Articles
Radiation and Environmental Science Centre

2017

Development of Methodology for Raman Microspectroscopic Analysis of Oral Exfoliated Cells

Hugh Byrne
Technological University Dublin, hugh.byrne@tudublin.ie

Fiona Lyng
Technological University Dublin, Fiona.lyng@tudublin.ie

Alison Malkin
Technological University Dublin, alison.malkin@tudublin.ie

Ola Ibrahim
Technological University Dublin, ola.ibrahim@mydit.ie

Genecy Calado
Technological University Dublin

Follow this and additional works at: https://arrow.tudublin.ie/radart

See next page for additional authors Part of the Medicine and Health Sciences Commons

Authors

Hugh Byrne, Fiona Lyng, Alison Malkin, Ola Ibrahim, Genecy Calado, Steven Flint, and Isha Behl

Development of methodology for Raman microspectroscopic analysis of oral exfoliated cells

Isha Behl[1,2], Genecy Calado[1,2], Ola Ibrahim[1,2], Alison Malkin[3], Stephen Flint [5], Hugh J. Byrne[4], Fiona M. Lyng[1,2]

[1]DIT Centre for Radiation and Environmental Science, FOCAS Research Institute, Dublin Institute of Technology

[2]School of Physics, Dublin Institute of Technology

[3]School of Biological Sciences, Dublin Institute of Technology

[4]FOCAS Research Institute, Dublin Institute of Technology

[5]Oral medicine unit, Dublin Dental Hospital

Abstract

Oral squamous cell carcinoma ranks as the 15th most common cancer worldwide. The present study was undertaken to standardise a protocol for the analysis of oral exfoliated cells using Raman microspectroscopy. For this purpose, samples were obtained from two different sites, based on prevalence of disease *(*ventral side of the tongue and buccal mucosa). Different oral rinsing agents were employed and it was concluded that non-alcoholic mouthwash adequately removes food debris. Samples were collected using various collection tools and compared. It was observed that endo-cervical brushes yielded cells from deeper layers of the epithelium. Furthermore, monolayer formation of cells was carried out adopting cytospin and ThinPrep techniques and only the ThinPrep method provided flat and separated cells on the glass slide. Raman spectra were acquired from the nuclear and cytoplasmic regions of the cell using an XploRA confocal Raman instrument (HORIBA Jobin Yvon) with a 532nm laser as source. Glass spectral contamination was removed using non negatively constrained least squares (NNLS) algorithms. Corrected spectra were subjected to principal components analysis (PCA) which was able to differentiate the nucleus and cytoplasm regions of the cell; based on nucleic acid and protein features respectively. However, no classification of the two anatomically different sites was observed according to PCA or PCA-LDA (Linear discriminant analysis) using either the nuclear or cytoplasmic spectra. Nevertheless, the study has developed a standardised protocol for sample collection, sample preparation, spectral acquisition and data processing for future studies of oral exfoliated cells based on Raman microspectroscopy.

Keywords- Oral squamous cell carcinoma, Raman microspectroscopy, Non negatively constrained least squares analysis, Principal components analysis, Linear discriminant analysis

Introduction

Head and neck cancers are the 6th most common cancer worldwide. Oral squamous cell carcinoma (OSCC) ranks as the 15th most common cancer in the world and the 10th most frequent cancer in males [1]. Tongue and buccal mucosa are common sites of oral cancers in developed and developing countries respectively [2]. **90% of cancers that develop in the oral cavity are squamous cell carcinomas which involve epithelial cells [2-4].** The major aetiological factors include tobacco and alcohol consumption [2-4]. OSCC is preceded by leukoplakic (white) and erythroplakic (red) patches, known as precancerous lesions. Despite the easy accessibility of the oral cavity, the overall survival rate is lower than 50% [1-4]. The morbidity and mortality of oral cancer can be reduced when identified and treated at the precancerous stage as compared to more advanced stages. Therefore, identification of such precancerous and early cancerous pathologies is crucial to patient survival.

The primary OSCC diagnostic technique involves conventional oral examination (COE), which is followed by confirmatory diagnosis with histopathological examination of tissue. Biopsy is an invasive procedure which can cause psychological stress on patients. It also suffers certain shortcomings related to site selection, particularly when the lesions are large and can often vary in disease severity from one part of the lesion to another. In this case, it is important to select the most appropriate site for biopsy and multiple biopsies may be required from the lesion [5]. Brush biopsy can be a less-invasive means of diagnosing dysplasia and early carcinoma in such cases. Dysplastic and cancerous cells tend to have fewer and weaker connections to each other and to their neighbouring normal cells in the surrounding tissue. Therefore, these cells tend to exfoliate and can easily be collected from the surface of the lesion. A sample of these cells applied to a microscope slide will often contain abnormalities if harvested from a dysplastic or cancerous lesion. Exfoliative cytology is a simple, less-invasive technique that is well accepted by patients and can help with diagnosis of

precancerous and early carcinomatous lesions [3, 6]. Although conventional cytopathology has greatly improved the early detection of disease due to easy sampling and its less-invasive nature, like histopathology, this methodology also suffers from subjective interpretation. Hence, there is an unmet clinical need for new methods of screening and diagnosis which are rapid, objective and can help with prognosis of precancerous conditions. In view of the inherent accessibility of the oral cavity, there is ample scope for development of new, effective and non-invasive diagnostic methodologies, such as optical spectroscopic based diagnostic tools.

Optical spectroscopy methods are sensitive at the molecular level and are ideally poised to provide alternative methods for diagnosis and screening of oral cancers [7]. The potential of vibrational spectroscopy for clinical diagnosis based on exfoliated epithelial cells has already been demonstrated for the case of cervical cancer [8-10]. A limited number of infrared absorption spectroscopy studies have demonstrated the feasibility of using oral exfoliated cells. In 2009, Diem *et al* [11] reported biochemical variations using spectral cytopathology (SCP) and spectral histopathology (SHP) with multivariate statistical methods for differentiating normal from cancerous conditions. Schubert *et al* [12] developed the PapMap algorithm to process imaging data in order to decrease experimental time and begin to develop an automated process using Infrared microspectroscopy. In this study, they have shown that data collected in imaging mode is comparable to data collected in single point mode. In the same year, Papamarkakis *et al.* [13] demonstrated that oral cells could be classified according to anatomical region through principal components analysis (PCA). The cells from the tongue and the floor of the mouth were spectrally distinct. These compositional changes were attributed to the unique expression of keratins and the abundant expression of collagen in the cells of the tongue and floor of the mouth, respectively. In addition, samples from patients with reactive atypical changes or malignancy associated changes were seen to

be spectrally similar to a sample from a patient with squamous cell carcinoma, rather than samples from healthy volunteers. Furthermore, spectral changes were investigated in oral cells infected with the herpes simplex virus. This was due to shifts in proteins and phosphate levels which were associated with the degradation of host proteins and production of viral proteins coupled with the integration of viral genome with that of the host in the herpes simplex viral infected cells. Further studies from the same group [14, 15] showed that spectra from exfoliated cells from the tongues of healthy volunteers could be discriminated from exfoliated cells from the tongues of patients with oral dysplasia and cancer. Despite these studies using infrared spectroscopy, to the best of our knowledge, there have been no studies on single oral exfoliated cells using Raman spectroscopy. In order to develop new methods to differentiate between normal, pre-cancerous and cancerous oral exfoliated cells using Raman microspectroscopy, it is important to establish a standardised protocol for sample collection, preparation and Raman spectral acquisition and data processing. Therefore, the present study was undertaken with the aim of establishing a standardised protocol which includes selection of collection sites, mouthwash standardisation, sample collection tools and preparation, Raman spectra acquisition and glass correction.

