

# Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

## Raman exfoliative cytology for oral precancer diagnosis

Aditi Sahu  
Poonam Gera  
Venkatesh Pai  
Abhishek Dubey  
Gunjan Tyagi  
Mandavi Waghmare  
Sandeep Pagare  
Manoj Mahimkar  
C. Murali Krishna

**SPIE.**

Aditi Sahu, Poonam Gera, Venkatesh Pai, Abhishek Dubey, Gunjan Tyagi, Mandavi Waghmare, Sandeep Pagare, Manoj Mahimkar, C. Murali Krishna, "Raman exfoliative cytology for oral precancer diagnosis," *J. Biomed. Opt.* **22**(11), 115003 (2017), doi: 10.1117/1.JBO.22.11.115003.

# Raman exfoliative cytology for oral precancer diagnosis

Aditi Sahu,<sup>a,b,†,‡</sup> Poonam Gera,<sup>c,‡</sup> Venkatesh Pai,<sup>d</sup> Abhishek Dubey,<sup>d</sup> Gunjan Tyagi,<sup>a</sup> Mandavi Waghmare,<sup>e</sup> Sandeep Pagare,<sup>e</sup> Manoj Mahimkar,<sup>b,d</sup> and C. Murali Krishna<sup>a,b,\*</sup>

<sup>a</sup>Tata Memorial Center, Advanced Center for Treatment, Research and Education in Cancer, Chilkapati Lab, Kharghar, Navi Mumbai, India

<sup>b</sup>Homi Bhabha National Institute, Training School Complex, Anushakti Nagar, Mumbai, India

<sup>c</sup>Tata Memorial Center, Advanced Center for Treatment, Research and Education in Cancer, Tissue Biorepository, Kharghar, Navi Mumbai, India

<sup>d</sup>Tata Memorial Center, Advanced Center for Treatment, Research and Education in Cancer, Mahimkar Lab, Kharghar, Navi Mumbai, India

<sup>e</sup>D. Y. Patil School of Dentistry, Department of Oral medicine and Radiology, Nerul, Navi Mumbai, India

**Abstract.** Oral premalignant lesions (OPLs) such as leukoplakia, erythroplakia, and oral submucous fibrosis, often precede oral cancer. Screening and management of these premalignant conditions can improve prognosis. Raman spectroscopy has previously demonstrated potential in the diagnosis of oral premalignant conditions (*in vivo*), detected viral infection, and identified cancer in both oral and cervical exfoliated cells (*ex vivo*). The potential of Raman exfoliative cytology (REC) in identifying premalignant conditions was investigated. Oral exfoliated samples were collected from healthy volunteers ( $n = 20$ ), healthy volunteers with tobacco habits ( $n = 20$ ), and oral premalignant conditions ( $n = 27$ , OPL) using Cytobrush. Spectra were acquired using Raman microprobe. Spectral acquisition parameters were:  $\lambda_{\text{ex}}$ : 785 nm, laser power: 40 mW, acquisition time: 15 s, and average: 3. Postspectral acquisition, cell pellet was subjected to Pap staining. Multivariate analysis was carried out using principal component analysis and principal component-linear discriminant analysis using both spectra- and patient-wise approaches in three- and two-group models. OPLs could be identified with  $\sim 77\%$  (spectra-wise) and  $\sim 70\%$  (patient-wise) sensitivity in the three-group model while with 86% (spectra-wise) and 83% (patient-wise) in the two-group model. Use of histopathologically confirmed premalignant cases and better sampling devices may help in development of improved standard models and also enhance the sensitivity of the method. Future longitudinal studies can help validate potential of REC in screening and monitoring high-risk populations and prognosis prediction of premalignant lesions. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.22.11.115003]

Keywords: oral cancer; exfoliated cells; oral premalignant conditions; Raman spectroscopy; principal component analysis; principal component-linear discriminant analysis; Raman exfoliative cytology.

Paper 170385PRR received Jun. 25, 2017; accepted for publication Oct. 19, 2017; published online Nov. 14, 2017.

## 1 Introduction

Oral cancers are a major health problem worldwide, especially in developing countries such as India. Tobacco (smokeless/smoking), areca nut, and alcohol consumption are the major etiological factors for development of oral cancers.<sup>1–4</sup> The five-year survival rates for these cancers are currently around 50% to 55%.<sup>5</sup> These rates depend mainly on the stage of cancer at diagnosis.<sup>6</sup> It is known that early diagnosis can improve overall outcomes of oral cancer. During carcinogenesis, while some early oral cancers demonstrate subtle presentation, some invasive forms of oral squamous cell carcinoma (OSCC) often present with clinically perceptible mucosal changes. OSCC develops as a result of accumulation of various molecular and genetic alterations.<sup>7</sup> The cellular abnormalities accumulate and give rise to clinical alterations in the oral epithelium, which is termed as precancer. These symptomatic precancers mainly occur as persistent red or white patches (erythroplakia/leukoplakia), stiffening of mucosa known as oral submucous fibrosis (OSMF), mucosal changes, such as tobacco pouch keratosis

(TPK) and lichen planus, or nonhealing ulcers.<sup>8–10</sup> These precancerous lesions and conditions may demonstrate 6% to 45% dysplasia histopathologically.<sup>11</sup> These conditions also carry a higher than normal risk of malignant transformation, and this transformation rate ranges between 5% and 37%.<sup>12,13</sup>

Biopsy followed by histopathology of suspicious mucosal areas is the gold standard approach for diagnosing epithelial dysplasia in high-risk lesions.<sup>14</sup> Removal of lesions with moderate or severe dysplasia is advocated, whereas mild dysplasia is followed up for reversal or progression. Due to associated disadvantages such as invasiveness, subjectivity, longer output times, and less patient compliance, biopsy may not be a suitable approach for screening and monitoring of low-risk, benign, or innocuous-appearing lesions. Such lesions also demonstrate a lower degree of mild dysplasia; thus, histopathological grading of these lesions can be highly subjective.<sup>15,16</sup> Further, clinical and histological features alone cannot accurately predict whether potentially malignant disorders of the oral mucosa remain stable, regress, or progress to malignancy.<sup>16,17</sup> Thus, several adjunctive methods have been introduced to facilitate detection and assessment of these lesions. These include Toluidine blue and visualization aids, such as Vizilite, VELscope, Microlux DL, and Orascoptic DK.<sup>18</sup> Exfoliative cytology being a simple and less-invasive procedure of sample collection has also

\*Address all correspondence to: C. Murali Krishna, E-mail: [mchilakapati@actrec.gov.in](mailto:mchilakapati@actrec.gov.in)

†The author is now a research scholar in The Milind Rajadhyaksha Lab at Memorial Sloan Kettering Cancer Center, New York, New York, United States.

‡The authors contributed equally to this paper.

been actively pursued for precancer and cancer diagnosis. Conventional oral exfoliative cytology was initially explored, but low sensitivity rates hampered clinical applicability. With the advent of modified sample collection and staining methods, introduction of computerized image analysis, deoxyribonucleic acid (DNA) cytometry, and addition of molecular tests, such as detection of p53 mutation and loss of heterozygosity, this area has received renewed interest in the last few years.<sup>19–22</sup> Ancillary methods, such as DNA image cytometry and AgNOR analysis, are undertaken to allow more precise cytological diagnosis. Studies combining cytology with DNA cytometric evaluation have shown significantly higher sensitivity and specificity. Such multimodal cell analyses may also be useful when only few cytologically abnormal (atypical) cells are present.<sup>23–27</sup> However, the clinical utility of these methods is still under validation. Alternatively, these cells can be subjected to optical spectroscopic analyses, which yield inherent biochemical signatures. As biochemical changes precede morphological changes, spectroscopic methods may facilitate early identification of oral precancers, detect their propensity for malignant transformation, and predict response to treatment.

Optical spectroscopic approaches, such as Raman spectroscopy (RS), yield global biochemical fingerprint of the sample. Due to high sensitivity, RS has shown potential in the diagnosis of various diseases, including cancers.<sup>28</sup> Although *in vivo* RS has shown promise in classification of several cancers and precancers,<sup>29,30</sup> including oral premalignant and cancer conditions,<sup>31</sup> RS of exfoliated cells has several practical advantages. RS of exfoliative cytology has shown potential in classifying normal and abnormal cells and detected human papillomavirus infection in cervical exfoliated cells.<sup>32,33</sup> RS has also differentiated between healthy and oral cancer cells<sup>34</sup> and shown preliminary classification between healthy and premalignant exfoliated cells in an exploratory study.<sup>35</sup> In this study, the feasibility of differentiating exfoliated cells from healthy volunteers, healthy tobacco users, and subjects with premalignant lesions was explored using RS coupled with cytopathology, this multimodal approach was termed as Raman exfoliative cytology (REC). Cytology using Pap staining was carried out on same cell samples after spectra acquisition to understand the cellular characteristics of the sample. This approach will help in assessing potential of spectroscopic and cytological methods for prospective risk assessment of premalignant or clinically suspicious lesions. Findings are presented in this study.

