Article type: Full Article

Raman spectral discrimination in human liquid biopsies of oesophageal transformation to adenocarcinoma

Ishaan Maitra[1,*], Camilo L. M. Morais[1], Kássio M. G. Lima[1,2,*], Katherine M. Ashton[3], Ravindra S. Date[3], and Francis L. Martin[1,*]

[1] *School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK*

[2] *Institute of Chemistry, Biological Chemistry and Chemometrics, Federal University of Rio Grande do Norte, Natal, Brazil*

[3] *Lancashire Teaching Hospitals NHS Foundation Trust, Royal Preston Hospital, Preston PR2 9HT, UK*

***Correspondence:**

Dr. Ishaan Maitra, School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK. E-mail: imaitra@uclan.ac.uk

Prof. Kássio M. G. Lima, Institute of Chemistry, Biological Chemistry and Chemometrics, Federal University of Rio Grande do Norte, Natal, Brazil. E-mail: kassiolima@gmail.com

Prof. Francis L. Martin, School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK. E-mail: flmartin@uclan.ac.uk

KEYWORDS: biofluids, chemometrics, classification techniques, oesophageal cancer, Raman spectroscopy

Abstract

The aim of this study was to determine whether Raman spectroscopy combined with chemometric analysis can be applied to interrogate biofluids (plasma, serum, saliva and urine) towards detecting oesophageal stages through to oesophageal adenocarcinoma (normal/squamous epithelium, inflammatory, Barrett´s, low-grade dysplasia [LGD], high-grade dysplasia [HGD], and oesophageal adenocarcinoma [OAC]). The chemometric analysis of the spectral data was performed using principal component analysis (PCA), successive projections algorithm (SPA) or genetic algorithm (GA) followed by quadratic discriminant analysis (QDA). The GA-QDA model using a few selected wavenumbers for saliva and urine samples achieved 100% classification for all classes. For plasma and serum, the GA-QDA model achieved excellent accuracy in all oesophageal stages (>90%). The main GA-QDA features responsible for sample discrimination were: 1012 cm^{-1} (C-O stretching of ribose), 1336 cm^{-1} (Amide III and CH_2 wagging vibrations from glycine backbone), 1450 cm^{-1} (methylene deformation), and 1660 cm^{-1} (Amide I). The results of this study are promising and support the concept that Raman on biofluids may become a useful and objective diagnostic tool to identify oesophageal disease stages from squamous epithelium to OAC.

Abbreviations: **BO**, Barrett's oesophagus; **CCD**, charge-coupled device; **GA**, genetic algorithm; **GA-QDA**, genetic algorithm quadratic discriminant analysis; **GI**, gastro-intestinal, **GORD**, gastro-oesophageal reflux disease; **HGD**, high-grade dysplasia; **LGD**, low-grade dysplasia; **OAC**, oesophageal adenocarcinoma; **PCA**, principal component analysis; **PCA-QDA**, principal component analysis quadratic discriminant analysis; **QDA**, quadratic discriminant analysis; **SPA-QDA**, successive projections algorithm quadratic discriminant analysis.

1 INTRODUCTION

Barrett's oesophagus (BO) is the only known precursor to oesophageal adenocarcinoma (OAC) to date as a result of chronic inflammation from gastro-oesophageal reflux disease (GORD). GORD increases the risk of OAC by greater than 40-fold compared with the general population [1]. At present there are no reliable predictive biomarkers that might enable us to risk-stratify BO patients and identify those who would benefit most from endoscopic management [2]. Early detection and prevention are the key strategies to manage OAC. Early detection of cancer or dysplasia in BO allows intervention at an early stage. The argument as to which BO patients are most likely to benefit from surveillance and management hinges on the high prevalence of BO and the low cancer incidence among unselected BO cases, versus the burden of invasive treatment and the high morbidity and mortality from OAC [3].

There is a continuing effort in the search of new technology that can detect early biochemical signs of malignancy and therefore significantly reduce morbidity and mortality. Biomarkers including carbohydrates, proteins, lipids and nucleic acids can be used for risk assessment, diagnosis, prognosis, and for the prediction of treatment efficacy [4]. Exploration of biofluids has only been recently explored by vibrational spectroscopic techniques [4]. They are easily accessible, minimally invasive and exhibit fingerprint spectra that have characteristic bands reflecting their biochemical structure.