2. Materials and methods

Oral cytology samples from ten healthy donors were used. Ethical approval was granted by the Dublin Institute of Technology Research Ethics Committee. The methodology involves collection of the sample using an appropriate collection tool, followed by formation of a monolayer of sample on a glass slide in order to acquire Raman spectra from the single cells. After spectral acquisition, the sample is subjected to Papanicolaou (Pap) staining. Raman spectra are processed and correlated with the stained sample (Figure 1). Every aspect of the protocol was taken into consideration individually to standardise the protocol.
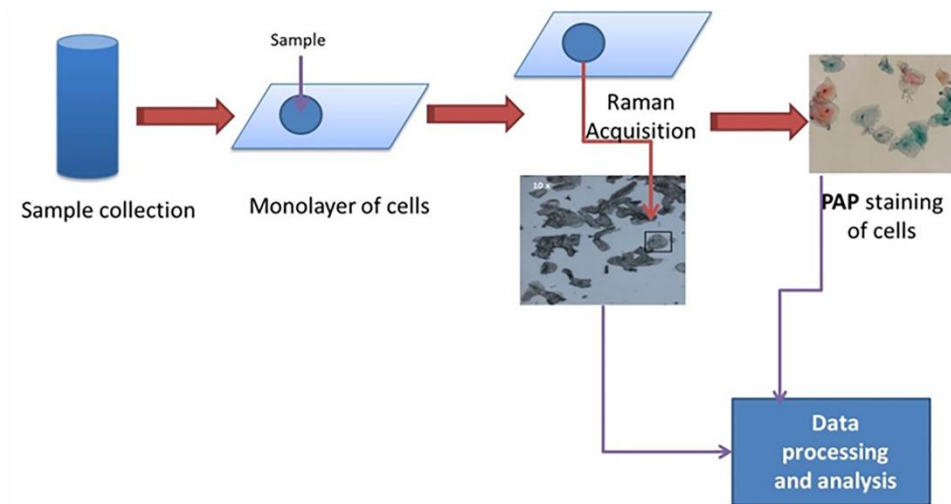
Figure 1: Schematic diagram of the basic protocol

2.1 Cell collection site

Two collection sites were chosen, the ventral side of the tongue and the buccal mucosa (Figure 2), as these are both common incidence sites for oral cancer [2-4].
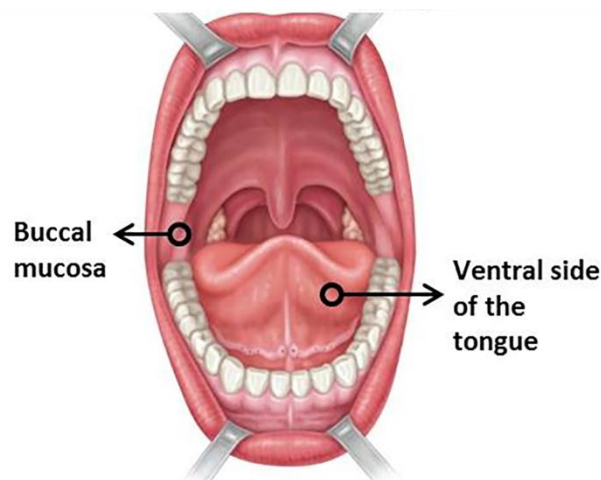

Figure 2: Collection site of the oral exfoliated cells
(http://www2.highlands.edu/academics/divisions/scipe/biology/labs/cartersville/2122/diagrams/digestivediagram3.jpg)

2.2 Mouthwash standardisation

Samples were collected at random times of the day, in order to minimise any potential impacts of food consumption and to develop a protocol independent of the sample collection time.

Samples were collected in four ways – a) without rinsing the mouth, and after b) rinsing the mouth using water, c) rinsing the mouth with alcohol based mouthwash (Corsodyl,

GlaxoSmithKline, Middlesex, UK) and d) rinsing the mouth with non-alcohol based mouthwash (Kin gingival, Labotatorios Kin S.A, Barcelona, Spain) as rinsing agents. Rinsing time was kept between 30 seconds and 1 minute before collecting the sample.

2.3 Collection of the samples

Standardisation of the protocol for sample collection involves developing a method that ensures an adequate cell number, including cells from the basal layer, for proper sampling. Histologically, normal mucosa is non keratinized epithelium comprised of three layers- basal, intermediate and superficial. Cells migrate from the basal to the superficial layer; the cells mature and finally desquamate after reaching the surface [16]. At the cellular level, Papanicolaou (Pap) stained smears of normal squamous epithelial cells show numerous superficial, intermediate and basal cells as shown in Figure 3.

Papanicolaou stain (Pap stain) is a multichromatic staining cytological technique developed by George Papanikolaou, the father of cytopathology. Papanicolaou stain includes both acidic and basic dyes. Acidic dye stains the basic components of the cell and basic dye stain the acidic components of the cell. The polychromatic PAP stain involves five dyes in three solutions.

1. **Hematoxylin : Natural dye hematoxylin is the nuclear stain which stains cell nuclei blue. It has affinity for chromatin, attaching to sulphate groups on the D.N.A. molecule. Harris' hematoxylin is the commonest cytologically although Gills' hematoxylin and Hematoxylin S can be used.**

2. **Orange Green 6: This is the first acidic counterstain (cytoplasmic stain) which stains matured and keratinized cells. The target structures are stained orange in different intensities.**

3. **Eosin Azure: This is the second counterstain which is a polychrome mixture of eosin Y, light green SF and Bismarck brown. <u>Eosin Y</u> gives a pink colour to**

cytoplasm of mature squamous cells, nucleoli, cilia and red blood cells. Staining solutions commonly used in cytology are EA 31 and EA 50, while EA 65 <u>Light green SF</u> stains blue to cytoplasm of metabolically active cells like parabasal squamous cells, intermediate squamous cells and columnar cells. <u>Bismarck brown Y</u> stains nothing and sometimes it is often omitted.
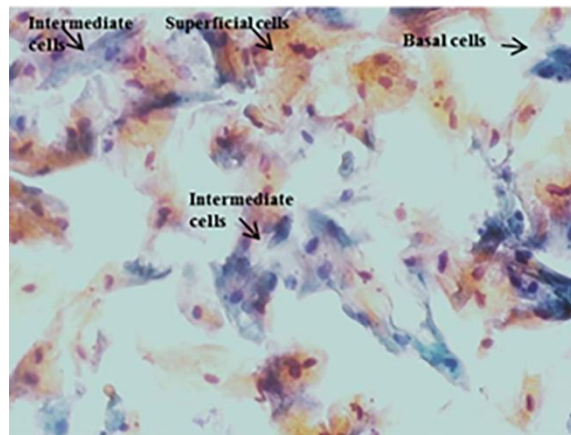

Figure 3: Pap stained oral exfoliated cells depicting cells from different layers of epithelium (50X): Basal cells which are cuboidal, small in size with a large nucleus and stain blue, Intermediate cells which are polygonal, squamous cells which are shed singly or in large sheets and stain blue and Superficial cells which are flat cells with smaller nuclei and stain pink.

As mild dysplastic cells can be present immediately above the basal layer, it was important to ensure that basal cells are present in the smear. The ideal way to collect the cells and ensure a sufficiently deep sample is to scrape from the surface and sub-surface layers by vigorous abrasion in one spot until bleeding starts. Due to limitations related to ethical issues, only moderate pressure was applied while scraping to avoid bleeding. To ensure proper sampling without bleeding, the collection site was vigorous rubbed (25 times) with the collection tool, using moderate pressure. Five types of sample collection tool were used; cotton swabs (Johnson & Johnson Limited, Wokingham, UK), perio-interdental brushes (Aquilant Scientific, Dublin, Ireland), broom type cyto-brush (Aquilant Scientific, Dublin, Ireland), oral specialised brush (Rover's Orcellex brush, Rovers Medical Devices B.V.,

Lekstraat, Netherlands), and endocervical cyto-brushes (Aquilant Scientific, Dublin, Ireland) to ensure proper sample collection. All collection tools were tested individually.