## 2 Materials and Methods

### 2.1 Exfoliative Cytology

Oral exfoliated cell specimens were obtained from healthy subjects without tobacco habits, i.e., HVs ( $n = 20$ ), healthy subjects with tobacco habits, i.e., habit controls or HT ( $n = 20$ ), and subjects with oral premalignant lesions (OPL), such as leukoplakia, TPK, or OSMF  $n = 27$  visiting Advanced Center for Treatment, Research and Education in Cancer, Tata Memorial Center (TMC), Kharghar, Navi Mumbai, and D. Y. Patil Dental College and Hospital, Nerul, Navi Mumbai, after obtaining written informed consent. All subjects were clinically diagnosed as either healthy (no mucosal abnormalities) or having premalignant lesions. The premalignant lesions were only clinically diagnosed, and no biopsy was performed to confirm diagnosis. Clinical diagnosis is the gold standard for detection of all OPLs, and only high-risk lesions are

**Table 1** Subject accrual summary.

Sr. No.	Category	Subject accrual	Tobacco habits	Number of specimens for data analysis	Number of spectra
1	HV	20	No	20	218
2	Habit controls (HT)	20	Yes	33	213
3	OPLs	27	Yes	36	234
	-Leukoplakia				
	-TPK				
	-OSMF				

subjected to biopsy, as per clinical guidelines. This study was approved by the Institutional Ethics Committee of both participating institutes, TMC and D. Y. Patil Dental College and Hospital. Subject accrual summary is provided in Table 1. Both HT and OPL groups were characterized by similar tobacco habits in terms of frequency and duration, which on an average ranged between 3 and 4 times/day (frequency) and 8 to 10 years (duration). While tobacco chewing was the most common habit, most of the subjects had mixed tobacco habits (chewing and smoking). Exfoliated cells were collected using Cytobrush (Himedia, India) in normal saline from both buccal mucosae of all subjects. For healthy volunteers, the left and right buccal mucosa samples were treated as single entity, while both left and right exfoliated specimens were treated as different entities in HT and OPL groups, owing to differences in tobacco exposure/presence of OPL conditions, respectively. For data analysis, only sites with tobacco exposure (for HT) and site of premalignant lesion (for OPL) were included. Therefore, exfoliated cell samples from left mucosa, right mucosa, or both mucosae were used depending on the site of tobacco exposure (for HT) and the site for premalignant lesion (for OPL). Therefore, 20 samples from 20 cases of HV, 33 samples from 20 cases of HT, and 36 samples from 27 cases of OPL were used for data analysis. Samples were transported in ice at 4°C and immediately processed upon arrival. Exfoliated cells were harvested after vortexing and centrifugation at 6000 rpm for 1 min. Cell counting was performed using Neubauer's chamber. For cell counting, the harvested cell pellet was suspended in 1 ml of saline and mixed to obtain a uniform suspension. A 10- $\mu$ l representative sample was drawn from this suspension, mixed with 10  $\mu$ l of Trypan blue stain, and pipetted onto the counting chamber. The cells were counted at 10 $\times$  magnification in the white blood cell area using a light microscope. Meanwhile, the remaining 990- $\mu$ l cell suspension was centrifuged to obtain the cell pellet for Raman spectral acquisition.

### 2.2 Raman Spectral Acquisition

The pellet was placed on a CaF<sub>2</sub> window and spectra were recorded using a fiber-optic Raman microprobe (Horiba-Jobin-Yvon, France) as described previously.<sup>36</sup> Briefly, the system consists of a laser (785 nm, Process Instruments) as an excitation source and an HE-785 spectrograph (Horiba-Jobin-Yvon, France) coupled with a charge-coupled device (CCD) (Synapse, Horiba-Jobin-Yvon) as dispersion and

detection elements, respectively. Optical filtering of unwanted noise, including Rayleigh signals, is accomplished through a “superhead,” the other component of the system. Optical fibers were employed to carry the incident light from the excitation source to the sample and also collect the Raman scattered light from the sample and relay it to the detection system. The Raman microprobe was assembled by coupling a 40× microscopic objective (Nikon, Japan) to the superhead. The instrument has no movable parts, and the spectral resolution of the system is  $4\text{ cm}^{-1}$ , as per the manufacturer. The spot size for laser excitation is 3 to 5  $\mu\text{m}$ . Spectral acquisition parameters were as follows: laser power: 40 mW, acquisition time: 15 s, and spectra were averaged over three accumulations. Spectra were acquired from the whole cell pellet and not individual cells. Several areas of the pellet were sampled to acquire 6 to 7 spectra so as to obtain maximum cellular information from the cell pellet. After spectra recording, cells were smeared onto glass slides, fixed in 95% alcohol, and subjected to Papanicolaou staining.

### 2.3 Pap Staining

Pap stain was performed mainly to understand the cellular composition of the sample. Pap employs hematoxylin (nuclear stain), Orange G, and eosin-azure (cytoplasmic stains) and stains cells based on their degree of maturation and keratin content. Post-Raman spectral acquisition, cell pellet was smeared onto the glass slide, fixed in 95% ethanol, and Pap stained using a protocol described elsewhere.<sup>34</sup> The slides were read by a certified pathologist subsequent to mounting. Dysplasia was identified by karyomegaly (nuclear enlargement), increased nuclear:cytoplasmic ratio, cellular and nuclear pleomorphism, nuclear membrane irregularities, and nuclear crowding. Dyskeratosis was defined as abnormal keratinization observed in the lower layers of the epithelium.<sup>37</sup> A final class of pathology diagnosis was designated as per Bethesda classification recommended for cervical PAP reporting.<sup>38</sup>

### 2.4 Spectral Preprocessing and Data Analysis

Raman spectra were corrected for CCD response and spectral contaminations from substrate and fiber signals. To remove interference of the slow-moving background, first derivatives of spectra (Savitzky-Golay method and window size 3) were computed. Interpolated ( $800$  to  $1800\text{ cm}^{-1}$ ), first derivative, and vector normalized spectra were subjected to multivariate unsupervised principal component analysis (PCA) and supervised principal component-linear discriminant analysis (PC-LDA). PCA is a routinely used method for data compression and visualization.<sup>39</sup> PCA was carried out using first 10 factors to cover maximum variance in the data. The final scatter plot is selected after plotting score combinations for several factor combinations and also verifying the factors using loading plots. While PCA aims to identify features that represent variance among complete data, LDA provides data classification based on an optimized criterion, which is aimed for better class separability. LDA can be used in conjunction with PCA (PC-LDA) to increase the efficiency of classification. For PC-LDA, PCA scores obtained using a set of significant PCs with maximum variance among data are used as input data for LDA-based classification. The advantage of doing this is to remove or minimize noise from the data and concentrate on variables important for classification. To avoid overfitting of the data, as a rule of

thumb, only less than half number of components than the spectra in the smallest group should be utilized. Generally, too few factors result in lack of information while more number of factors can increase the noise in the data, which leads to overfitting. Use of around 10 components has been observed to be a balance between excluding important information and including noise in data.<sup>40–43</sup> PC-LDA models were validated by leave-one-out cross validation (LOOCV). For PC-LDA, seven factors were used as a balance between including maximum variance and excluding overfitting, after analyzing first 10 factors in the analysis. The same factors were used for each cross validation. Algorithms for these analyses were implemented in MATLAB<sup>®</sup> (MathWorks Inc., Natick, Massachusetts)-based in-house software.<sup>44</sup>

Average and standard deviation spectra were computed by baseline correction of background-subtracted spectra by fitting a fifth-order polynomial function. These baseline corrected, smoothed (Savitzky-Golay, 3), and normalized spectra were used for spectral comparisons.

## 3 Results and Discussion

Screening and monitoring of oral precancer conditions, such as leukoplakia, OSMF, and TPK, can lead to early detection and better management of oral cancers. Biopsy followed by histopathology may not be suitable for screening and risk assessment of such lesions.<sup>15,16</sup> RS coupled with cytopathology has shown promise in differentiating normal and cancer cells in oral and cervix cancers.<sup>32,34</sup> In this study, the feasibility of REC to detect early precancerous changes was investigated by employing HT, HV, and OPL groups. Data analysis was carried out using PCA and PC-LDA using both spectra- and patient-wise approaches in three- and two-group models. Cytological evaluation was employed to corroborate Raman spectral findings.

### 3.1 Cytological Findings

Commonly encountered precancer conditions in the oral cavity include white lesions termed as leukoplakia/TPK, red lesions called erythroplakia, and progressively fibrosed areas, such as OSMF. On cytological diagnosis, leukoplakia may be benign or premalignant. In benign leukoplakia, yellow, yellow-orange stained cells without nuclei (anucleate squames) may be observed. In premalignant leukoplakias, nuclear abnormalities in well-differentiated squames are usually observed. These nuclear changes may be indicative of mild or moderate dysplasia. In the less common red lesions, i.e., erythroplakia, small cancer cells with minimal or absent keratinization manifest in the smears, indicative of severe dysplasia or carcinoma *in situ*.<sup>37,45</sup>

Cytological analysis of the cell pellet employed for spectral acquisition was carried out after Pap staining and was graded for presence of normal cells, orangeophilic cells, parakeratotic cells, anucleate squames, dyskeratotic, and dysplastic cells by a certified pathologist. The final pathological class was assigned as per the Bethesda system. Representative slides for the three groups, i.e., HV, HT, and OPL, are shown in Figs. 1(a)–1(c). In the HV smears, normal proportions of cells from stratum intermedium, stratum superficial, and few cells from stratum basale were observed. A very small percentage of parakeratotic and orangeophilic cells were detected in some smears. HT smears showed higher numbers of orangeophilic and occasional parakeratotic cells and anucleate squames. In OPL smears, very high frequency of orangeophilic cells, numerous parakeratotic

and anucleate squames were observed for most cases. In some cases, mild changes in the nuclear features, nuclear:cytoplasmic ratios, and dyskeratosis-like features were also observed. These subtle changes may be indicative of early dysplasia. Higher parakeratotic cells and orthokeratinized cells indicate higher mitotic index in the lower layers of the epithelium.<sup>46,47</sup> Thus, higher parakeratinization and orthokeratinization observed in OPL may be associated with epithelial dysplasia and hyperplasia.