Raman spectroscopy is complementary to infrared (IR) technology and has advantageous properties when analysing biofluids. Contrary to conventional detection methods, optical spectroscopy on plasma, serum, saliva or urine provides an opportunity to diagnose diseases non-invasively. Raman spectroscopy is based on inelastic vibrational scattering, which can detect the secondary constitution of molecules [5]. No labelling is necessary and the technique provides high spectral sensitivity [6]. Spectroscopy on biofluids has been focused on plasma and serum analysis due to the large readily available biobanks in

research laboratories. Multiple studies have been performed demonstrating the potential of Raman spectroscopy for differentiating normal subjects from patients with colorectal [7], hepatocellular [8], cervical [9], and breast cancers [10]. Raman spectroscopy in serum has been performed to differentiate controls from oral cancers (tongue cancers) [11]; however, to date, no research has been established using Raman spectroscopy to investigate oesophageal transformation to adenocarcinoma from human body fluids.

This paper proposes an accurate, fast, and inexpensive method using biofluids (plasma, saliva, serum and urine) for detecting oesophageal stages through to OAC (normal; inflammatory; Barrett´s; low-grade dysplasia (LGD), high-grade dysplasia (HGD); and, OAC) using Raman spectroscopy.

2 MATERIAL AND METHODS

2.1 Sample Collection

Patients were identified from Upper gastro-intestinal (GI) multi-disciplinary team meetings and pathology hospital databases which had been created by the Pathology Laboratory manager and one of the Consultant GI Histopathologists. Potential patients were identified prospectively and consent for biofluids (blood for plasma and serum; urine and saliva) was taken between October 2017 and June 2019 in a clinic or endoscopy setting. The biofluid specimens were categorised as follows: i) plasma: n=35 normal, n=18 inflammatory, n=27 Barrett´s, n=6 LGD, n=12 HGD and n=22 OAC (set A); ii) saliva: n=35 normal, n=18 inflammatory, n=26 Barrett´s, n=5 LGD, n=10 HGD and n=24 OAC (set B); serum: n=36 normal, n=19 inflammatory, n=28 Barrett´s, n=6 LGD, n=12 HGD and n=23 OAC (set C); and, urine: n=38 normal, n=19 inflammatory, n=27 Barrett´s, n=6 LGD, n=11 HGD and n=26 OAC (set D).

Ethical approval was granted by the East of England - Cambridge Central Research Ethics Committee from 2015 (Archival gastro-intestinal tissue, blood, saliva and urine

collection; REC reference: 18/EE/0069; IRAS project ID: 242639). Ethics was also granted from the parent University (STEMH 909 application). All biofluids taken at source patient contact were stored in their containers in the fridge at 4°C to 7°C. Prior to freezing, blood samples were centrifuged at 20°C at 2200 rpm for 15 min to obtain plasma and serum samples (local protocol). Saliva samples were taken from patients 3 to 6 h prior to ingestion of solids or liquids. All biofluids were then snap frozen and stored in the freezer at -80°C.

Prior to slide preparation, biofluids samples were left to thaw in the fridge at 7°C. Thirty mL of individual biofluids (plasma, serum and saliva) were pipetted onto aluminium foil-lined FisherBrand™ slides for Raman spectroscopy analysis. Urine was centrifuged at 2200 rpm at 20°C with the supernatant pipetted onto each slide. Each slide was labelled with a specific GI (Gastrointestinal number) used to anonymise samples. All slides were left to dry prior to transportation in wooden slide boxes to the spectroscopy laboratory for analysis. Samples were stored in a de-humidified glass container to prevent condensation and physical damage.

2.2 Raman Spectroscopy

Raman point spectra acquisition was performed with an InVia Renishaw Raman spectrometer coupled with a charge-coupled device (CCD) detector and a Leica microscope. A 200-mW laser diode was used at a wavelength of 785 nm with a grating of 1200 lines/mm. Exposure time was set at 10 s, with 5% laser power, and 2 accumulations at a spectral range between 2000−400 cm^{-1}. Twenty-five point spectra were taken per sample using a 20× objective to focus the laser beam on the sample.

