In conclusion, we have extended the electrical discharge-based flow visualization technique to visualize the flow fields around blunt cones in a hypersonic shock tunnel. The experimental results obtained in HST2 shock tunnel at IISc at a flow Mach number of 5.75 are compared with the numerically computed results for a 120° apex angle blunt cone. The visualized detached strong shock wave ahead of the body matches with predicted shock wave with a matching stand-off distance. The variation of shock stand-off distance as a function of the binary scaling parameter and the stagnation enthalpy is currently being investigated.

1. Jagadeesh, G., Srinivasa Rao, B. R., Nagashetty, K., Reddy, N. M. and Reddy, K. P. J., *Curr. Sci.*, 1996, **71**, 128–130.
2. Jagadeesh, G., Srinivasa Rao, B. R., Nagashetty, K., Reddy, N. M. and Reddy, K. P. J., *J. Flow Visualization Image Processing*, 1997, **4**, 51–57.
3. Jagadeesh, G., Reddy, N. M., Nagashetty, K. and Reddy, K. P. J., AIAA Paper 98-2601, June 1998.
4. Jagadeesh, G., Reddy, N. M., Nagashetty, K. and Reddy, K. P. J., *J. Spacecraft Rockets*, 2000, **37**, 137–139.

Received 9 May 2000; accepted 20 September 2000

A pilot study on the use of autofluorescence spectroscopy for diagnosis of the cancer of human oral cavity

S. K. Majumder*, S. K. Mohanty*, N. Ghosh*, P. K. Gupta*,‡, D. K. Jain† and Fareed Khan†

Biomedical Applications Section, Centre for Advanced Technology, Indore 452 013, India
†Department of Surgery, M. Y. Hospital, Indore 452 001, India

The results of a pilot study to evaluate the potential of autofluorescence spectroscopy for the diagnosis of the cancer of oral cavity are presented. The study was carried out using a N_2 laser-based system developed in-house and involved 25 patients with histopathologically confirmed squamous cell carcinoma of oral cavity. A general multivariate statistical algorithm was developed to analyse and extract clinically useful information from the oral tissue spectra acquired *in vivo*. The algorithm could differentiate, over the sample size investigated, the squamous cell carcinoma of the oral cavity from normal squamous tissue with a sensitivity and specificity of 86% and 63%, respectively towards cancer. The relatively poor specificity is presumably because most of the patients investigated had advanced cancers, due to which some of the visually uninvolved sites treated as normal may not be truly normal.

ORAL cancer is one of the most common cancers in India and several other South Asian countries and its incidence is on a rise due to consumption of tobacco and pan masala. There is, therefore, a growing need to develop sensitive and less invasive methods for screening cancerous or precancerous conditions of oral cavity.

Considerable work has been carried out on the use of laser-induced fluorescence (LIF) from native tissues for diagnosis of cancer[1–4]. This technique has the potential to probe *in situ* and quantitatively the biochemical changes that occur, as the tissue becomes neoplastic. Although LIF technique is particularly well suited for early detection of the cancer of oral cavity, due to easy accessibility of this organ, the evaluation of this diagnostic modality has not received requisite attention. In an earlier study on tissues of the oral cavity resected at biopsy[5], we found significant differences in the fluorescence signatures of malignant and normal sites. Based on the spectrally integrated intensity alone as the discrimination parameter, a sensitivity and specificity of ~90% was obtained in this study involving 47 patients.

In this communication, we report the results of a pilot study on 25 patients with histopathologically confirmed squamous cell carcinoma (SCC) of the oral cavity. A general multivariate statistical algorithm was developed to analyse and extract clinically useful information from the oral tissue spectra acquired *in vivo*. The algorithm could differentiate the SCC of the oral cavity from normal squamous tissues with a sensitivity and specificity towards cancer, of 86% and 63%, respectively. The relatively poor specificity is presumably because most of the patients investigated had advanced cancers, due to which some of the visually uninvolved sites treated as normal may not be truly normal.