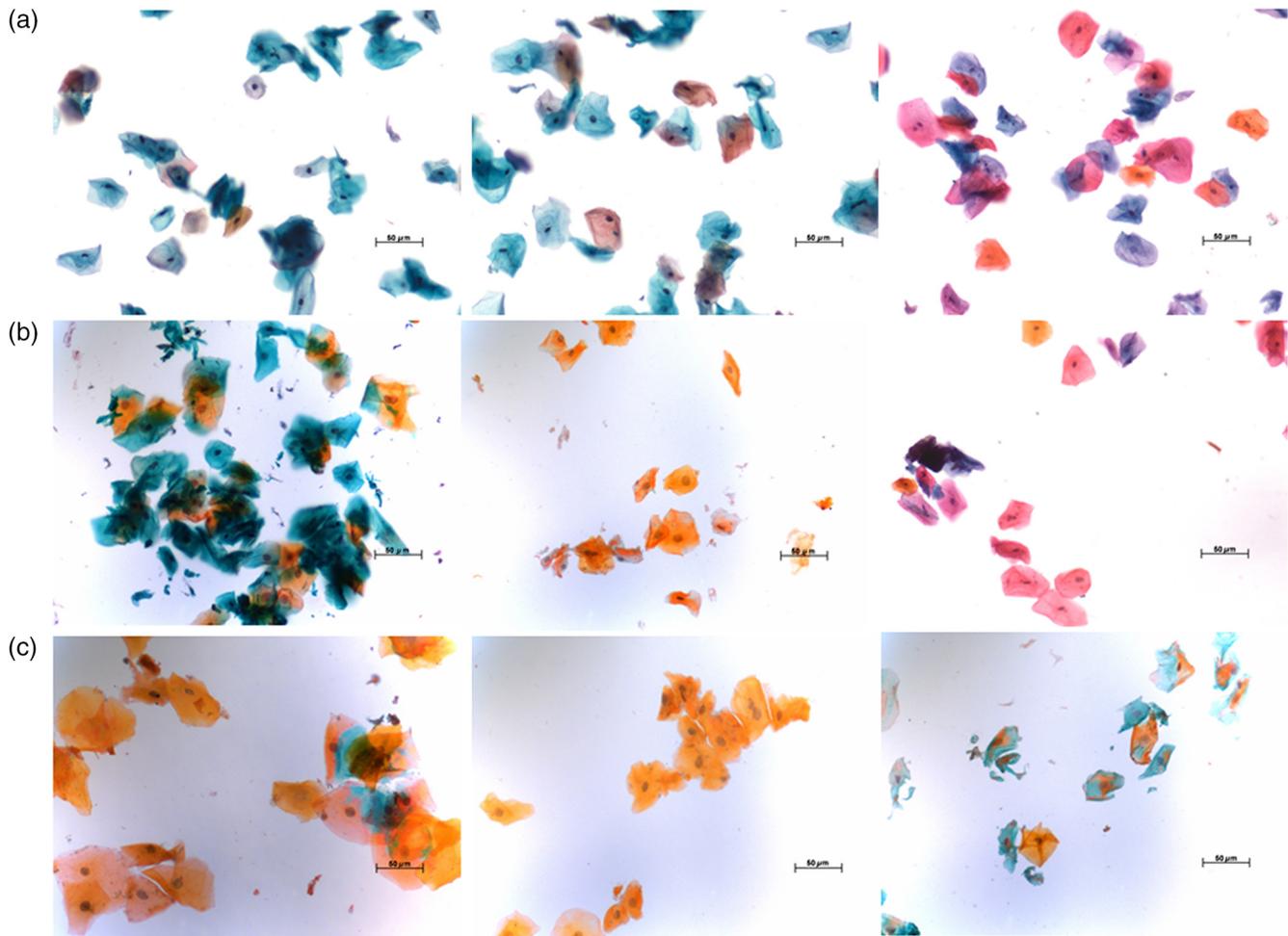
### 3.2 Spectral Features

The average and standard deviation spectra for the HV, HT, and OPL groups are shown in Figs. 2(a)–2(c). Prominent spectral features include bands at 830, 857, 1008, 1090, 1216, and in the region of 1260 to 1340, 1452, and 1660  $\text{cm}^{-1}$ . Thus, contributions from phenylalanine, DNA-phosphate backbone-related features, amide III,  $\text{CH}_2$  twisting in proteins and lipids,  $\text{CH}_2$  bending in protein and lipids, and amide I features from proteins were observed. Intensity-related variations at most of these peaks and shifts in the amide III region were observed across the groups. Broad amide I was observed with increase in the pathological status, from healthy non-tobacco to oral premalignant conditions. The broad features between 1260 and

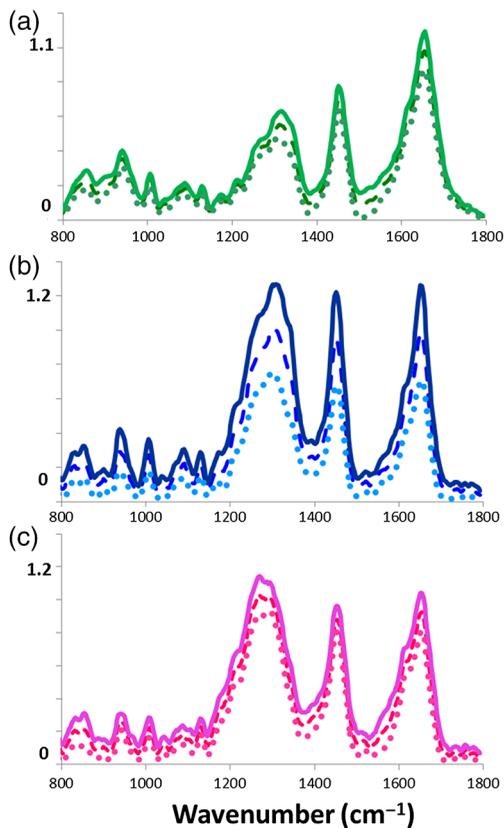
1340  $\text{cm}^{-1}$  contain spectral features corresponding mainly to amide III;  $\text{CH}_3$ ,  $\text{CH}_2$ ,  $\text{CH}_3\text{CH}_2$  wagging/bending/twisting modes in proteins/DNA bases/lipids; ring breathing and other vibrational modes of purine bases adenine, guanine in DNA. These changes could be ascribed to changes in the secondary structure of proteins, cellular changes, and hypercellularity due to inflammation and also due to the fact that pathological condition cells have large amount of surface proteins, receptor proteins, enzymes, and antigens, which may have led to broadening of the amide I and overall protein-dominant features. Assignments were based on available literature.<sup>48,49</sup>

### 3.3 Multivariate Analysis

Multivariate data analysis was carried using unsupervised PCA and supervised PC-LDA, followed by LOOCV. Data analysis was carried out in two- and three-group models. In the first step, three-group model analysis using healthy (HV), healthy with tobacco (HT), and premalignant (OPL) groups was carried out. The HV group was used as a control group. In the second step, a more clinically relevant two-group model using only HT and OPL groups was explored to investigate classification of premalignant conditions from tobacco-exposed mucosa. Data



**Fig. 1** Representative images of Pap-stained cytological smears (200 $\times$ ) from (a) healthy volunteers showing cells from superficial and intermediate layers with occasional presence of orangeophilic cells, (b) habit controls showing increased number of orangeophilic cells and occasional parakeratotic cells and anucleate squames, and (c) OPLs showing very high orangeophilia, parakeratotic, anucleate cells, cells with minor dyskeratosis, and mild increase in nuclear:cytoplasmic ratio.



**Fig. 2** Mean and standard deviation spectra of exfoliated cells from (a) healthy volunteers (HV), (b) habit controls (HT), and (c) premalignant (OPL) subjects, for spectral comparisons (solid line: mean plus standard deviation; broken line: mean, and dotted line: mean minus standard deviation).

analysis was carried out using both spectra- and patient-wise approaches on 218 HV spectra (20 samples), 213 HT spectra (33 samples), and 234 OPL spectra (27 samples).

### 1. Exploring three-group model for classifying healthy, healthy with tobacco, and oral premalignant groups.

Data analysis was first carried out in three-group model using the HV, HT, and OPL groups to explore differences between healthy, tobacco-exposed mucosa, and premalignant lesions in buccal mucosa. Healthy group (HV) was used as a pure healthy control—with no tobacco exposure in buccal mucosa. PCA and PC-LDA followed by LOOCV were carried using the spectra- and patient-wise approaches.