2.4 Sample preparation

Two cytology methods were tested for preparing a monolayer of cells on the glass slide, cytospin (Cytocentrifugation) and ThinPrep (Liquid based cytology). The Shandon CytoSpin III Cytocentrifuge (Thermo Shandon LTD., Cheshire, UK) and the ThinPrep processor (T2000, Hologic, Inc. LTD, Marlborough, USA) systems were used for the sample preparation.

The cytospin method is based on the principle of centrifugal force, whereby cells are transferred onto a microscope slide and the fluid from the suspension is absorbed onto a filter card. This allows the cells to adhere to the slide in a monolayer. After collection of the sample, the collection tool was rinsed in a centrifugal tube containing PreservCyt (ThinPrep Pap Test, Screenlink Healthcare, Dublin, Ireland). The PreservCyt solution helps to break up mucus and non-diagnostic debris while keeping true cell clusters intact. The centrifugal tube was then subjected to centrifugation at 1500 rpm (Heraeus Megafuge 16, Thermo scientific, Dublin, Ireland) for 1 min to pellet the cells. Subsequently, the supernatant was removed and cells were re-suspended in 1ml fresh PreservCyt solution. Then, the sample was loaded in the cytocentrifuge (Shandon CytoSpin III) at 800 rpm for 5 mins to form a monolayer of cells, ~6mm in diameter, on a glass slide (Sparks Lab Supplies Ltd, Dublin, Ireland). The slide was then post-fixed using industrial methylated spirit (IMS) (Lennox Laboratory Supplies, Dublin, Ireland) and air dried for 10-15mins.

In the case of ThinPrep, the sample was directly collected in ThinPrep vials (Screenlink Healthcare, Dublin. Ireland) and further processed in the ThinPrep processor (T2000). The processor homogenises the sample by spinning the filter, creating shear forces in the fluid that are strong enough to disaggregate randomly joined material, break up blood, mucus and

non-diagnostic debris, while keeping true cell clusters intact. The cells were collected onto the membrane of the TransCyt filter (Screenlink Healthcare, Dublin. Ireland) and transferred onto a ThinPrep glass slide (Screenlink Healthcare, Dublin. Ireland), to form a monolayer of cells which was approximately 20 mm in diameter. The slide was then ejected automatically into a fixative bath of industrial methylated spirit (IMS) and air dried for 10-15mins.

2.5 Raman spectra acquisition

Raman spectral acquisition parameters were as previously described for cervical Thinprep samples [14]. An XploRA confocal Raman instrument (HORIBA Jobin Yvon) was used for spectral acquisition. Manual calibration of the grating was done by using the 520.7cm^{-1} Raman line of crystalline silicon. Dark current measurement and recording of the substrate and optics signal was also performed, for data correction. As source, a 532 nm laser of ~12 mW power was focused by a 100X (MPlanN, Olympus, NA = 0.9) objective onto the sample; and the resultant Raman signals were detected using a spectrograph with a 1200 g⁄mm grating coupled with a CCD. Raman spectra were acquired in the 400 to 1800 cm^{-1} region with an integration time of 30 s ⁄ spectrum, and averaged over two accumulations. **Spectra were acquired from the center of the nucleus and at random from the cytoplasmic regions of the cells (Figure 4). As approximately 10-25 cells were recorded per slide, depending on the quality of the sample, it is expected that any heterogeneity of the cytoplasm should be averaged out.**
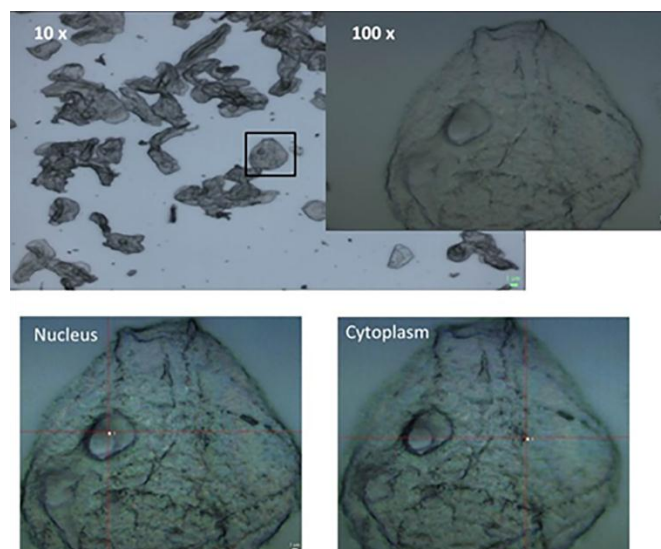
Figure 4: Representative image of the region of Raman spectral acquisition on cell.

6. *Data preprocessing and analysis*

Pre-processing and analysis of data was carried out using Matlab software (Mathworks). The scripts have been developed and adapted for the uploading of the spectra and their pre-processing, including smoothing (Savitzky- Golay, k=5; w=13), baseline correction (rubber band) and vector normalization [17, 18]. The glass substrate contributes its own spectrum, which can significantly obscure the cellular spectra. However, using the 532 nm wavelength as a source and a Raman microscope equipped with a 100x objective, the contribution from the substrate is considerably reduced [17, 19]. Furthermore, any residual contribution from the glass in the finger print region (400-1800 cm^{-1}) has been removed by using a non-negatively constrained least squares (NNLS) method. This method was previously applied to remove wax contamination digitally from wax embedded tissue sections [20, 21]. The NNLS algorithm is based on the Kuhn–Tucker condition for Problem of Least Square, which forces the values to be zero or positive [22]. The algorithm weights the values of glass as well as cell components in the acquired spectra from the cell and further subtracts the glass from the cell spectra applying non-negative constraints. Reference glass spectra were recorded on a clean ThinPrep glass slide. The glass corrected spectra were again subjected to pre-processing steps such as smoothing, baseline and normalisation corrections. Subsequently,

corrected spectra for nucleus and cytoplasm were subjected to Principal Components Analysis (PCA). PCA is an unsupervised method allowing evaluation of the variability existing in the data sets [18]. As shown in Supplemental Figure 1, although the variability of the substrate is encompassed by a lower number (n= ~60) of glass spectra, the PC loadings show that a larger number (n= ~300) of glass spectra better describe the full spectral weighting. Therefore, a matrix of n=300 glass spectra was employed for spectral correction. Notably, the process adds a negligible amount of time to the data processing step.

PCA-LDA (Linear discriminant analysis) was also carried out, by which PCA performs a feature reduction of the data and LDA classifies the data into one of two or more classes. PCA-LDA models were validated using leave-one-out cross-validation. In the leave-one-out method, one spectrum is removed from the data set and the algorithm is driven using the remaining spectra. The algorithm is then tested using the removed spectrum. This process is repeated for every spectrum in the data set, such that an estimate of the potential accuracy of future algorithms developed using the method in question can be calculated [23-25].