2.3 Data Analysis and Chemometric Methods

The data import, pre-treatment and construction of chemometric classification models (principal component analysis quadratic discriminant analysis [PCA-QDA], successive projections algorithm quadratic discriminant analysis [SPA-QDA] and genetic algorithm quadratic discriminant analysis [GA-QDA]) were implemented in MATLAB R2014a

software (MathWorks, USA) by using the PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., USA) and laboratory-made routines. The raw spectra were pre-processed by cutting between 1800 and 800 cm^{-1} (939 wavenumbers at 4 cm^{-1} spectral resolution). Cosmic rays were corrected using the Renishaw WiRE software system, and baseline distortions due to fluorescence interference were corrected with the asymmetric least squares (ALS) baseline correction algorithm [12]. For PCA–QDA, SPA–QDA and GA–QDA models, the samples were divided into training (60%), validation (20%) and prediction sets (20%) by applying the classic Kennard–Stone (KS) uniform sampling algorithm [13] to the spectra as shown in Table 1. The optimum number of variables for SPA-QDA and GA-QDA were performed with an average risk G of QDA misclassification. Such a cost function was calculated in the validation set as:

$$\boldsymbol{G} = \frac{\mathbf{1}}{\boldsymbol{N_V}} \sum_{\boldsymbol{n=1}}^{\boldsymbol{N_V}} \boldsymbol{g_n} \tag{1}$$

where $\boldsymbol{N_V}$ is the number of validation spectra and $\boldsymbol{g_n}$ is defined as:

$$\boldsymbol{g_n} = \frac{\boldsymbol{r^2(x_n, m_{I(n)})}}{\mathbf{min}_{\boldsymbol{I(m)\neq I(n)}}\, \boldsymbol{r^2(x_n, m_{I(m)})}} \tag{2}$$

where $\boldsymbol{I(n)}$ is the index of the true class for the nth validation object $\boldsymbol{x_n}$. In this definition, the numerator is the squared Mahalanobis distance between object $\boldsymbol{x_n}$ (of class index $\boldsymbol{I(n)}$) and the sample mean $\boldsymbol{m_{I(n)}}$ of its true class. The denominator in Eq. (2) corresponds to the squared Mahalanobis distance between object $\boldsymbol{x_n}$ and the centre of the closest wrong class, $\boldsymbol{m_{I(m)}}$. The minimum value of the cost function (maximum fitness) will be achieved when the selected variables from the original data are as close as possible to its true class and more distance as possible from its wrong class according to the validation samples. The GA routine was carried out during 100 generations with 200 chromosomes each. Crossover and mutation probabilities were set to 60% and 1%, respectively. Moreover, the algorithm was repeated three times, starting from different random initial populations. The best solution (in terms of the fitness value) resulting from the three realisations of the GA was employed.

The calculation of classification quality parameters is a recommended standard practice for test evaluation [14]. For this study, measures of test accuracy, such as sensitivity (proportion of positive samples correctly identified), specificity (proportion of negative samples correctly identified) and F-score (measurement of the model accuracy), were utilised. These quality metrics were calculated as follows:

$$\text{Sensitivity (\%)} = \frac{\mathbf{TP}}{\mathbf{TP+FN}} \times \mathbf{100} \quad (3)$$

$$\text{Specificity (\%)} = \frac{\mathbf{TN}}{\mathbf{TN+FP}} \times \mathbf{100} \quad (4)$$

$$\text{F-score} = \frac{\mathbf{2 \times SENS \times SPEC}}{\mathbf{SENS+SPEC}} \quad (5)$$

where TP stands for true positives, TN for true negatives, FP for false positives and FN for false negatives. SENS stands for sensitivity and SPEC for specificity.

All selected wavenumbers obtained from SPA-QDA and GA-QDA for all oesophageal stages of disease (*i.e.*, normal *vs.* inflammatory *vs.* Barrett´s *vs.* LGD *vs.* HGD *vs.* and OAC) were confirmed by a Student's *t*-test (two-tailed, 95% confidence interval).