#### a. Spectra-wise analysis.

In the first step, spectra-wise approach was adopted. All spectra were used as input for PCA to examine trends in the data. The PCA variance plot is shown in Fig. 3(a). Scores of factor 1 and 3 were explored for classification. Factor loadings for factor 1 and 3 are shown in Figs. 3(b) and 3(c). The scatter plot in Fig. 3(d) indicates three overlapping clusters for HV, HT, and OPL. As PCA is not a classification tool but is used for data compression and visualization to indicate trends in the data, PC-LDA was employed to explore classification among the groups. Seven factors [Fig. 4(a)] were used for the analysis, which accounted for  $\sim 75\%$  variance. Scores of factor 2 and 3 were employed for obtaining scatter plot, as shown in Fig. 4(b). As shown in the scatter plot, better classification was observed between the three groups. LOOCV was used to validate the

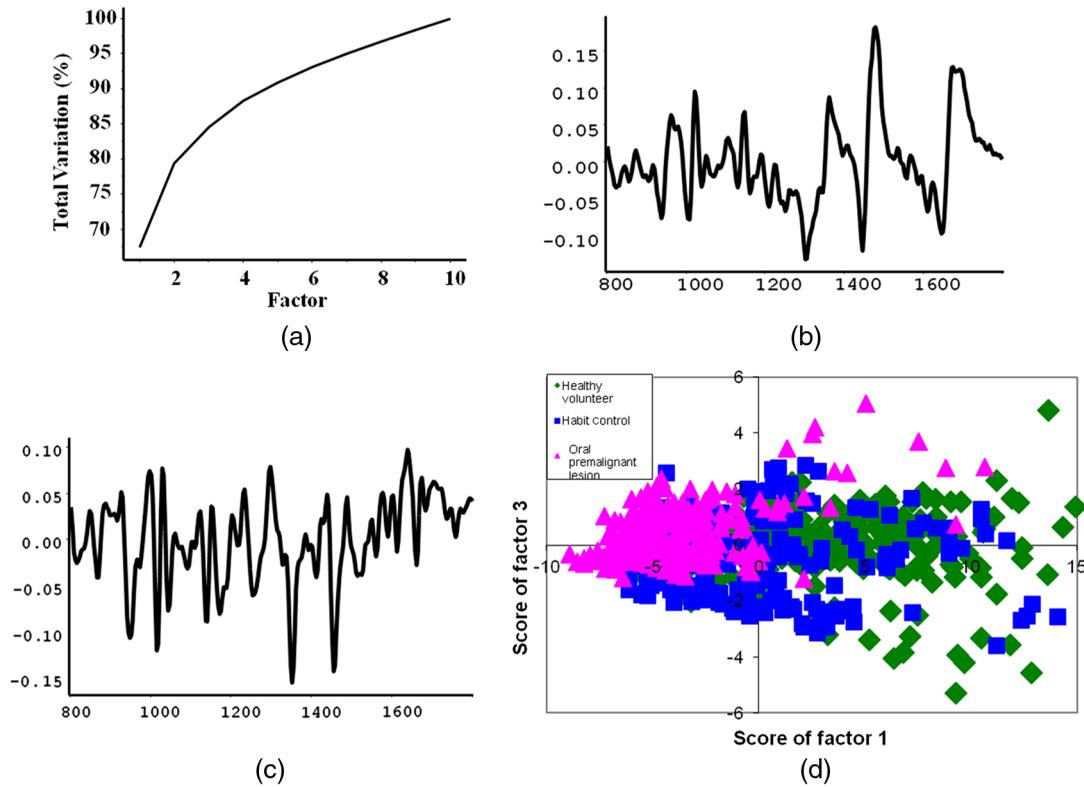
PC-LDA findings. As shown in the LOOCV confusion matrix (Table 2), 174/218 HV spectra were correctly classified while 35 and 9 misclassified with HT and OPL, respectively. In case of HT, 137/213 spectra were correctly classified. Misclassifications of HT were observed with both HV and OPL groups. For OPL, 179/234 spectra were correctly predicted while major misclassifications<sup>49</sup> were with HT. Thus, OPL could be identified with  $\sim 77\%$  sensitivity using the spectra-wise approach.

#### b. Patient-wise analysis.

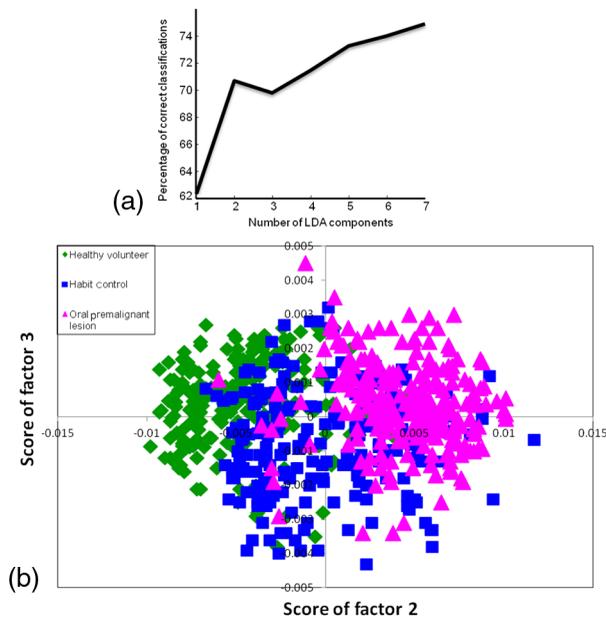
The patient-wise approach, wherein average spectrum of each sample was used for data analysis, was also explored. Average spectra were calculated to yield a true representative of the sample and also to circumvent the intrasample heterogeneity. These average spectra were then subjected to multivariate analysis PCA and PC-LDA. PCA was carried out using 10 factors [Fig. 5(a)]. Scores of factor 1 and 3 were used to obtain PCA scatter plot. Factor loadings are shown in Figs. 5(b) and 5(c). As shown in Fig. 5(d), slightly distinct clusters for HV, HT, and OPL were observed. PC-LDA was subsequently carried out to build standard models, using seven factors [Fig. 6(a)]. The scatter plot shown in Fig. 6(b) indicates almost exclusive clusters for the three groups. LOOCV confusion matrix findings demonstrate that 14/20 HV, 20/33 HT, and 25/36 OPL cases were correctly predicted (Table 3). For HV, six cases were misclassified with HT; for HT, seven and six were misclassified with HV and OPL, respectively. While for OPL, two and nine cases were misclassified with HV and HT, respectively, yielding a sensitivity of  $\sim 70\%$ .

For HV, major misclassifications were observed with the HT group. In case of HV, normal proportions of intermediate and superficial cells along with minor orangeophilia were present in most cases. HT smears had similar cytological profiles but higher orangeophilia and parakeratotic cells. As explained in our previous studies,<sup>34</sup> tobacco exposure and cellular response to this exposure may not be uniform in the entire buccal mucosa area, and therefore, the presence of normal cells in these smears may lead to overlap between HV and HT groups. HT spectra also misclassified with OPL spectra and vice versa. Apart from normal cells, HT smears also displayed parakeratotic cells and some cells with increased nuclear size but maintained N:C ratios indicating mild atypia due to chronic irritation. These features somewhat overlapped with the OPL group. Further, premalignant lesions may also be heterogenous at microscopic level, thus cytological similarities between clinically normal but tobacco-exposed areas in OPL lesions and HT cannot be ruled out. Further, OPLs included in this study were only clinically proven leukoplakia, OSMF, and TPK cases having long-term tobacco habits, and they were not histopathologically confirmed dysplastic lesions. These samples may have been benign keratotic lesions, which also demonstrate hyperkeratosis in the cytological smears. Thus, the overlap between HT and OPL can be understood.

To understand the association between the cytological characteristics and Raman classification, the cytological evaluation, Pap, and Raman summary have been tabulated for each sample (data not shown). Concordance was observed between the cytology and spectral findings for most cases. There were some instances of over- and underdiagnoses with both approaches, but the combined approaches yielded a better classification. In case of OPL (and HT), overdiagnosis through Raman with respect to cytology was observed in some cases. However,



**Fig. 3** Three-group model PCA for exploring differences between exfoliated cells from healthy volunteers, habit controls, and premalignant subjects using spectra-wise approach: (a) variance plot, (b) loadings of factor 2, (c) loadings of factor 3, and (d) scatter plot.



**Fig. 4** Three-group model PC-LDA for exploring differences between exfoliated cells from healthy volunteers, habit controls, and premalignant subjects using spectra-wise approach: (a) scree plot and (b) scatter plot.

the prediction as OPL group in Raman may be indicative of some early premalignancy-related molecular changes in the cells, which were not detected by Pap owing to probing of only morphological changes (mostly late-onset) and/or lower

**Table 2** PC-LDA for exploring differences between exfoliated cells from healthy volunteers, habit controls, and premalignant subjects using spectra-wise approach—confusion matrix of LOOCV results (In the confusion matrix, the diagonal elements represent true positive predictions while the exdiagonal elements represent false positive and false negative predictions).