A schematic of the N_2 laser-based portable fluorimeter developed for *in vivo* clinical studies is shown in Figure 1. It consists of a sealed-off N_2 laser (7 ns, 80 μJ, 10 Hz), a spectrograph coupled to a gateable intensified CCD camera (4 Quik 05A, Stanford Computer Optics Inc, USA) and a fiber optic probe to excite and collect fluorescence from the tissue. The probe, developed in-house, is an optical fiber bundle with two legs; one contains a single quartz fiber (NA 0.22, core diameter 400 μm) and the other contains six quartz fibers (NA 0.22, core diameter 400 μm). The two legs merge to form a common fiber bundle which consists of a central fiber, surrounded by a circular array of six fibers. The central fiber delivers excitation light to the tissue surface and the six surrounding fibers collect tissue

‡For correspondence. (e-mail: pkgupta@cat.ernet.in)

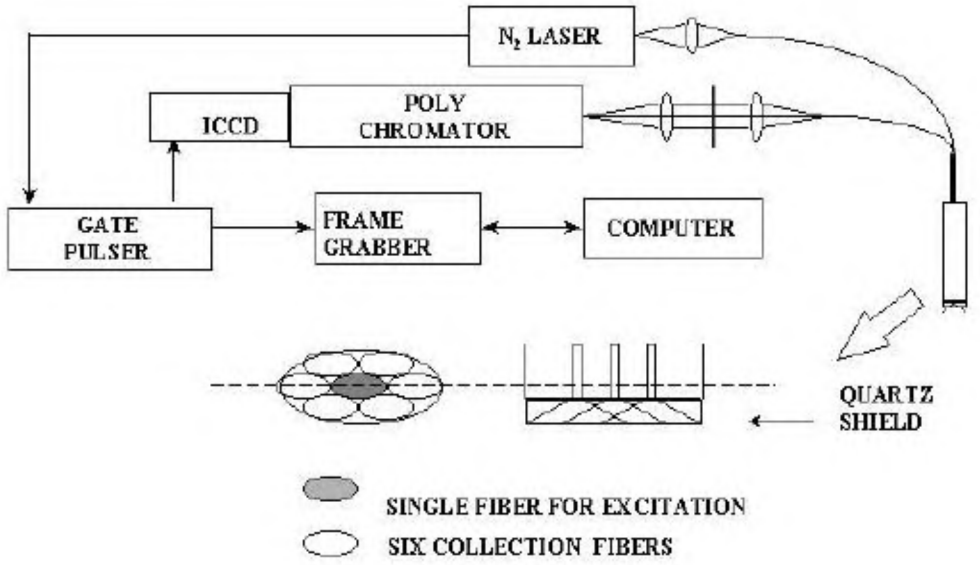


Figure 1. A schematic of the N_2 laser-based portable fluorimeter developed at the Centre for Advanced Technology, Indore, for *in vivo* clinical studies.

fluorescence from the surface area directly illuminated by the excitation light. The light coming from the distal ends of the six collection fibers was imaged via a quartz lens system on the entrance slit of the spectrograph whose output end was coupled to the gateable intensified CCD camera. The delay between N_2 laser pulse and the CCD camera shutter was adjusted to pick up fluorescence and discriminate against ambient background light. The gate-width used was 200 ns. The common end of the fiber bundle was enclosed in an SS tube (9 mm outer diameter and 60 mm length). The tip of the probe was shielded by a 2 mm thick quartz optical flat, to provide a fixed distance between tissues and the fibers for improved collection of fluorescence and also to protect contamination of the fiber tips with body fluids. The spectral data acquisition was computer controlled. The autofluorescence spectra were recorded with the tip of the fiber optic probe placed in contact with the tissue surface.

The study involved 25 patients with histopathologically confirmed SCC of the oral cavity. An informed consent was obtained from all the patients who underwent spectroscopic examination.