3. Results

Samples were collected from the tongue and buccal mucosa with / without rinsing agent. As expected, debris was observed when samples were collected without rinsing the mouth (Figure 5a). No significant difference was observed using water as a rinsing agent (Figure 5b), whereas no contamination was observed using an alcohol based mouthwash (Figure 5c). However, it was observed that more superficial cells were obtained after using the alcohol based mouthwash as a rinsing agent when compared to rinsing with water. Use of an alcohol free mouthwash as rinsing agent helped in the reduction of food debris, while the sample cellular profile was retained (Figure 5d).
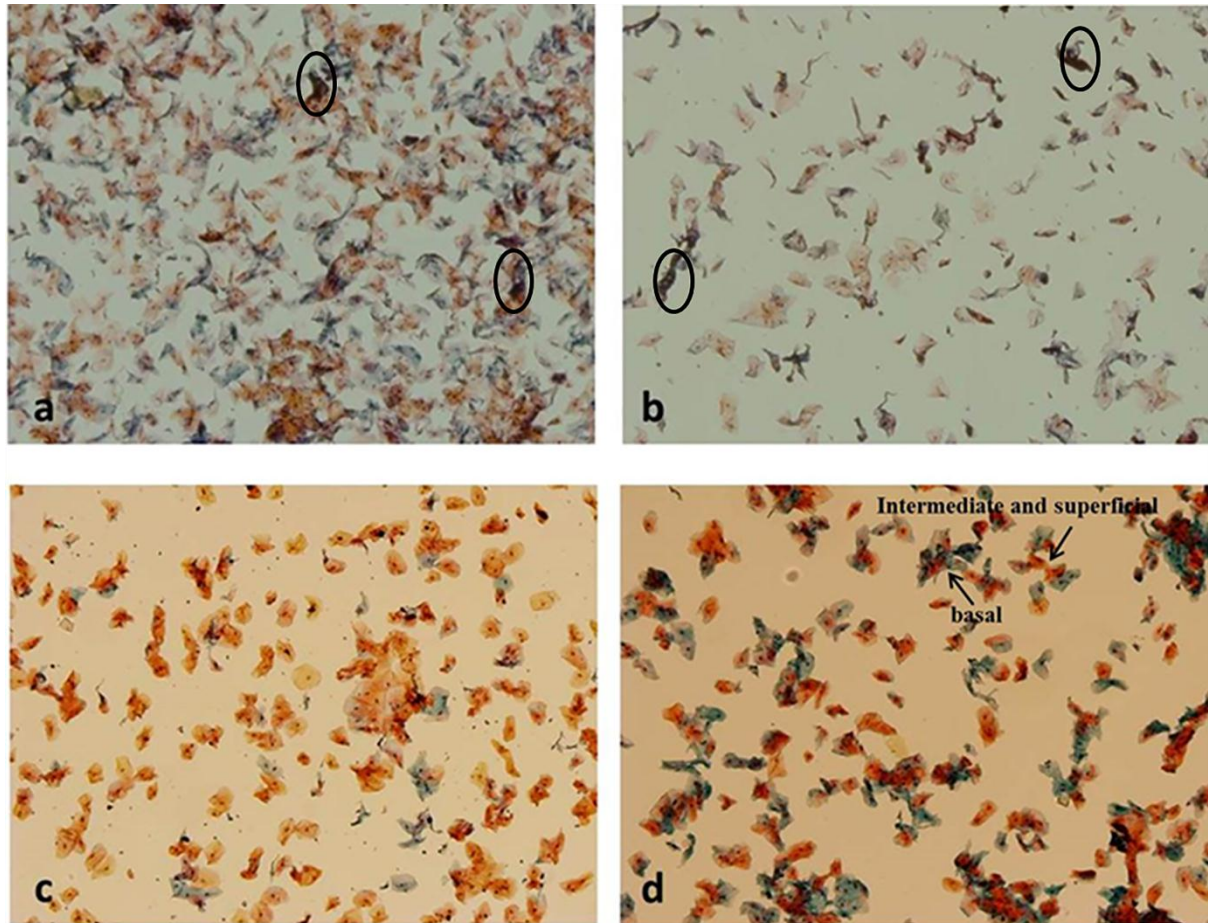
Figure 5: Pap stained oral exfoliated cells following various rinsing agents for rinsing mouth before collecting sample a) No rinsing agent (10X) b) water (10X) c) alcohol based mouthwash (10X) and d) non-alcohol based mouthwash (10X). **Debris is indicated by circles in a and b.**

Five different types of sample collection tool were tested to identify the most appropriate tool for sample collection; cotton swabs, perio-interdental brushes, broom type cyto-brush, oral specialised brush, and endocervical cyto-brushes. Firstly, exfoliated oral cell samples were collected using cotton swabs. They were preferred over other sampling tools to avoid any potential bleeding. Figure 6a shows that a good number of cells can be obtained with the cotton swabs, but cotton fibre contamination can be observed on the slide. Moreover, the cells had a tendency to stick to the cotton swabs. Next, oral exfoliated cell samples were collected using perio-interdental brushes (Figure 6b). Difficulties were encountered with bending of the brush due to the weak neck. As a result, the sample contained more mature cells from the superficial or intermediate layers rather than immature cells from the basal

layer. Additionally, the number of cells obtained was rather low. Thirdly, a broom like cytobrush was used (Figure 6c) and similar issues were encountered, as with the perio-interdental brushes, due to the soft bristles. More mature cells were obtained rather than immature cells. Further, a specialised oral brush (Rovers Orcellex brush) was also explored as a collection tool (Figure 6d). This brush had similar issues, namely a low number of cells and mainly superficial cells collected due to soft bristles. Lastly, endocervical cytobrushes were used for sampling (Figure 6e). Using this brush, an appropriate number of cells, including cells from the deeper layer of the epithelium, was obtained.

Figure 6: Various types of collection tools A) a) cotton swab b) Pap stained oral exfoliated cells (10X) c) Pap stained oral exfoliated cells showing cotton contamination (10X) B) a) perio-dental brush b) Pap stained oral exfoliated cells (10X) C) a) broom like cytobrush b) Pap stained oral exfoliated cells (10X) D) a) Rovers Orcellex brush b) Pap stained oral

exfoliated cells (10X) E) a) endocervical cyto-brush b) Pap stained oral exfoliated cells (10X) c) Pap stained oral exfoliated cells (40X) demonstrating basal, intermediate and superficial cells.

In order to optimise sample preparation, the cytospin and Thinprep methods were explored. For the cytospin method, cell clumping and folding was prominent. Reducing the cell concentration overcame the problem of cell clumping but **folding of the cells was still a problem. Since the results were not found to be consistent, the ThinPrep technique was explored, and it was observed that using ThinPrep method, cells were evenly distributed on the slide and moreover the cells were flat consistently (Figure 7). The concentration of cells was sample dependent, but typically 2500-6000 cells were collected in the ThinPrep vial.**
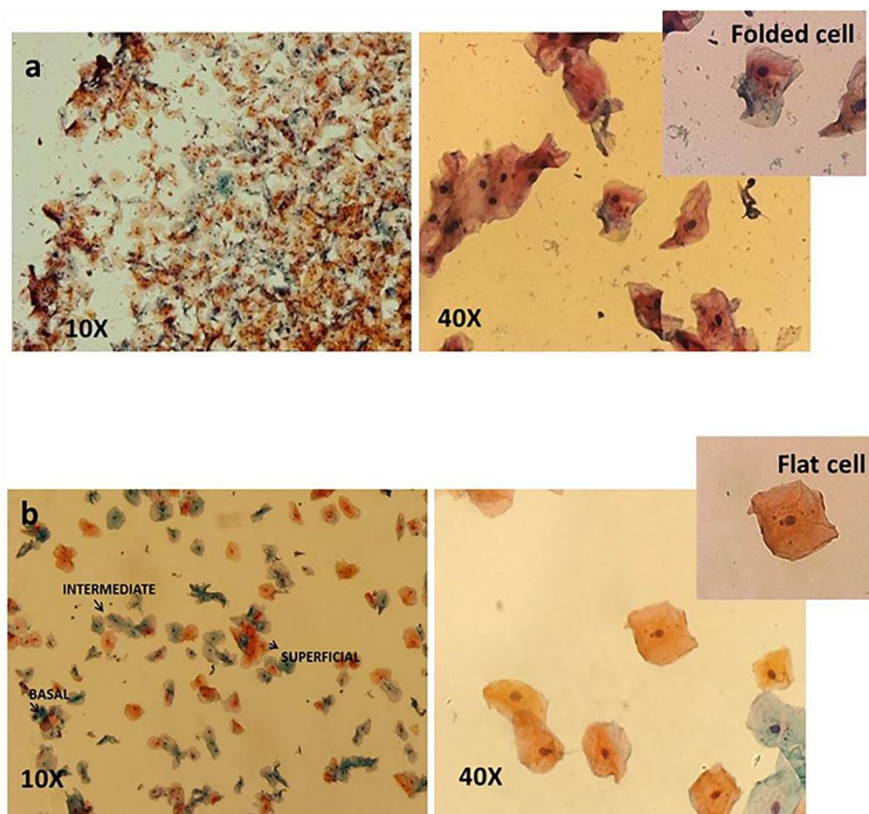

Figure 7: Depiction of two different methods for sample preparation a) Pap stained oral exfoliated cells following cytospin showing clumping of cells (10X) and folding of cells (40X) b) Pap stained oral exfoliated cells following ThinPrep showing good distribution of cells and cells from deeper layer of epithelium (10X) and flat cells (40X).