	Healthy volunteer	Habit control	Premalignant	Total
Healthy volunteer	174	35	9	218
Habit control	38	137	38	213
Premalignant	6	49	179	234

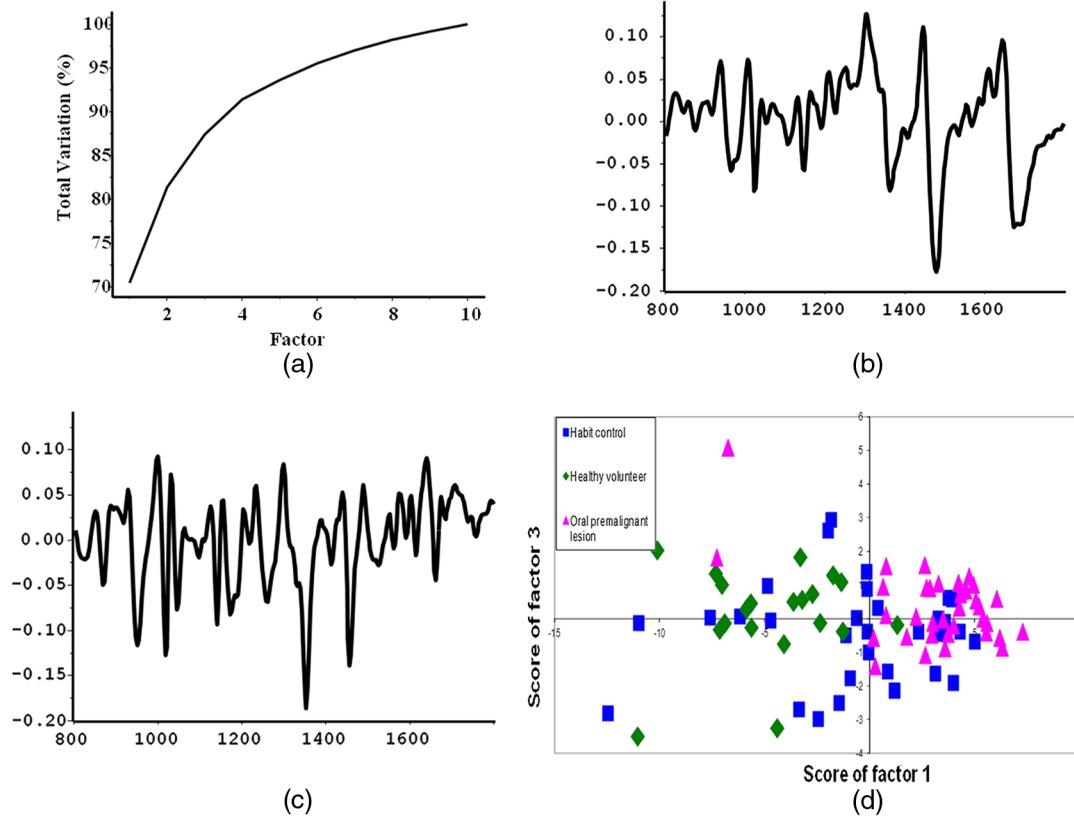
sensitivity. Thus, this combinatorial REC approach may be a better tool for identifying early premalignant changes with high confidence.

2. Exploring the two-group model of healthy tobacco and premalignant groups for clinical applications.

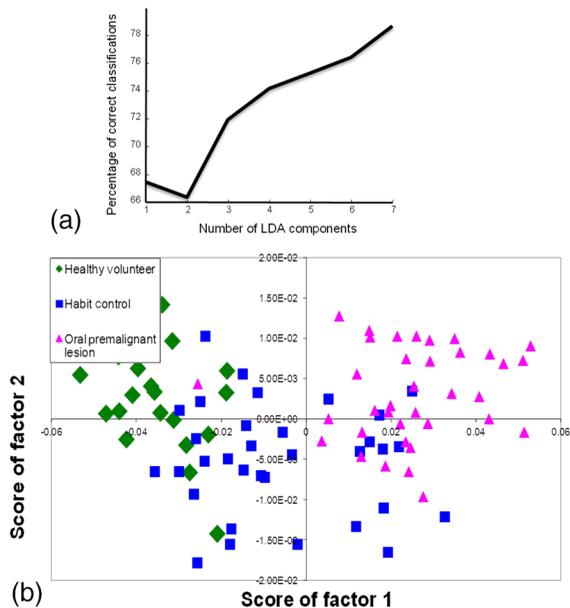
In the next step, classification between healthy tobacco users and premalignant subjects was investigated. Such two-group model is more clinically relevant for screening of high-risk populations using REC. Both, spectra- and patient-wise approaches were employed for the two-group model.

a. Spectra-wise analysis.

PCA was carried out on 213 HT and 234 OPL spectra from 33 and 36 exfoliated samples, respectively, using 10 factors [Fig. 7(a)]. Scatter plot was plotted for scores of factor 1 and



**Fig. 5** Three-group model PCA for exploring differences between exfoliated cells from healthy volunteers, habit controls, and premalignant subjects using patient-wise approach: (a) variance plot, (b) loadings of factor 1, (c) loadings of factor 3, and (d) scatter plot.



**Fig. 6** Three-group model PC-LDA for exploring differences between exfoliated cells from healthy volunteers, habit controls, and premalignant subjects using patient-wise approach: (a) scree plot and (b) scatter plot.

3. Factor loadings for factors 1 and 3 are shown in Figs. 7(b) and 7(c). Two overlapping clusters were observed in Fig. 7(d). PC-LDA followed by LOOCV was carried out in the next step. Four factors accounting for ~79% variance in the data

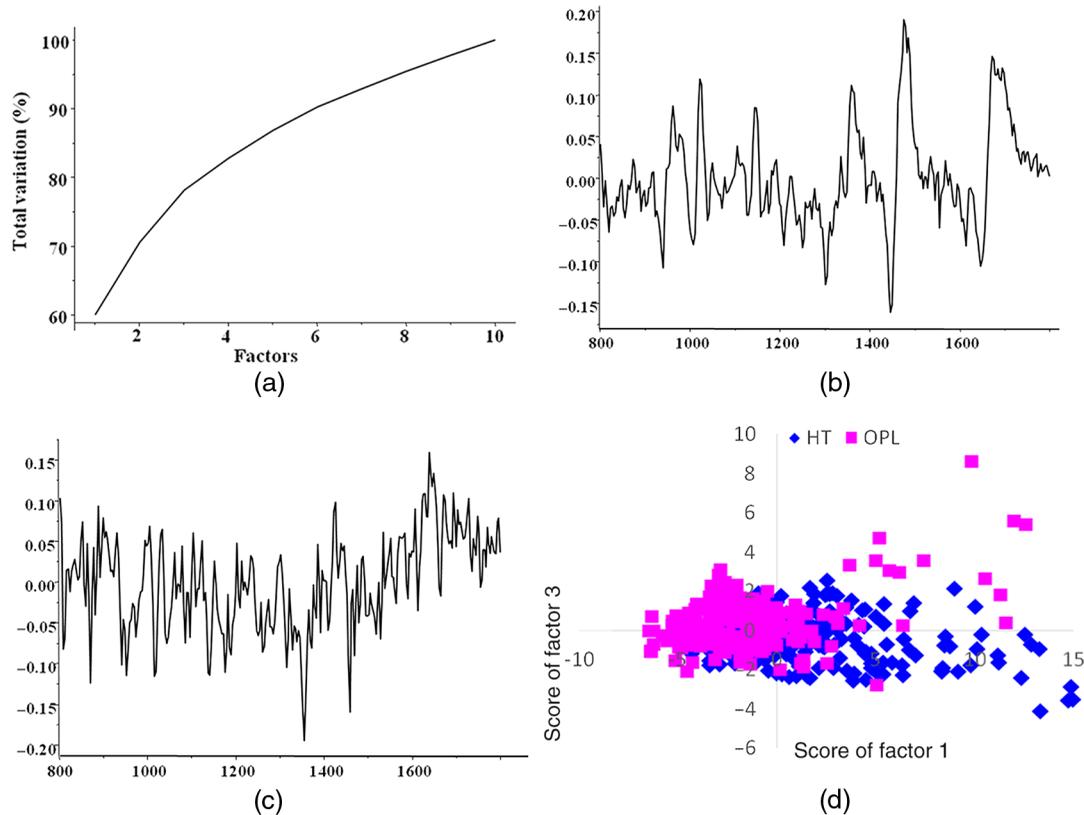
**Table 3** PC-LDA for exploring differences between exfoliated cells from healthy volunteers, habit controls, and premalignant subjects using patient-wise approach—confusion matrix of LOOCV (in the confusion matrix, the diagonal elements represent true positive predictions while the exdiagonal elements represent false positive and false negative predictions).

