic. Thus, a theoretical calculation of the overall results gave the sensitivity 89% and specificity 100%.

**Acknowledgments.** We greatly acknowledge the support by Prof. Sune Svanberg.

**REFERENCES**