Raman spectra were acquired from the nucleus and cytoplasm of cells from the buccal mucosa, as shown in Figure 8. Glass contamination was evident to a greater extent in the spectra of the thinner cytoplasm of the cell rather than the nucleus. NNLS, described in section 2.6, was used to remove glass contamination from the spectra, and was seen to successfully eliminate the glass bands at 445, 560, 790, 1100 cm^{-1} (Figure 8).
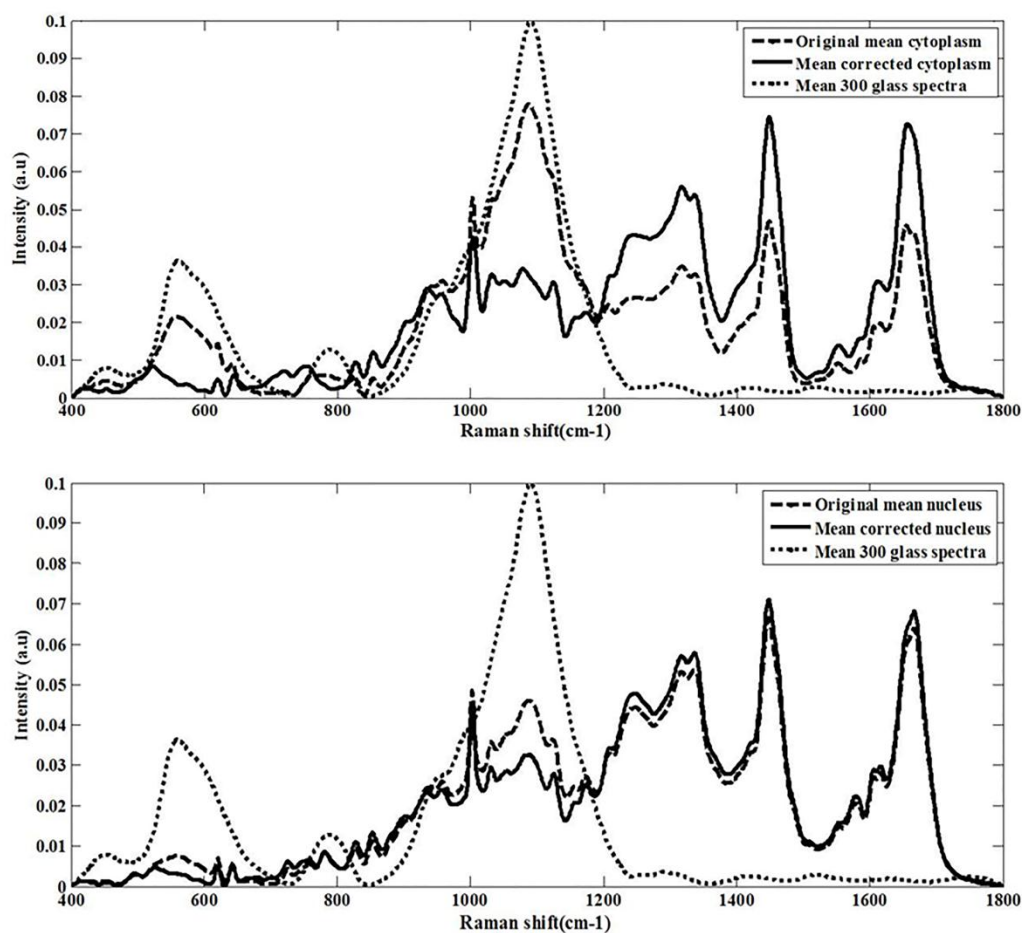

Figure 8: Depicting glass corrections for a) cytoplasm and b) nucleus of buccal mucosa cells. (---Original mean spectra of cytoplasm and nucleus, ⎯ Mean corrected cytoplasm and nucleus, ...Mean 300 glass spectra)

Prominent bands of proteins (755, 1004, 1033, 1450, 1660 cm^{-1}) were observed in the mean spectra of the cytoplasm whereas nucleic acid bands (720, 780, 1340, 1576 cm^{-1}) were observed in the mean spectra of the nucleus (Figure 9a) [18, 26]. In the PCA scatter plot of nucleus and cytoplasm, segregation of the cytological regions was observed according to Principal component 1 (PC1), which explains 63.58% of the variance, while principal

component 2 (PC2) explains 8.35% (Figure 9b). The loading of PC1 is dominated by positive protein bands associated with the cytoplasm at 936, 1007, 1038, 1455, 1654 cm^{-1}. Negative peaks of the loading of PC1 correspond to the nucleus and nucleic acid bands were observed at 726, 784, 1094, 1375, 1485,1 576 cm^{-1}, as shown in Figure 9c [18, 26]. No glass contamination can be seen, as no broad features of glass can be seen in the loading of PC1.
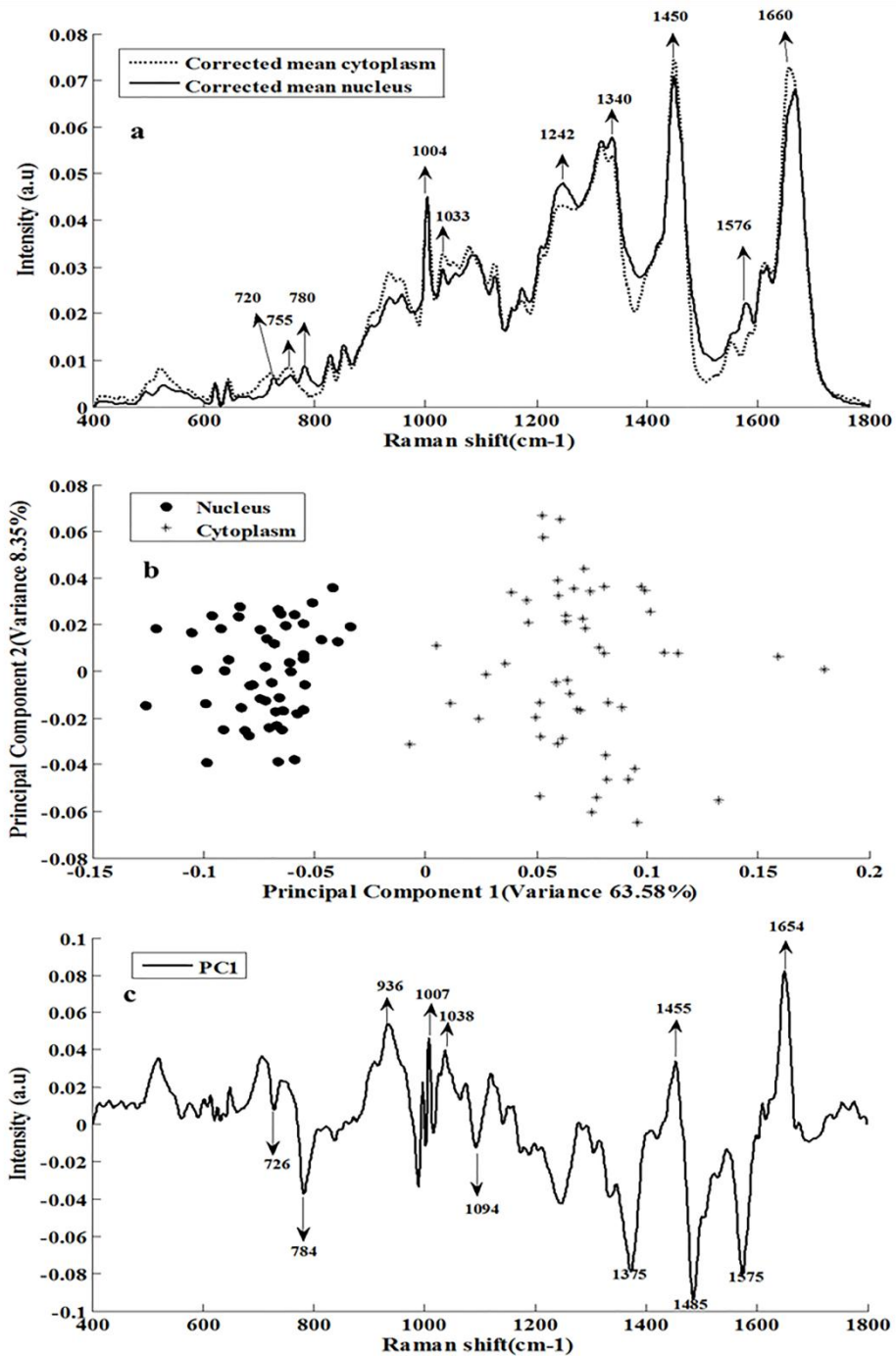