3 RESULTS

3.1 Plasma Dataset

Fig. 1A shows the average raw Raman spectra derived from blood plasma for all groups (Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC), respectively. Raman spectra of all oesophagus sample stages were compared after cosmic rays and baseline correction. The shape and trend of the six groups were very similar: there are four main Raman peaks at the position at around 1004 cm^{-1}, 1335 cm^{-1}, 1450 cm^{-1} and 1660 cm^{-1} after fluorescence background removal (Figure 1B). Electron-rich groups (*e.g.*, C═O, C═N, and C═C) are the major source of features in Raman spectroscopy [15], and many Raman peaks are caused by the same molecular functional group belonging to different biomolecules in the material [16]. However, there are still some visible spectral differences, notably the bands at

1004 cm^{-1} (collagen), 1335 cm^{-1} (CH_3CH_2 wagging mode of collagen), 1450 cm^{-1} (methylene deformation) and 1660 cm^{-1} (Amide I). Most of these bands decrease in amplitude from the normal group to the OAC group. The difference between groups can be observed more clearly from the averaged pre-processed spectrum of each group (Figure 1B). After pre-processing of the spectral data, chemometric techniques (PCA-QDA, SPA-QDA and GA-QDA) were applied to systematically classify all groups based on their Raman spectra.

The classification of the six oesophageal stages was developed by discriminant analysis using the Raman spectra between 800 and 1800 cm^{-1}. The GA-QDA model with only 16 variables (namely, 884 cm^{-1}, 1188 cm^{-1}, 1206 cm^{-1}, 1235 cm^{-1}, 1296 cm^{-1}, 1307 cm^{-1}, 1365 cm^{-1}, 1383 cm^{-1}, 1402 cm^{-1}, 1440 cm^{-1}, 1461 cm^{-1}, 1608 cm^{-1}, 1641 cm^{-1}, 1656 cm^{-1}, 1715 cm^{-1}, and 1793 cm^{-1}; Fig. 1C and 1D) was found to give the highest classification accuracy in comparison with the other methods (PCA-QDA and SPA-QDA). The classification rate in the test set for Inflammatory and Barrett's samples using GA-QDA was equal to 100%. For the other classes, GA-QDA achieved accuracies, sensitivities and specificities above >85%. The PCA-QDA model using four PC scores (90% of the variance for all classes) achieved 100% accuracy, sensitivity, specificity and F-scores for LGD and OAC classes (Table 2). SPA-QDA also achieved a considerable high accuracy in classification of HGD (100%) when applied using 30 selected wavenumbers, as shown in Table 2. Table S1 lists the selected wavenumbers obtained with the GA-QDA model applied to the plasma samples along their respective tentative biomolecular assignments.

3.2 Saliva Dataset

Fig. 2A shows the average raw Raman spectra derived from saliva for all groups (Normal vs. Inflammatory vs. Barrett's vs. LGD vs. HGD vs. OAC). Raman spectra of all stages of oesophageal disease were compared after cosmic rays and baseline correction. There are three main Raman peaks evident with the shape and trend similar for all groups: 1331 cm^{-1}, 1413 cm^{-1} and 1556 cm^{-1} in the raw spectra after fluorescence background removal

(Figure 2B). In particular, strong peaks observed in the pre-processed spectra at 1336 cm^{-1} and 1664 cm^{-1} indicate Amide III and CH2 wagging vibrations from glycine backbone and Amide I, respectively. These peaks are inherent to Raman spectra of saliva [17]. Peaks at 852 cm^{-1} and 1128 cm^{-1} correspond to C-N stretching, CH3 rocking and C-O vibrations, respectively.

The classification of the six oesophageal stages was developed by discriminant analysis using the Raman pre-processed spectra between 800 and 1800 cm^{-1}. The correct classification for the test set using GA-QDA was equal to 100% for all groups based on only 16 selected wavenumbers (namely, 804 cm^{-1}, 848 cm^{-1}, 873 cm^{-1}, 943 cm^{-1}, 1012 cm^{-1}, 1020 cm^{-1}, 1091 cm^{-1}, 1163 cm^{-1}, 1198 cm^{-1}, 1326 cm^{-1}, 1397 cm^{-1}, 1404 cm^{-1}, 1453 cm^{-1}, 1528 cm^{-1}, 1552 cm^{-1} and 1765 cm^{-1}), as shown in Figure 2C. An excellent classification by the GA-QDA model for saliva was achieved (only 3 errors in the training set and 5 errors in the validation set) (Figure 2D). The PCA-QDA model using seven PC scores (90% of the variance for all classes) achieved good results specifically for inflammatory and HGD groups as can be seen in Table 3. In the SPA-QDA model, a considerable high accuracy in classification of OAC (100%) has been achieved using 30 selected wavenumbers (Table 3). Table S2 lists the selected wavenumbers obtained by GA-QDA for saliva samples with their tentative biomolecular assignments.