	Healthy volunteer	Habit control	Premalignant	Total
Healthy volunteer	14	6	0	20
Habit control	7	20	6	33
Premalignant	2	9	25	36

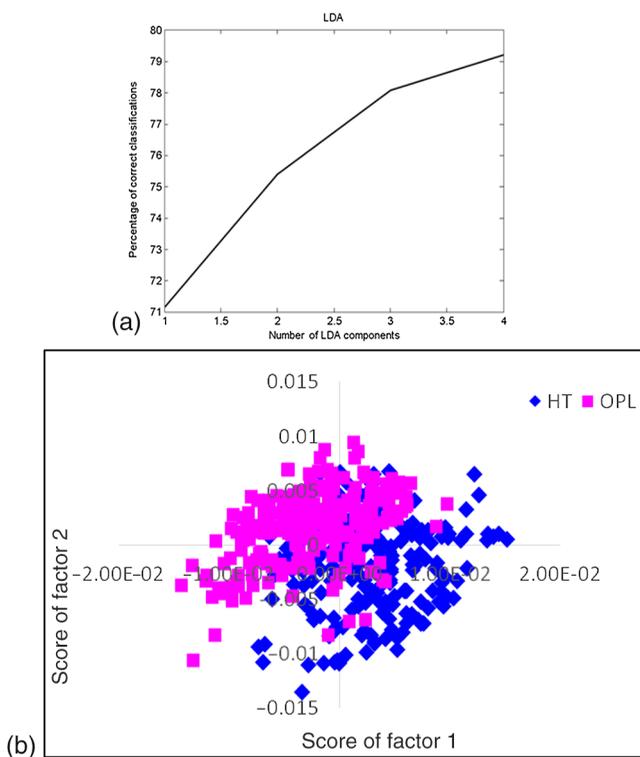
were employed for PC-LDA [Fig. 8(a)]. Scatter plot for scores of factor 1 and 2 is shown in Fig. 8(b). Two clusters with overlap in healthy tobacco and premalignant groups were observed. In the LOOCV confusion matrix (Table 4), 153 HT spectra were correctly classified while 60 misclassified with OPL. Two hundred OPL spectra were correctly classified while 34 misclassified with HT. Thus, sensitivity for OPL detection was found to be 86%.

b. Patient-wise analysis.

The two-group model analysis was also carried out using patient-wise approach, where all spectra from a sample are averaged to obtain a representative spectrum. This analysis was carried out using 33 HT and 36 OPL samples. PCA was carried out using 10 factors [Fig. 9(a)]. The factor loadings for



**Fig. 7** Two-group model PCA for exploring differences between exfoliated cells from habit controls and premalignant subjects using spectra-wise approach: (a) variance plot, (b) loadings of factor 1, (c) loadings of factor 3, and (d) scatter plot.



**Fig. 8** Two-group model PC-LDA for exploring differences between exfoliated cells from habit controls and premalignant subjects using spectra-wise approach: (a) scree plot and (b) scatter plot.

factor 1 and 3 are presented in Figs. 9(b) and 9(c). Scores of these factors were used to plot the scatter plots, as shown in Fig. 9(d). Two clusters with some overlap were observed. PC-LDA was carried out in the next step using three factors accounting for ~84 variance [Fig. 10(a)]. Scatter plot was plotted using scores of factor 1 and 2, as shown in Fig. 10(b). Two clusters with minimal overlap were observed. As per the LOOCV confusion matrix (Table 5), 22/33 HT were correctly classified while 30/36 OPL were correctly classified. Thus, OPL could be differentiated from HT with 83% sensitivity.

A higher sensitivity or detection of OPL conditions was observed in the two-group model analysis. Such a model may be more useful for prospective screening of high-risk populations (long-term tobacco users with or without mucosal abnormalities) in field settings. There was some overlap observed in the HT and OPL groups. As mentioned previously, the cytological similarities between the HT and OPL groups and/or the presence of early premalignant changes in tobacco-exposed mucosa of HT could be a reason for the misclassifications. Use of biopsy-confirmed cases with high degree of dysplasia may help in the building of more robust standard models. In prospective studies, different clinical models for premalignant and different tobacco habit groups, such as tobacco chewers, areca nut chewers, and tobacco smokers, will be developed for better detection of premalignant changes in oral mucosa.

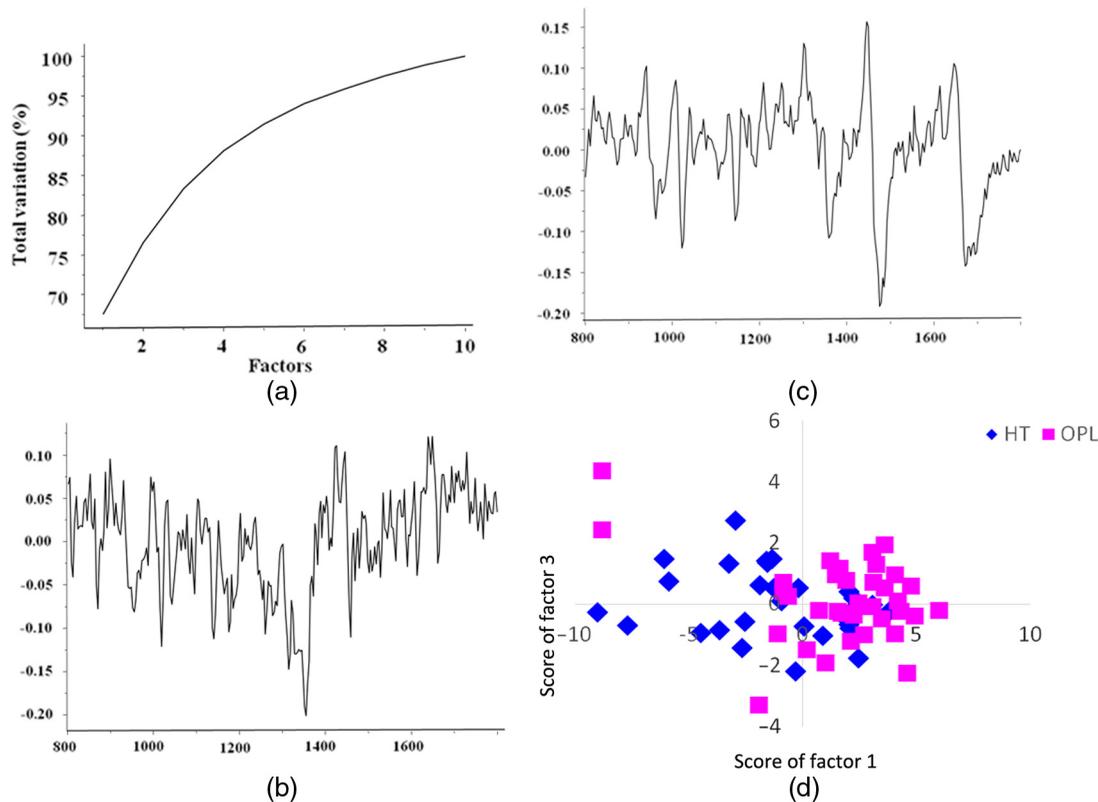
It is known that tissue biopsy is associated with lower patient compliance (9%), but brush biopsy has demonstrated nearly 100% compliance.<sup>50</sup> Studies on dysplasia/neoplasia surveillance in oral lichen planus malignant transformation led to the diagnosis in intraepithelial and microinvasive phases, leading to

**Table 4** PC-LDA for exploring differences between exfoliated cells habit controls and premalignant subjects using spectra-wise approach—confusion matrix of LOOCV results (in the confusion matrix, the diagonal elements represent true positive predictions while the exdiagonal elements represent false positive and false negative predictions).

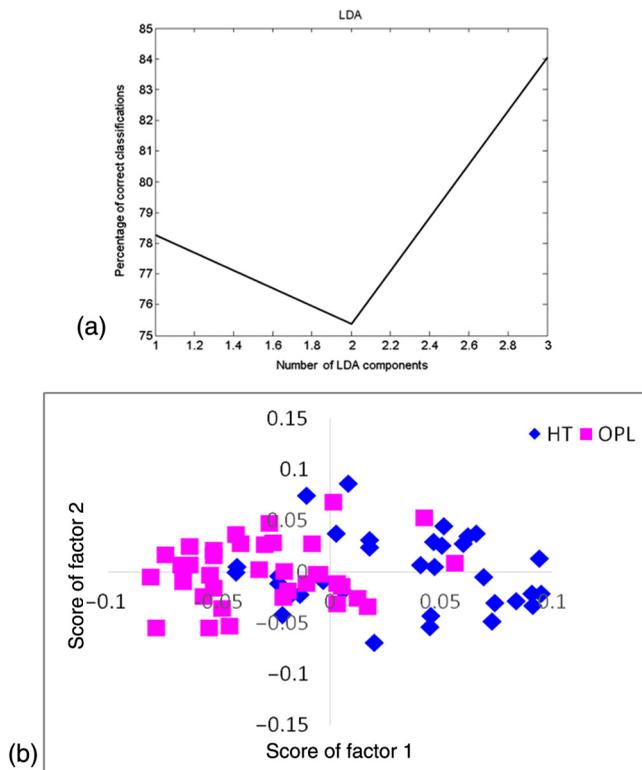
	Habit control	Premalignant	Total
Habit control	153	60	213
Premalignant	34	200	234

substantially high five-year survival rates.<sup>51</sup> Therefore, effective screening and risk assessment of oral premalignant conditions may also lead to early diagnosis of oral cancers and improve overall survival rates. Oral exfoliative cytology using Pap staining was initially explored for oral cancer screening, but it could not achieve expected results possibly due to inadequate cellular representation and subjective assessment. Although the reliability of oral exfoliative cytology was initially questionable, advent of newer modifications for sample collection and automated image analysis systems have renewed interest in the field by improving sensitivity of the method.<sup>52,53</sup> However, these tests rely only on morphological changes that may not be apparent in the early stages of carcinogenesis. Alternatively, optical spectroscopic approaches, such as RS that can yield global biochemical fingerprint of the sample, can be employed. RS has shown potential in the diagnosis of various diseases, including cancers,<sup>28</sup> and *in vivo* RS has even demonstrated classification of