Figure 9: a) Mean spectra of cytoplasm and nucleus for buccal mucosa b) PCA scatter plot and c) PC1 loading

Additionally, spectra of nucleus and cytoplasm were acquired for the two sites *i.e.* buccal mucosa and tongue. PCA of the combined datasets demonstrated a similar separation of nuclear and cytoplasmic regions, as in the case of the single site, shown in Figure 9, with a similar differentiation according to PC1 (Figure 10). However, no differentiation of the nuclear or cytoplasmic spectra of the different sites was observed. PCA-LDA was then employed to further explore the feasibility of differentiation between the two different anatomical sites. The results are summarised in the confusion matrix shown in Tables 1 & 2. In the case of the cytoplasm of tongue and buccal mucosa (Table 1), only 214 out of 394 spectra were correctly classified, a classification efficiency of 54.3%. Similar values of specificity (53.9%) and sensitivity (54.7%) were determined for the cytoplasmic regions of both sites, whereas in the case of the nuclear region (Table 2), only 226 out of 394 spectra were correctly classified, a 57.36% classification efficiency. Sensitivity and specificity values of 57% and 57.75% respectively were calculated for the nuclear regions of tongue and buccal mucosa.
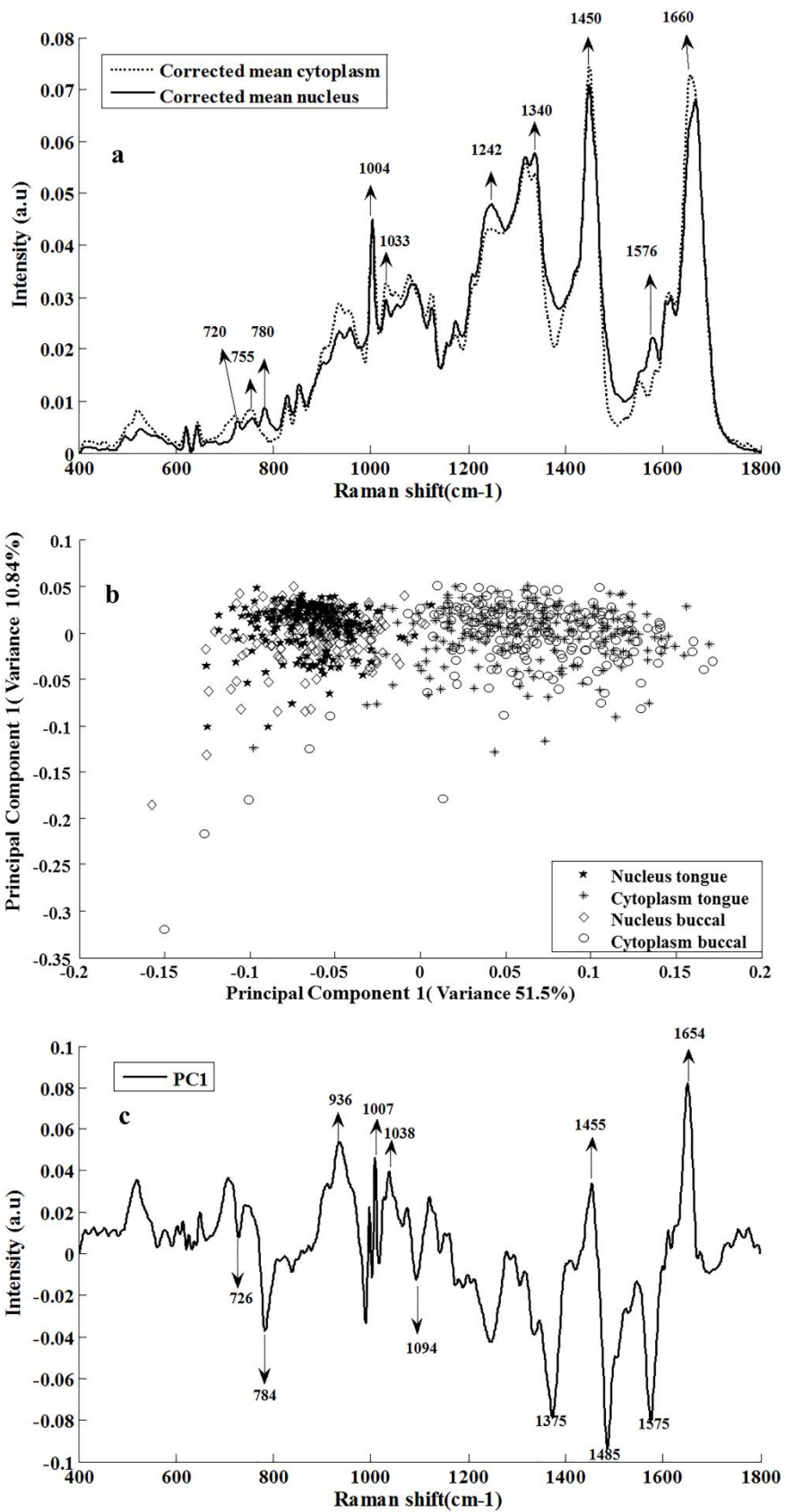
Figure 10: a) Mean spectra of nucleus and cytoplasm from tongue and buccal mucosa b) PCA scatter plot and c) PC1 loading

Table 1: Confusion matrix for discrimination of cytoplasmic region of tongue and buccal mucosa by PCA-LDA

	Cytoplasm Tongue	Cytoplasm Buccal
Cytoplasm Tongue	99	98
Cytoplasm Buccal	82	115
Classification Efficiency	54.3%	
Sensitivity	54.7%	
Specificity	53.9%	

Table 2: Confusion matrix for discrimination of nuclear region of tongue and buccal mucosa by PCA-LDA

	Nucleus Tongue	Nucleus Buccal
Nucleus Tongue	118	79
Nucleus Buccal	89	108
Classification Efficiency	57.36%	
Sensitivity	57%	
Specificity	57.75%	

4. DISCUSSION

In the present study, a methodology was developed for Raman spectroscopy of oral exfoliated cells. The complete protocol has been summarised in Figure 11. Every step of the protocol was standardised, which was initiated with the selection of the site for sample collection, chosen on the basis of frequency of occurrence of oral cancer. Cytological samples from the oral cavity are prone to contamination by food debris, due to mastication. It is important therefore to standardise the method to get rid of such food particles from the mouth before sample collection. This was demonstrated by collecting sample without rinsing the mouth. It was found that water as rinsing agent has no effect on clearing debris. Furthermore, it was observed that more superficial cells were obtained while using a typical mouthwash as a rinsing agent. This can be attributed to the fact that the mouthwash employed was alcohol based, which made superficial cells exfoliate more easily. Use of an alcohol free mouthwash as a rinsing agent allowed the collection of both superficial and intermediate cells as well as basal cells from the deeper layer of the epithelium. Thus, it was concluded that the non-alcoholic based mouthwash was the most appropriate rinsing agent.