3.3 Serum Dataset

Fig. 3A shows the average raw Raman spectra derived from serum for all groups (Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC). There are three main Raman peaks with the shape and trend similar for all groups: 1327 cm^{-1}, 1443 cm^{-1} and 1662 cm^{-1} in the raw spectra after fluorescence background removal (Figure 3B). In particular, strong peaks observed in the pre-processed spectra at 1004, 1337, 1450 and 1657 cm^{-1} indicate phenylalanine (proteins), Amide III, CH_2 bending and Amide I, respectively. The PCA-QDA model using seven PC scores (95% of the variance for all classes) achieved good

results especially for inflammatory, HGD and OAC groups as can be seen in Table 4. Inr the SPA-QDA model, a considerable high accuracy in the classification of OAC (100%) was achieved using 30 selected wavenumbers. The correct classification for the test set using GA-QDA was 100% for normal and HGD groups based on only 30 selected wavenumbers (namely, 821 cm^{-1}, 842 cm^{-1}, 894 cm^{-1}, 962 cm^{-1}, 989 cm^{-1}, 1035 cm^{-1}, 1042 cm^{-1}, 1047 cm^{-1}, 1062 cm^{-1}, 1132 cm^{-1}, 1146 cm^{-1}, 1162 cm^{-1}, 1165 cm^{-1}, 1249 cm^{-1}, 1279 cm^{-1}, 1282 cm^{-1}, 1321 cm^{-1}, 1362 cm^{-1}, 1402 cm^{-1}, 1414 cm^{-1}, 1415 cm^{-1}, 1450 cm^{-1}, 1471 cm^{-1}, 1550 cm^{-1}, 1688 cm^{-1}, 1711 cm^{-1}, 1726 cm^{-1}, 1727 cm^{-1}, 1731 cm^{-1} and 1789 cm^{-1}), as shown in Figure 3C. The selected wavenumbers by GA-QDA for serum with their respective tentative assignment are listed in Table S3.

3.4 Urine Dataset

Fig. 4A shows the average raw Raman spectra derived from urine for all groups (Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC). The Raman spectra of all oesophageal disease stages were analysed after cosmic rays and baseline correction. There are two main Raman peaks with the shape and trend similar for all groups: 1012 cm^{-1} and 1340 cm^{-1} in the raw spectra after fluorescence background removal (Figure 4B). In particular, strong peaks are observed in the pre-processed spectra at 1012 cm^{-1} and 1336 cm^{-1} indicating C-O stretching in ribose and polynucleotide chain (DNA purine bases), respectively. Classification of the six oesophageal stages was developed by discriminant analysis using the Raman spectra between 800 and 1800 cm^{-1}. The correct classification for the test set using GA-LDA was equal to 100% for all groups based on only 29 selected wavenumbers (namely, 845 cm^{-1}, 849 cm^{-1}, 858 cm^{-1}, 864 cm^{-1}, 877 cm^{-1}, 997 cm^{-1}, 1051 cm^{-1}, 1089 cm^{-1}, 1186 cm^{-1}, 1230 cm^{-1}, 1231 cm^{-1},, 1248 cm^{-1},, 1320 cm^{-1}, 1348 cm^{-1}, 1374 cm^{-1}, 1481 cm^{-1}, 1565 cm^{-1}, 1580 cm^{-1}, 1616 cm^{-1}, 1681 cm^{-1}, 1684 cm^{-1}, 1704 cm^{-1}, 1710 cm^{-1}, 1719 cm^{-1}, 1729 cm^{-1}, 1740 cm^{-1}, 1763 cm^{-1} and 1791 cm^{-1}), as can be seen in Figure 4C. The PCA-QDA model using six PC scores (93% of the variance for all classes) achieved good results specifically for

HGD and OAC groups (Table 3); and the SPA-QDA model achieved a considerable high accuracy (100%) for four classes (normal, inflammatory, Barrett's and HGD) using 30 selected wavenumbers (Table 3). Table S4 lists the selected wavenumbers obtained by GA-QDA for urine plasma samples with their respective tentative biomolecular assignments.