Prior to recording spectra from a patient, the probe was disinfected with CIDEX (Johnson and Johnson, India), washed with PBS and cleaned dry with a piece of sterilized cotton. A background spectrum was acquired with the probe placed in air. It was subtracted from all subsequently acquired spectra. From each site, spectra were recorded for 30 s (i.e. 300 pulses) and were averaged to yield a single spectrum per site. The spectra were collected in the 350–700 nm spectral range with a resolution of ~2 nm. The signal-to-noise ratio was >100:1 at the fluorescence maximum. Before and after each measurement of tissue fluorescence, a spectrum was also acquired with the probe placed on the face of a quartz cuvette containing a known concentration of Rh-6G in ethyl alcohol. The spectrum from each tissue site was normalized with respect to the peak of the averaged Rh-6G spectrum. The intensity of fluorescence from each site is reported in this normalized unit.

Spectra were collected from a total of 208 sites, of which 126 were from SCC and the remaining 82 were from the adjacent uninvolved normal region. On an average, 5 spectra from the cancerous tissue sites and 4 spectra from the normal tissue sites were recorded. In each patient, the normal tissue sites interrogated were from the adjacent apparently uninvolved region of the oral cavity.

The autofluorescence spectra recorded from different cancerous and normal sites of the oral cavity of a patient are shown in Figure 2 *a* and *b*, respectively. The considerable site-to-site variation in the spectra is apparent. The mean spectra for the cancerous sites and the normal sites from seven patients selected at random are shown in Figure 3 *a* and *b*, respectively. The differences in the spectra from some of the normal and cancerous tissue sites are not that apparent, being masked by the large intra- and inter-patient variability in the intensity and line shape. While some of this variation may represent intrinsic variation in tissue fluorescence, the variable nature of the contact of the probe with the tissue surface in the clinical situation, will also add to the variation. It is pertinent to note that in the *in vitro* studies on oral cavity tissues[5], where the variability due to the nature of contact of the probe with the tissue surface will be minimal, a percentage variation ($\sigma/\bar{x}$) of ~30% was observed in the spectrally integrated intensities ($\sum I$), from different sites of normal or cancerous tissues over the total sample size investigated. Here, $\bar{x}$ is the

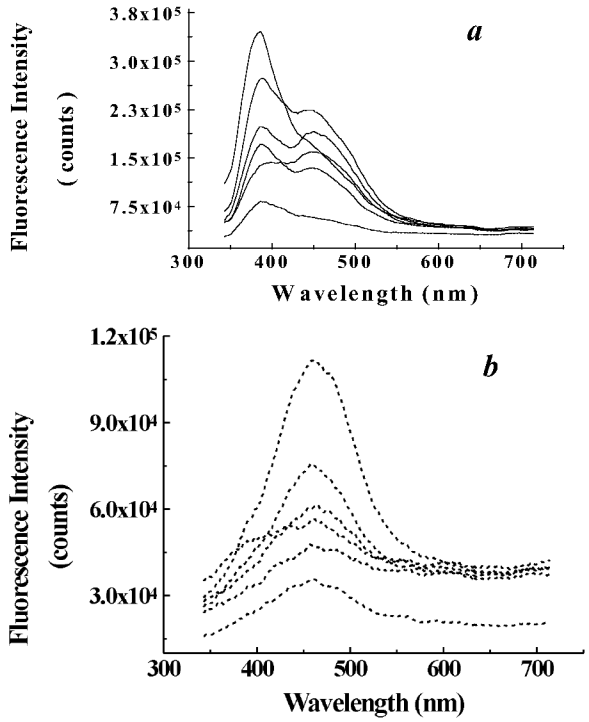


Figure 2. N_2 laser-excited autofluorescence spectra recorded from *a*, squamous cell carcinoma tissue sites; and *b*, uninvolved tissue sites of the same patient.