Next, different collection tools were tested to reach an appropriate depth of the epithelium resulting in a good number of cells on the slide for screening. As it is known that mild dysplasia cells are present near the basal layer, it is important to collect cells from the deeper layer of the epithelium. Fiber contamination was observed while using cotton swabs as a collection tool, whereas the other brushes had bristles which were too soft to collect cells from deeper layers of the epithelium. It was observed that use of endo-cervical brushes as a collection tool gave the best results in terms of number of cells obtained. In addition, the samples collected contained cells from the basal layer of the epithelium, which made it an ideal tool for sample collection.

While testing two different sample preparation methods, it was observed that the ThinPrep method forms the best monolayer of cells on the glass slide. This can be attributed to the difference in the principles of each technique. In the case of cytospin, the cells are cast on the slide by centrifugal force, which can cause them to fold during deposition. In the case of ThinPrep, cells are deposited on the slide using pressure due to which more cells are flat. Moreover, the overall diameter of the monolayer was considerably smaller in the cytospin samples when compared with the ThinPrep samples, and as a result cytospin samples had a higher surface coverage, and the cells had a tendency to clump together. Therefore, it was concluded that ThinPrep is a better technique than cytospin for oral cytology sample preparation for this study.

Strong broad bands of the glass substrate were observed while acquiring Raman spectra from the nucleus and cytoplasm. In order to correct for the glass contribution, the NNLS algorithm was used [22], using 300 spectra rather than a single glass spectrum to account for any variability of the glass [27]. Following this method of glass correction, the nucleus and cytoplasm of buccal cells could be discriminated using PCA with no glass interference in the PC loadings.

In 2010, Papamarkakis *et al.* observed that spectra from cells exfoliated from the tongue were differentiated from spectra from cells exfoliated from the buccal mucosa due to the presence of keratin in the tongue cells. [13]. In contrast, no significant differentiation was observed between the spectra of either the nuclear or cytoplasmic regions of buccal mucosa and tongue, using PCA in the present study. Further, PCA-LDA cross validation, also suggests that both the nucleus and cytoplasm of tongue and buccal mucosa cells have similar spectral profiles. The epithelium of both the buccal mucosa and the ventral side of the tongue are non-keratinising epithelium**,** hence have similar biochemical properties. The spectral similarity of the two different anatomical sites is thus not surprising [16]. **Furthermore,**

since mild dysplastic cells can be present in the intermediate layer, spectra were recorded from intermediate and superficial cells, immediately above the basal layer. It was earlier reported that intermediate and superficial layers of the oral epithelium share spectral similarities [28]. Thus, it can be concluded that the differences can be attributed to the similarity of the epithelium. Notably, in human lung cancer cell lines, significant discrimination was only achievable when the nucleoli of the cells were specifically targeted for Raman spectroscopic analysis [25].
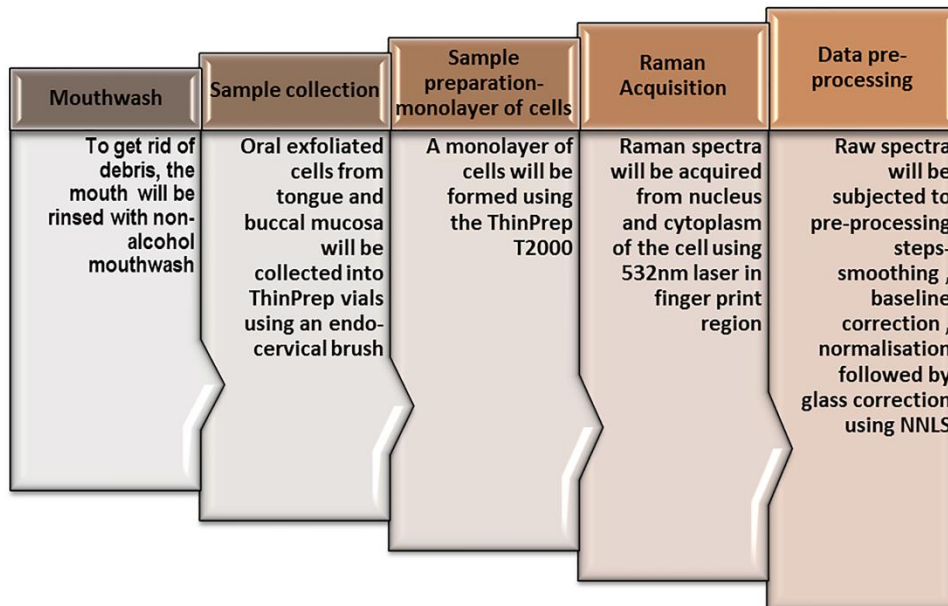

Figure 11: Optimized protocol of sample processing for the Raman microspectroscopic study

5. Conclusion:

It can be concluded that, for oral exfoliated samples for Raman spectroscopy, the mouth should be rinsed with a non-alcohol mouthwash before collecting samples using an endo-cervical brush. Sample preparation using the Thinprep method provides a monolayer of cells on the slide. Raman spectral acquisition in the fingerprint region from the nucleus and

cytoplasm, using 532 nm as source, gives good quality spectra with significant but not dominant contributions from the glass substrate, which, after preprocessing by smoothing, baseline correction and normalisation, can be removed using NNLS. PCA demonstrates a good differentiation of the nuclear and cytoplasmic regions, according to loadings consistent with the known biochemical differences between the regions. However, neither the nucleus nor the cytoplasm of the two different anatomical regions could be differentiated using PCA-LDA. This can be attributed to the fact that the epithelium of the buccal mucosa and the ventral side of the tongue are non-keratinising epithelium.

This study has developed a standardised protocol for the analysis of oral exfoliated cells based on Raman microspectroscopy. Thus, the protocol developed throughout this study will pave the way for future studies based on oral exfoliated cells and will contribute to the realisation of the application of vibrational spectroscopy for early detection of oral cancer.

6. Acknowledgment:

References:

1) World Health Organization: The GLOBOCAN 2012 database, http:// globocan.iarc.fr

2) Chi A. C., Day T. A., Neville B. W., "Squamous cell carcinoma and precursor lesions of the oral cavity: epidemiology and aetiology". CA Cancer J Clin (2015);65:401–421

3) Kumar S., Vezhavendhan N., Priya S., "Role of Oral Exfoliative Cytology in Oral Leukoplakia and Squamous Cell Carcinoma".Int J Clin Den Sci, (2011) , 2(1):93-97

4) Neville B. W., Day T. A., "Oral cancer and precancerous lesions". CA Cancer J Clin (2002) , 52:195-215

5) Poh C. F., Ng S., Berean K. W., Williams P. M., Rosin M. P., Zhang L., "Biopsy and Histopathologica Diagnosis of Oral Premalignant and Malignant Lesions". J. Canadian Dent. Asso. (2008), 74(3): 283-288

6) Mendes S. F., Ramos G. deO., Rivero E. R. C., Modolo F., Grando L. J., Meurer M. I. "Techniques for precancerous lesion diagnosis". J. Oncol, (2011) , 2011

7) Singh S.P., Ibrahim O., Byrne H. J., Mikkonen J. W., Koistinen A. P., Kullaa A. M., Lyng F. M. "Recent advances in optical diagnosis of oral cancers: Review and future perspectives". Head Neck (2016), 38: E2403–E2411

8) Ramos IR, Meade AD, Ibrahim O, Byrne HJ, McMenamin M, McKenna M, Malkin A, Lyng FM, Raman spectroscopy for cytopathology of exfoliated cervical cells, Faraday Discussions 2016, 187, 187-198

9) Lyng FM, Traynor D, Ramos IRM, Bonnier F, Byrne HJ. Raman spectroscopy for screening and diagnosis of cervical cancer, Analytical and Bioanalytical Chemistry 2015, 407(27):8279-89