normal, premalignant, and tumor conditions.<sup>31</sup> In this study, REC was used as a less-invasive and a more practical combination modality for assessing morphological as well as biochemical changes in the exfoliated cells obtained from healthy, healthy tobacco users, and subjects with premalignant lesions. Previous studies have shown the utility of RS in identifying single-cell type in mixed-cell populations,<sup>54</sup> differentiating between normal and cervical intraepithelial neoplasia cytology using ThinPrep approach,<sup>55</sup> and also in discriminating normal, dysplastic, and oral SCC cell lines.<sup>56</sup> Therefore, RS may also have the potential to identify minor biochemical changes and smaller number of dysplastic cells observed in early precancerous lesions. Spectral comparisons indicate intensity-related differences, minor shifts, and broadening in spectral features, mainly in the amide III,  $\delta\text{CH}_2$  region, and amide I. Higher DNA and protein features, including broadening of amide I, have been ascribed to cellular proliferation or hyperplastic/dysplastic cellular state.<sup>56,57</sup> Dysplastic cells also show higher DNA, protein, and tryptophan features as compared to normal cells.<sup>58</sup> The DNA and protein-dominant features in habit control and premalignant groups in this study may be a putative outcome of tobacco exposure-induced hypercellularity.<sup>59–61</sup> Thus, the changes in the spectra are consistent with alterations in the protein and DNA content accompanying changes in cellular physiology during proliferation. Another tissue study has also demonstrated higher contribution of proteins and DNA in dysplastic and carcinoma biopsy samples, as compared to normal mucosa.<sup>62</sup> These spectral features also show similarities with the *in vivo* findings wherein healthy with and without tobacco groups demonstrated high-lipid features, but healthy tobacco



**Fig. 9** Two-group model PCA for exploring differences between exfoliated cells from habit controls and premalignant subjects using patient-wise approach: (a) variance plot, (b) loadings of factor 1, (c) loadings of factor 3, and (d) scatter plot.



**Fig. 10** Two-group model PC-LDA for exploring differences between exfoliated cells from habit controls and premalignant subjects using patient-wise approach: (a) scree plot and (b) scatter plot.

**Table 5** PC-LDA for exploring differences between exfoliated cells habit controls and premalignant subjects using patient-wise approach—confusion matrix of LOOCV results (in the confusion matrix, the diagonal elements represent true positive predictions while the exdiagonal elements represent false positive and false negative predictions).

	Habit control	Premalignant	Total
Habit control	22	11	33
Premalignant	6	30	36

also showed shifts in amide III and broadening of amide I with respect to healthy without tobacco. The premalignant spectra demonstrated similar features to tobacco users, such as broadening of amide III,  $\delta\text{CH}_2$ , and amide I. These spectral characteristics were attributed to changes in secondary structures of the proteins (shifts in amide III, amide I, and broadening), inflammatory changes, and hypercellularity due to tobacco exposure.<sup>31</sup> Further, the spectral findings are also in line with the cytological findings that also indicate higher cell proliferation, as evidenced by increased keratinization in healthy tobacco and more in premalignant groups.<sup>46,47</sup> In this study, the identification of premalignant conditions in both spectra- and patient-wise analyses was observed. Higher sensitivity was observed in the clinically relevant two-group model  $\sim 77\%$  (spectra-wise) and  $\sim 70\%$  (patient-wise) in the three-group model versus 86% (spectra-wise) and 83% (patient-wise) in the two-group model, respectively. Use of biopsy-confirmed dysplastic conditions may help in the development of robust standard models and lead to better

identification of precancerous changes. Further, misclassification of HT with biopsy-confirmed OPL may also give additional insights about the clinically normal HT mucosa, which may be susceptible to premalignant or malignant lesion developments. The classification efficiency and utility of the method may also be improved by using better collection devices. Collection devices, such as Cytobrush, are suitable for collection of mostly superficial and intermediate layers. Additionally, sampling of keratotic lesions, such as leukoplakia and TPK, gives an overall lower cellular yield and more so from the basal layers. Therefore, the early changes occurring in the basal layers may not be detected using Cytobrush; even in this study, basal and parabasal cells were rarely encountered. Thus, early cytological changes may be better detected by using devices that enable transepithelial sampling. Therefore, with improvements in sample collection and use of biopsy-confirmed premalignant lesions, the efficacy of the method can be improved.

## 4 Conclusions

Oral cancer is associated with poor survival rates, attributed mainly to delays in diagnosis. Premalignant lesions, such as leukoplakia, TPK, erythroplakia, and OSMF, usually precede invasive OSCC. Detection of high-risk premalignant conditions can help in better clinical outcomes. Clinical monitoring of such lesions, combined with less-invasive methods, needs to be implemented to identify high-risk lesions. In this study, potential of REC in identifying premalignant conditions from healthy cells was explored using healthy without tobacco (HV), healthy with tobacco habits (HT), and different premalignant subjects (OPL)—leukoplakia, OSMF, and TPK. With increase in pathological status, changes mainly in the amide III, DNA bases, and broadening of amide I were observed in the spectral features. Data analysis was carried out by PCA and PC-LDA using spectra- and patient-wise approaches, in both two- and three-group models. OPLs could be identified with at least 70% sensitivity in the three-group model and at least 83% sensitivity in the two-group model. By incorporating only biopsy-confirmed lesions and improving sampling devices, the efficiency of REC in detecting high-risk lesions and monitoring for malignant transformation may be improved. The combination of REC and cytomorphology yields better classification. These preliminary studies indicate promise of this approach in serving as a useful adjunct for screening, monitoring, and risk assessment of oral premalignant conditions, especially using the two-group model for evaluating cancer risk in high-risk populations (subjects with extensive, long-term tobacco habits). REC may also be advantageous for lesions that are clinically not suspicious for malignancy but nevertheless need surveillance. After an increasingly dysplastic REC result, a mucosal biopsy may be advised for confirmatory purposes, and consequently, the biopsy findings can be used to make therapeutic decisions. After validation on large and diverse cohorts and stringent longitudinal validation studies, such a less-invasive and rapid approach can serve as a useful adjunct in screening and monitoring of oral precancerous lesions.

## Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

## Acknowledgments

The Raman spectrometer employed in the study was procured from DBT project BT/PR11282/MED/32/83/2008 (Development of *in vivo* laser Raman spectroscopy methods for diagnosis of oral precancerous and cancerous conditions), Department of Biotechnology, Government of India.