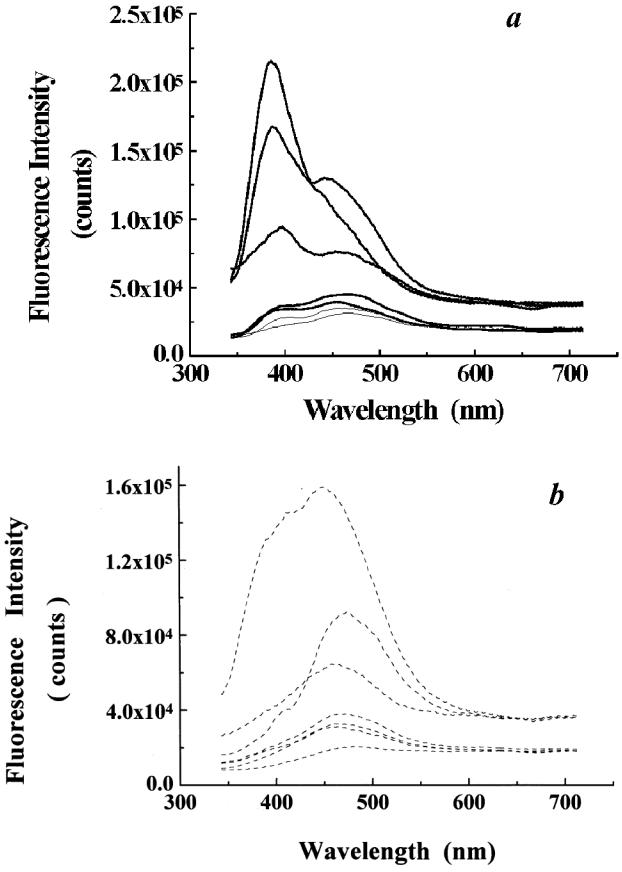


Figure 3. Mean autofluorescence spectra of (*a*) squamous cell carcinoma tissue sites; and (*b*) uninvolved tissue sites, from seven patients selected at random.

mean of $\sum I$ values from different sites of one category and σ is the standard deviation. In comparison, in the present *in vivo* study, the percentage variation $(\sigma/\bar{x})$ in $(\sum I)$ was ~60%. In order to ensure good discrimination, it is necessary to minimize these variations which may obscure the intercategory differences. In order to minimize the inter- and intra-patient variability, a two-step procedure was adopted. In the first step, the spectrum from each site was normalized with respect to the spectrally integrated intensity from that site. This normalization is expected to remove from the spectrum, the influence of the variable nature of the contact of the probe with the tissue surface and also lead to enhancement of spectral features. Next, mean-scaling was performed by calculating the mean spectrum for a patient (the average of the mean spectra from both cancerous and normal sites in that patient) and subtracting it from the spectrum of every site of the patient. Since mean-scaling displays the differences in the fluorescence spectra for a particular diagnostic category with respect to the mean spectrum for that patient, it is expected to maximize the spectral differences between the two diagnostic categories. Indeed, normalization of the spectra with respect to their respective spectrally integrated intensities followed by mean scaling made the spectral differences between the two diagnostic categories much more apparent. In Figure 4 *a* and *b*, we show the autofluorescence spectra of the cancerous and normal tissue sites after normalization. The corresponding normalized, mean-scaled spectra are shown in Figure 5. As the mean normalized spectrum has been subtracted from each normalized spectrum corresponding to cancerous and normal oral cavity tissue sites from this patient, the mean now lies at $y = 0$ over the entire emission wavelength range. Evaluation of Figure 5 indicates that between 350 and 500 nm, the normalized, mean-scaled fluorescence intensity of the normal squamous tissue is less than the mean, and that of SCC is greater than the mean.