10) Bonnier F, Traynor D, Kearney P, Clarke C, Knief P, Martin C, O'Leary JJ, Byrne HJ and Lyng F, Processing ThinPrep cervical cytological samples for Raman spectroscopic analysis, Anal. Methods, 2014, 6, 7831

11) Diem M., Papamarkakis K., Schubert J., Bird B, Romeo., M. J., Milikovic M., "The Infrared Spectral Signatures of Disease: Extracting the Distinguishing Spectral Features Between Normal and Diseased States", Appl. Spec.(2009), 63 (11), 307-318

12) Schubert J. M., Mazur A. I., Bird B., Miljkovic´ M., Diem M. "Single point vs. mapping approach for spectral cytopathology (SCP)". J. Biophoton (2010) 3(8–9): 588–596

13) Papamarkakis K., Bird B., Schubert J. M., Miljkovic M., Wein R., Bedrossian K., Laver N., Diem M. "Cytopathology by optical methods: spectral cytopathology of the oral mucosa". Lab. Inv. (2010), 90(4):589-98

14) Diem M., Mazur A., lenau K., Schubert J., Bird B., Milikovic M., Krafft C., popp J., "Molecular pathology

via IR and Raman spectral imaging", J. biophotonics (2013), 6(11-12), 855-886

15) Miljkovi´c M., Bird B., Lenau K., Mazur A. I., Diem M., "Spectral cytopathology: new aspects of data collection, manipulation and confounding effects", Analyst(2013), 138, 3975–3982

16) Kumar V., Abbas A. K, Fausto N., Robbins and Cotran Pathologic Basis of Disease, 7th Edition

17) Bonnier F., Traynor D., Kearney P.,Clarke C.,   Knief P., Cara Martin, John J. O'Leary, Hugh J. Byrne and Fiona Lyng, "Processing ThinPrep Cervical Cytological Samples for Raman Spectroscopic Analysis" Anal. Methods (2014), 6: 7831-7841

18) Bonnier F., Knief P., Lim B., Meade A., Dorney J., "Imaging live cells grown on a three dimensional collagen matrix using Raman microspectroscopy". Analyst (2010), 2, 135, 3169-3177

19) Kerr L. T., Byrne H. J., Hennelly B. M., "Optimal choice of sample substrate and laser wavelength for Raman spectroscopic analysis of biological specimen",Anal. Methods (2015), 7, 5041-5952

20) Vrabie V., Gobinet C., Piot O., Tfayli A., Bernard P., Huez R., Manfait M., "Independent component analysis of Raman spectra: Application on paraffin-embedded skin biopsies", Biomedical Signal Processing and Control(2007), 2, 40-50.

21) Tfayli A., Gobinet C., Vrabie V., Huez R., Manfait M.,Piot O., "Digital dewaxing of Raman signals: discrimination between nevi and melanoma spectra obtained from paraffin-embedded skin biopsies", Applied Spectroscopy (2009), 63, 564-570.

22) Chen J., Member, Richard C., Bermudez J. M., Honeine P., Member, "Variants of Non-Negative Least-Mean-Square Algorithm and Convergence Analysis". IEEE Transactions on Signal Processing (2014), 62(15): 3990-4005

23) Stone M.. "Asymptotics for and against cross-validation" Biometrika(1977), 64 (1): 29–35. 107.

24) Efron, B., Tibshirani, R. "Improvements on cross-validation: The .632 + Bootstrap Method"(1997). Journal of the American Statistical Association 92 (438): 548–560

25) Farhane Z., Bonnier F., Casey A., Maguire A., O'Neill L. , Byrne H.J.,"Cellular discrimination using in vitro Raman micro spectroscopy: the role of the nucleolus", Analyst (2015), 140, 5908-5919

26) Bryne H. J., Sochkalingum G. D., Stone N. "Raman Microscopy: Complement or Competitor?", Biomedical Applications of Synchrotron Infrared Microspectroscopy (2011), chapter 4: 105-142

27) Ibrahim O., Maguire A., Meade A. D., Flint S., Toner M., Byrne H. J., Lyng F. M. "Improved protocols for pre-processing Raman spectra of formalin fixed paraffin preserved tissue sections", Analyst (submitted)

28) Behl I, Kukreja L, Deshmukh A, Singh S.P., Mamgain H, Hole A. R., C. Murali Krishna, " Raman

mapping of oral buccal mucosa: A spectral histopathology approach*" J. Biomed. Opt.*, (2014) ,19(12), 126005
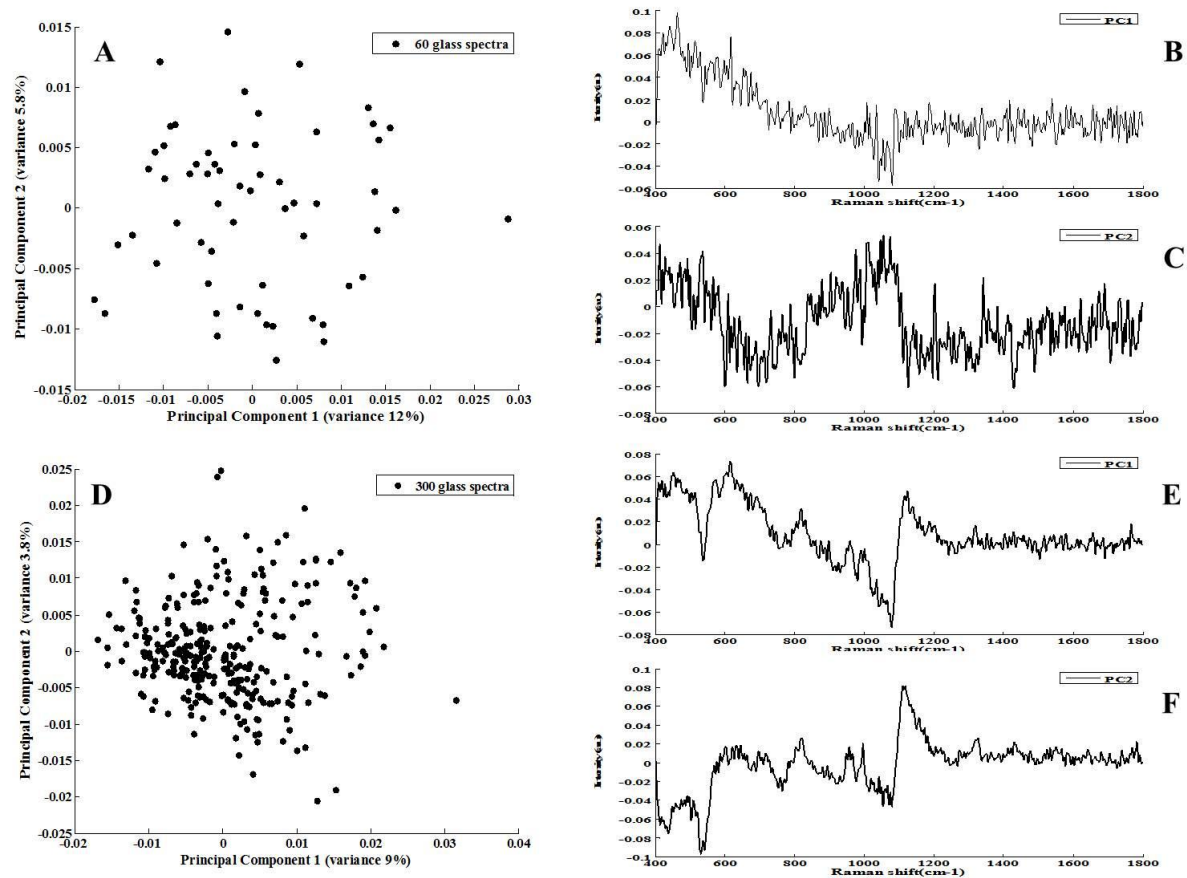
Supplemental Figure:

Figure 1: A) PCA scatter plot, B) PC1 loading and C) PC2 loading for 60 glass spectra; D) PCA scatter plot, E) PC1 loading and F) PC2 loading for 300 glass spectra