## References

- W. J. Blot et al., "Smoking and drinking in relation to oral and pharyngeal cancer," *Cancer Res.* **48**(11), 3282–3287 (1988).
- N. Johnson, "Tobacco use and oral cancer: a global perspective," *J. Dent. Educ.* **65**(4), 328–339 (2001).
- S. Warnakulasuriya, C. Trivedy, and T. J. Peters, "Areca nut use: an independent risk factor for oral cancer. The health problem is under-recognized," *BMJ* **324**(7341), 799 (2002).
- Y. C. Ko et al., "Betel quid chewing, cigarette smoking and alcohol consumption related to oral cancer in Taiwan," *J. Oral Pathol. Med.* **24**(10), 450–453 (1995).
- S. Silverman, "Demographics and occurrence of oral and pharyngeal cancers: the outcomes, the trends, the challenge," *J. Am. Dent. Assoc.* **132**, 7S–11S (2001).
- B. W. Neville and T. A. Day, "Oral cancer and precancerous lesions," *CA Cancer J. Clin.* **52**(4), 195–215 (2002).
- H. Williams, "Molecular pathogenesis of oral squamous carcinoma," *Mol. Pathol.* **53**(4), 165–172 (2000).
- I. Van der Waal, "Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management," *Oral Oncol.* **45**(4), 317–323 (2009).
- I. van der Waal, "Potentially malignant disorders of the oral and oropharyngeal mucosa; present concepts of management," *Oral Oncol.* **46**(6), 423–425 (2010).
- S. Warnakulasuriya, N. Johnson, and I. Van der Waal, "Nomenclature and classification of potentially malignant disorders of the oral mucosa," *J. Oral Pathol. Med.* **36**(10), 575–580 (2007).
- T. Axéll et al., "Oral white lesions with special reference to precancerous and tobacco-related lesions: conclusions of an international symposium held in Uppsala, Sweden, May 18–21 1994," *J. Oral Pathol. Med.* **25**(2), 49–54 (1996).
- H. M. Mehanna et al., "Treatment and follow-up of oral dysplasia—a systematic review and meta-analysis," *Head Neck* **31**(12), 1600–1609 (2009).
- P. G. Arduino et al., "Outcome of oral dysplasia: a retrospective hospital-based study of 207 patients with a long follow-up," *J. Oral Pathol. Med.* **38**(6), 540–544 (2009).
- C. F. Poh et al., "Biopsy and histopathologic diagnosis of oral premalignant and malignant lesions," *J. Can. Dent. Assoc.* **74**(3), 283–288 (2008).
- A. Karabulut et al., "Observer variability in the histologic assessment of oral premalignant lesions," *J. Oral Pathol. Med.* **24**(5), 198–200 (1995).
- G. Pitiyage et al., "Molecular markers in oral epithelial dysplasia: review," *J. Oral Pathol. Med.* **38**(10), 737–752 (2009).
- J. B. Epstein et al., "Screening for and diagnosis of oral premalignant lesions and oropharyngeal squamous cell carcinoma role of primary care physicians," *Can. Fam. Physician* **54**(6), 870–875 (2008).
- L. L. Patton, J. B. Epstein, and A. R. Kerr, "Adjunctive techniques for oral cancer examination and lesion diagnosis: a systematic review of the literature," *J. Am. Dent. Assoc.* **139**(7), 896–905 (2008).
- R. Mehrotra et al., "Oral cytology revisited," *J. Oral Pathol. Med.* **38**(2), 161–166 (2009).
- J. J. Sciubba and U. C. O. S. Group, "Improving detection of precancerous and cancerous oral lesions: computer-assisted analysis of the oral brush biopsy," *J. Am. Dent. Assoc.* **130**(10), 1445–1457 (1999).
- N. K. Proia et al., "Smoking and smokeless tobacco-associated human buccal cell mutations and their association with oral cancer—a review," *Cancer Epidemiol. Biomarkers Prev.* **15**(6), 1061–1077 (2006).
- R. Mehrotra, "The role of cytology in oral lesions: a review of recent improvements," *Diagn. Cytopathol.* **40**(1), 73–83 (2012).
- T. W. Remmerbach et al., "Cytologic and DNA-cytometric early diagnosis of oral cancer," *Anal. Cell. Pathol.* **22**(4), 211–221 (2001).
- D. Maraki, J. Becker, and A. Boecking, "Cytologic and DNA-cytometric very early diagnosis of oral cancer," *J. Oral Pathol. Med.* **33**(7), 398–404 (2004).
- T. W. Remmerbach et al., "Toward a multimodal cell analysis of brush biopsies for the early detection of oral cancer," *Cancer Cytopathol.* **117**(3), 228–235 (2009).
- T. W. Remmerbach et al., "Earliest detection of oral cancer using non-invasive brush biopsy including DNA-image-cytometry: report on four cases," *Anal. Cell. Pathol.* **25**(4), 159–166 (2003).
- T. W. Remmerbach et al., "Diagnostic value of nucleolar organizer regions (AgNORs) in brush biopsies of suspicious lesions of the oral cavity," *Anal. Cell. Pathol.* **25**(3), 139–146 (2003).
- D. I. Ellis et al., "Illuminating disease and enlightening biomedicine: Raman spectroscopy as a diagnostic tool," *Analyst* **138**(14), 3871–3884 (2013).
- E. Hanlon et al., "Prospects for *in vivo* Raman spectroscopy," *Phys. Med. Biol.* **45**(2), R1 (2000).
- M. B. Fenn et al., "Raman spectroscopy for clinical oncology," *Adv. Opt. Technol.* **2011**, 213783 (2011).
- S. Singh et al., "In vivo Raman spectroscopic identification of premalignant lesions in oral buccal mucosa," *J. Biomed. Opt.* **17**(10), 105002 (2012).
- S. Rubina et al., "Raman spectroscopic study on classification of cervical cell specimens," *Vib. Spectrosc.* **68**, 115–121 (2013).
- E. Vargis et al., "Near-infrared Raman microspectroscopy detects high-risk human papillomaviruses," *Transl. Oncol.* **5**(3), 172–179 (2012).
- A. Sahu et al., "Raman spectroscopy and cytopathology of oral exfoliated cells for oral cancer diagnosis," *Anal. Methods* **7**, 7548–7559 (2015).
- A. Sahu et al., "Raman spectroscopy and oral exfoliative cytology," *Proc. SPIE* **8926**, 89262N (2014).
- A. Sahu et al., "Raman spectroscopy of oral buccal mucosa: a study on age-related physiological changes and tobacco-related pathological changes," *Technol. Cancer Res. Treat.* **11**(6), 529–541 (2012).
- L. G. Koss and G. R. Durfee, "Diagnostic cytology and its histopathologic bases," *Am. J. Med. Sci.* **242**(6), 790 (1961).
- D. Soloman, "The 1988 Bethesda system for reporting cervical/vaginal cytologic diagnoses: developed and approved at the National Cancer Institute Workshop in Bethesda, MD, December 12–13, 1988," *Diagn. Cytopathol.* **5**(3), 331–334 (1989).
- K. Varmuza and P. Filzmoser, *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press, Boca Raton, Florida (2009).
- P. Crow et al., "The use of Raman spectroscopy to differentiate between different prostatic adenocarcinoma cell lines," *Br. J. Cancer* **92**(12), 2166–2170 (2005).
- T. Harvey et al., "Factors influencing the discrimination and classification of prostate cancer cell lines by FTIR microspectroscopy," *Analyst* **134**(6), 1083–1091 (2009).
- T. C. Bakker Schut et al., "Real-time tissue characterization on the basis of *in vivo* Raman spectra," *J. Raman Spectrosc.* **33**(7), 580–585 (2002).
- J. G. Kelly et al., "Biospectroscopy to metabolically profile biomolecular structure: a multistage approach linking computational analysis with biomarkers," *J. Proteome Res.* **10**(4), 1437–1448 (2011).
- A. Ghanate et al., "Comparative evaluation of spectroscopic models using different multivariate statistical tools in a multicancer scenario," *J. Biomed. Opt.* **16**(2), 025003 (2011).
- G. MacDonald, "Classification and histopathological diagnosis of epithelial dysplasia and minimally invasive cancer," in *Satellite Symp. on Epithelial Dysplasia and Borderline Cancer of the Head and Neck: Controversies and Future Directions. Joint BSOMP, BSOM, BAHNO Meeting*, Oxford (2003).
- B. Sahiar, D. Daftary, and F. S. Mehta, "Cytological and histological keratinization studies in leukoplakias of the mouth," *J. Oral Pathol. Med.* **4**(1), 19–26 (1975).
- I. Gujral et al., "Cytological study of leukoplakic lesions in oral cavity," *J. Mahatma Gandhi Inst. Med. Sci.* **15**(2), 49–54 (2010).
- F. S. Parker, *Applications of Infrared, Raman, and Resonance Raman Spectroscopy in Biochemistry*, Springer, Heidelberg, Germany (1983).
- Z. Movasaghi, S. Rehman, and I. U. Rehman, "Raman spectroscopy of biological tissues," *Appl. Spectrosc. Rev.* **42**(5), 493–541 (2007).
- W. Gray and G. Kocjan, *Diagnostic Cytopathology*, Churchill Livingstone/Elsevier, London, United Kingdom (2010).

51. M. D. Mignogna, S. Fedele, and L. L. Russo, "Dysplasia/neoplasia surveillance in oral lichen planus patients: a description of clinical criteria adopted at a single centre and their impact on prognosis," *Oral Oncol.* **42**(8), 819–824 (2006).
52. P. B. Sugerman and N. W. Savage, "Exfoliative cytology in clinical oral pathology," *Aust. Dent. J.* **41**(2), 71–74 (1996).
53. G. Ogden, "The future role for oral exfoliative cytology—bleak or bright?" *Oral Oncol.* **33**(1), 2–4 (1997).
54. C. M. Krishna et al., "Micro-Raman spectroscopy of mixed cancer cell populations," *Vib. Spectrosc.* **38**(1), 95–100 (2005).
55. F. Bonnier et al., "Processing ThinPrep cervical cytological samples for Raman spectroscopic analysis," *Anal. Methods* **6**(19), 7831–7841 (2014).
56. L. F. C. Carvalho et al., "Raman micro-spectroscopy for rapid screening of oral squamous cell carcinoma," *Exp. Mol. Pathol.* **98**(3), 502–509 (2015).
57. K. W. Short et al., "Raman spectroscopy detects biochemical changes due to proliferation in mammalian cell cultures," *Biophys. J.* **88**(6), 4274–4288 (2005).
58. S. Mian et al., "Raman spectroscopic analysis of normal, dysplastic and cancerous oral mucosa: a tissue engineering approach," in *The University of Sheffield Engineering Symp. Conf. Proc.*, Vol. 1 (2014).
59. H. S. Sekhon, J. L. Wright, and A. Churg, "Cigarette smoke causes rapid cell proliferation in small airways and associated pulmonary arteries," *Am. J. Physiol. Lung Cell. Mol. Physiol.* **267**(5), L557–L563 (1994).
60. V. Nanjappa et al., "Chronic exposure to chewing tobacco selects for overexpression of stearyl-CoA desaturase in normal oral keratinocytes," *Cancer Biol. Ther.* **16**(11), 1593–1603 (2015).
61. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, World Health Organization and International Agency for Research on Cancer, *Smokeless Tobacco and Some Tobacco-Specific N-nitrosamines*, World Health Organization, Lyon, France (2007).
62. L. Xue et al., "Raman spectral characteristics of oral squamous cell carcinoma, epithelial dysplasia and normal mucosa," *Zhonghua Kou Qiang Yi Xue Za Zhi.* **50**(1), 18–22 (2015).

Biographies for the authors are not available.