In order to extract the diagnostic content of the autofluorescence spectra of the cancerous and normal oral tissues, Principal Component Analysis (PCA)[6] was employed. This procedure dimensionally reduces the spectral data into a smaller orthogonal set of linear combinations of the emission variables that account for most of the variance of the spectral data set. For PCA, an input data matrix $X(r \times c)$ was created, such that each row of the matrix corresponded to the processed fluorescence spectrum from each tissue site and each column corresponded to the processed fluorescence intensity at each emission wavelength.

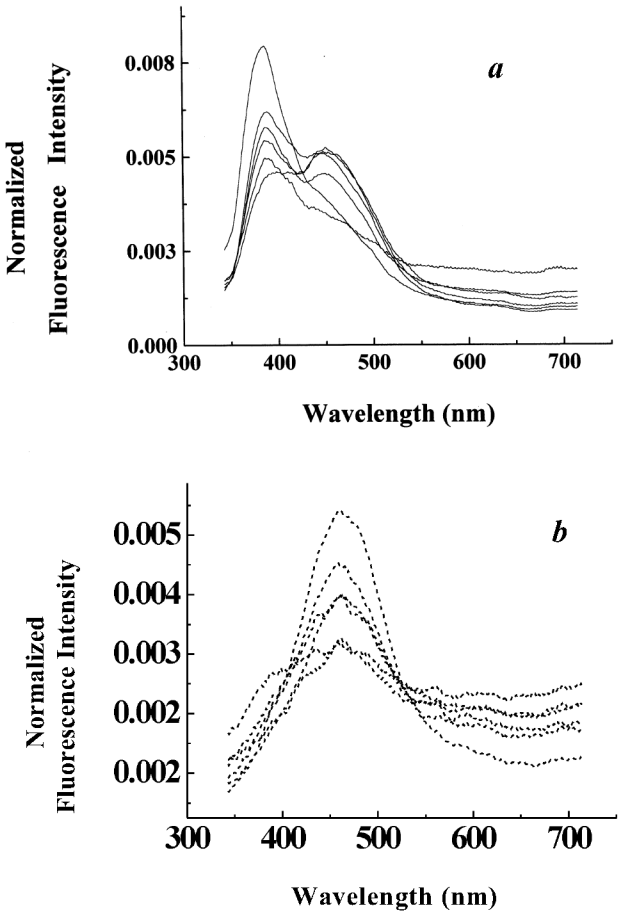


Figure 4. Normalized autofluorescence spectra from (*a*) squamous cell carcinoma tissue sites, and (*b*) uninvolved tissue sites of the same patient.

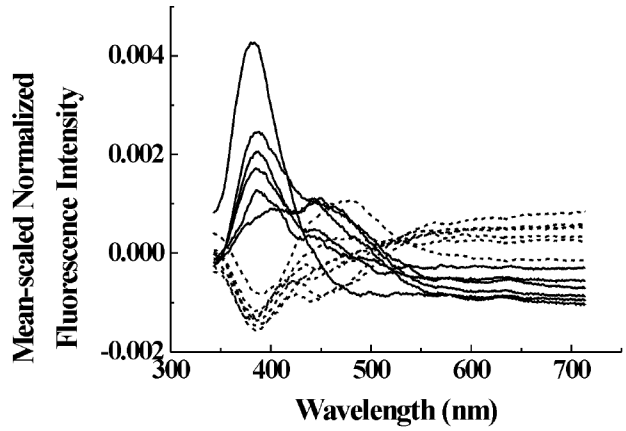


Figure 5. Mean-scaled, normalized autofluorescence spectra from squamous cell carcinoma tissue sites (solid line) and from normal (uninvolved) squamous tissue sites (dashed line) of the oral cavity of the same patient.

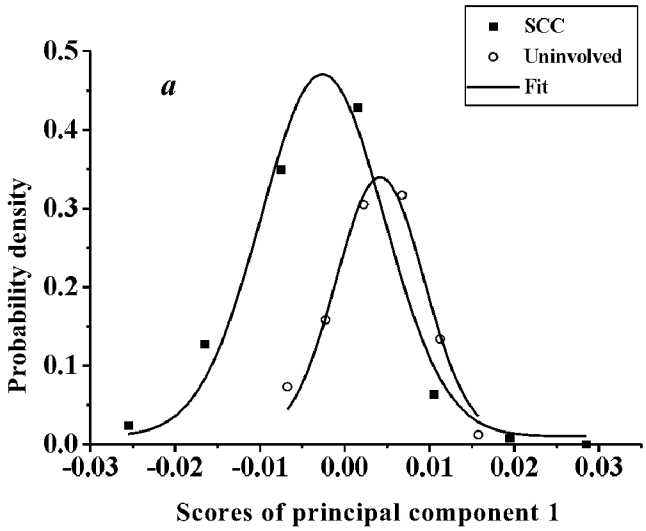
The average values of principal components were calculated for each principal component for cancerous and normal squamous tissue. An unpaired two-tailed Student's t test[7] was employed to test the statistical significance of the differences between the means of the principal component scores, for cancerous and normal oral cavity tissue. PCA of normalized, mean-scaled spectra resulted in four principal components that collectively accounted for 99% of the total variance of the spectral data. Of the four principal components, only two (PC-1 and PC-3) were found to have significantly different ($P < 0.0001$) values for SCC and normal squamous tissue. The differences in the values for the other two components (PC-2 and PC-4) for the two diagnostic categories were not significant ($P > 0.01$). Therefore, further analysis was carried out using PC-1 and PC-3, which together accounted for 67% of the total variance, PC-1 accounting for 63% of the total variance and PC-3, the remaining 4%.

Logistic discrimination was used to develop a classification algorithm to discriminate between the SCC tissue sites and the normal squamous tissue sites. For two diagnostic categories, G_1 and G_2, the posterior probability of being a member of G_1, given scores x_1 and x_2 for the two principal components, is given by[8]:

$$P(G_1 \mid x_1, x_2) = \frac{P(x_1, x_2 \mid G_1)P(G_1)C(2 \mid 1)}{P(x_1, x_2 \mid G_1)P(G_1)C(2 \mid 1) + P(x_1, x_2 \mid G_2)P(G_2)C(1 \mid 2)},$$

where, $P(x_1, x_2|G_i)$ is the conditional joint probability that a tissue sample of type i will have scores x_1 and x_2 for the two principal components PC-1 and PC-3, respectively, and $P(G_i)$ is the prior probability of finding tissue type i in the sample population. $P(x_1, x_2 \mid G_i)$ is the product of the conditional probability of the two individual principal components of having scores x_1 and x_2, respectively. The prior probability is an estimate of the likelihood that a tissue sample of type i belongs to a particular group when no information about it is available. The total number of cases in each group were used to estimate the prior probabilities, assuming the sample size to be representative of sample population. $C(i \mid j)$ is the cost of misclassifying a sample into group j, when it actually belongs to group i. The cost of misclassification can be varied from 0 to 1 with the condition that, the sum of the cost of misclassification for all the diagnostic categories under consideration equals one.

The posterior probabilities were determined by calculating the prior and the conditional joint probabilities and the costs of misclassification of the SCC tissue sites. The cost of misclassification of a particular tissue type was varied from 0 to 1 in increments of 0.05, and the optimal cost was identified when the number of misclassified samples was a minimum. If there was more than one cost at which the total number of misclassified samples was a minimum, then the cost at which the number of SCC tissue sites misclassified was a minimum was selected. The conditional probabilities



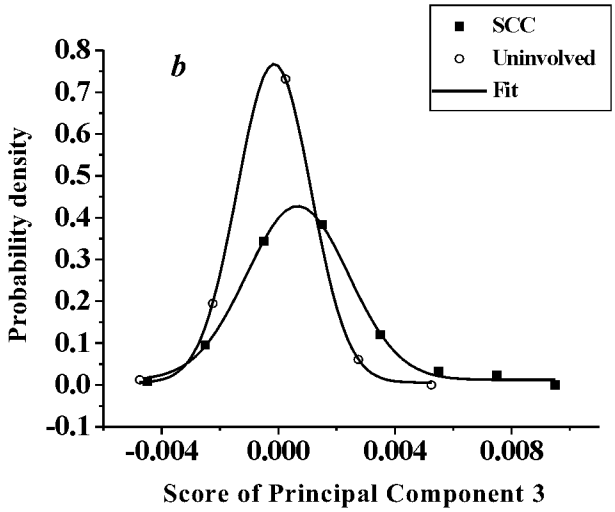


Figure 6. Measured probability distributions of the scores of *a*, principal component-1 (PC-1) and *b*, principal component-3 (PC-3), for squamous cell carcinoma (SCC) tissue sites and normal (uninvolved) squamous tissue sites. Solid line shows the fit to normal probability density function.

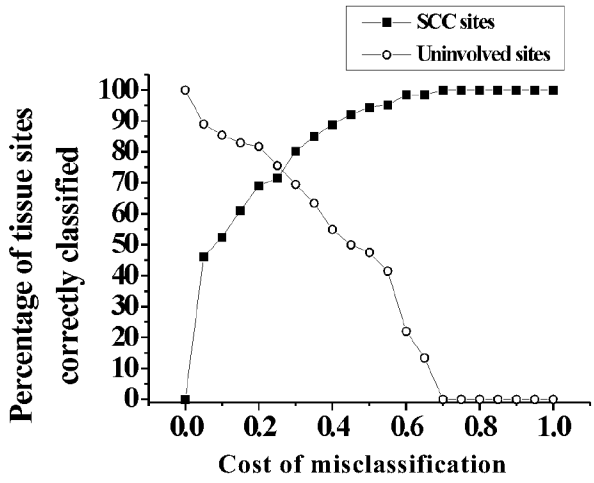


Figure 7. Percentage of the squamous cell carcinoma (SCC) and the normal squamous tissue sites classified correctly versus cost of misclassification of SCC tissue sites.

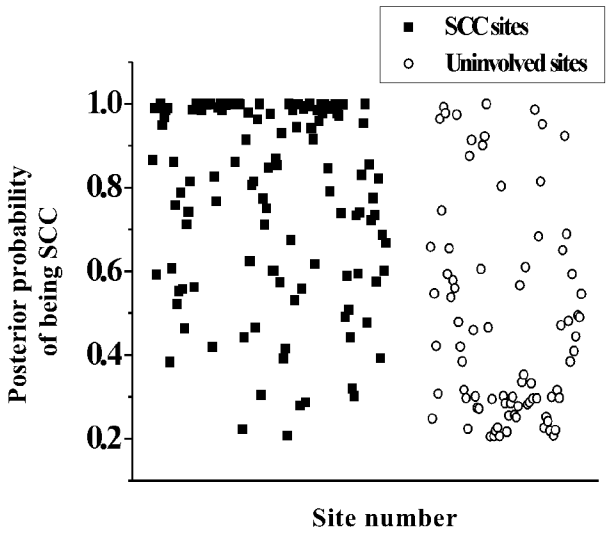


Figure 8. Posterior probabilities of being classified as squamous cell carcinoma (SCC) for all tissue sites investigated.

were determined from the probability distribution of the principal component scores of each tissue type fitted to a normal probability density function[9], which is characterized by peak position (x) and width (σ). The best fit of the normal probability density function to the probability distribution of the principal component scores of each tissue type was obtained in a least square sense, using x and σ as free parameters of the fit.

Figure 6 shows the measured probability distribution of PC-1 and PC-3 for normal squamous tissues and SCC of the oral cavity and its fit to a normal probability density function. Prior probabilities, determined by calculating the percentage of each tissue type from the data set, were 40% for normal squamous tissues and 60% for squamous carcinoma. The cost of misclassification was optimized at 0.36. Figure 7 shows the percentage of normal and cancerous oral tissues, correctly classified versus cost of misclassification of cancerous tissues, for the entire data set. It is evident from the figure that an increase in the cost of misclassification of cancerous tissues results in an increase in the proportion of correctly classified cancerous oral tissues and a decrease in the proportion of correctly classified normal oral tissues. In Figure 8 we plot the posterior probabilities for all the cancerous and the normal tissue sites of being classified as SCC. The figure indicates that 86% of the cancerous oral tissues have a posterior probability >0.5 and 63% of the normal oral tissues have a posterior probability <0.5.

The reason for the relatively lower specificity values appears to be the fact that most of the patients who participated in this study were at an advanced stage of malignancy. Some of the visually uninvolved sites

assumed to be normal may not be truly normal, since it is known that the 'normal' appearing region surrounding the cancerous tumour may have biochemical changes occurring due to the field-effect of the malignancy[10]. Indeed, when the discrimination analysis was carried out on the basis of the spectra averaged over all cancerous sites and the spectra averaged over all normal sites from a patient, a sensitivity and specificity towards cancer of 100% was obtained. The remarkably good results obtained on site-averaged spectra may suggest that a few of the normal sites had signatures very different from the other sites of the group.

This pilot study shows that LIF can discriminate SCC of oral tissues from normal squamous tissues. However, since the objective of the LIF-based system is the early diagnosis of cancer, in the next stage of the study, a qualified doctor will locate the suspected region of the oral cavity using visual assessment or other conventional means. Before the biopsy is taken, autofluorescence spectra will be recorded from the suspected sites as well as several other surrounding normal sites. The biopsy specimens from the suspected regions will then be subjected to histopathological examination and the spectral information will be correlated with the tissue histopathology to evaluate the true clinical potential of the LIF-based approach.

To conclude, a pilot study to evaluate the clinical potential of autofluorescence spectroscopy for diagnosis of cancer of oral cavity has been carried out using a system developed at the Centre for Advanced Technology (CAT), Indore. In this study involving 25 patients with histopathologically confirmed cancer of oral cavity, a sensitivity and specificity of 86% and 63%, respectively, towards cancer was obtained. The relatively lower specificity value appears to be due to the advanced stage of cancer in the patients investigated, due to which uninvolved regions treated as normal may not be truly normal. It will be desirable to carry out these studies in patients with early stages of cancer, to truly evaluate the diagnostic potential of the technique.

1. Kincade, K., *Laser Focus World*, February 1996, pp. 71–79.
2. Servick-Muraca, E. and Benaron, D. (eds), *OSA Trends in Optics and Photonics on Biomedical Optical Spectroscopy and Diagnostics*, Optical Society of America, Washington DC, 1996, vol. 3.
3. Richards Kortum, R. and Servick-Muraca, E., *Annu. Rev. Phys. Chem.*, 1996, **47**, 556–606.
4. Wagnieres, G. A., Star, W. M. and Wilson, B. C., *Photochem. Photobiol.*, 1998, **68**, 603–632.
5. Majumder, S. K., Gupta, P. K. and Uppal, A., *Lasers Life Sci.*, 1999, **8**, 221–227.
6. Johnson, R. A. and Wichern, D. W. (eds), *Applied Multivariate Statistical Analysis*, Prentice-Hall International, Inc., New Jersey, 1988, 2nd edn, Chap 8, pp. 340–377.
7. Kleinbaum, D. G., Kupper, L. L. and Muller, K. E. (eds.), *Applied Regression and other Multivariable Methods*, Duxbury Press, Belmont, California, 1988, 2nd edn, Chap. 3, pp. 16–40.
8. Ramanujam, N., Mitchell, M. F., Mahadevan-Jensen, A., Thomsen, S., Malpica, A., Wright, T., Atkinson, N. and Richards-Kortum, R., *Lasers Surg. Med.*, 1996, **19**, 46–62.