4 DISCUSSION

There have been no studies using Raman vibrational spectroscopy on biofluids in order to identify stages of oesophageal transformation to OAC. This study has demonstrated that Raman spectroscopy coupled multivariate classification techniques (PCA-QDA, SPA-QDA and GA-QDA) on biofluids can be used to identify oesophageal stages of disease to adenocarcinoma with excellent accuracy, sensitivity and specificity. A key finding from this study suggests an alteration of some main biomolecules such as tryptophan and phenylalanine from the control group (patients with normal squamous epithelium) to the adenocarcinoma group. These two substances have been shown to have an anti-cancer effect [18]. We presume that these two chemical components when present in plasma and serum are key factors that counteract with cancer cells; and they may increase oncogenic substances that induce cancer formation. In addition, spectral markers such as C-O stretching of ribose (1012 cm^{-1}), Amide III and CH_2 wagging vibrations from glycine backbone (1336 cm^{-1}), methylene deformation (1450 cm^{-1}), and Amide I (1660 cm^{-1}) were found as key discriminant features in saliva and urine between normal and progressive stages of oesophagus disease until cancer using GA-QDA. Ribose sugars are precursors to biosynthetic pathways generated by the Warburg effect, which is responsible for keeping cancer cells alive by generating energy through glycolysis, where glucose is converted to lactose for energy followed by lactate fermentation even when oxygen is available [19]. Amide III vibrations attributed to β-sheet and α-helix conformation in proteins are highly associated to cancer [20-22]; as well as DNA methylation, which is an enzyme-induced chemical modification to the DNA structure where a methyl group is

covalently bonded to the cytosine base, and abnormalities in this phenomenon are related to carcinogenesis [23]. Amide I is known to be associated to cancer due to alterations in proteins backbone conformation [24].

Berger et al. [25] initially introduced the idea that Raman had potential for the analysis of biofluids. Biofluid assays have numerous advantages including high accessibility as well as low invasiveness. Sample processing is cheap and not laborious. This can be implemented in a clinical setting from routine investigations to intra-operative monitoring. In addition, Raman technology to analyse biofluids is particularly advantageous as samples do not need to be dried prior to analysis, which can change their biochemical composition. Raman spectroscopy has been performed on biofluid assays in a variety of cancer diagnostics. Taleb et al. [8] achieved an overall accuracy rate of 86–91% using multivariate analysis techniques applied to Raman spectral data of serum in hepatocellular carcinoma. Harvey et al. [26] conducted studies using Raman spectroscopy of urine to detect prostate cancer cells. The authors identified that the Raman spectra suggested a higher concentration of nucleic acids and proteins in bladder cells compared to the prostate cancer cells. An early work by Chan et al. [27] focused on analysing white blood cells as a potential diagnostic tool for haematological malignancies such as Lymphoma and Leukaemia. The authors showed that single-cell Raman micro-spectroscopy was able to discriminate between normal human lymphocytes and transformed Jurkat and Raji lymphocyte cell lines based on highly reproducible biomolecular fingerprints [27]. Multivariate statistical models based on the Raman spectra achieved a sensitivity of 98.3% for cancer detection, with 97.2% of the cells being correctly classified as belonging to the normal or transformed group [27].

Herein, the pre-processed Raman spectral datasets were analysed by classification methods based on QDA. QDA is a discriminant analysis algorithm based on a Mahalanobis distance calculation that uses a separate variance-covariance matrix for each class [28]. This increases the discrimination accuracy in complex biological medium where classes having

different variance structures are present [28]. PCA-QDA, SPA-QDA and GA-QDA were used on the six groups of samples for all biofluids separately. GA-QDA was found to be very effective in the discrimination between all groups, since this algorithm achieved a high-quality performance rate using few wavenumbers. The GA-QDA model obtained an accuracy of 100% for saliva and urine in all groups. Although the number of samples in LGD and HGD groups are still small, which is a limitation towards the predictive ability of this classifier, the overall results demonstrate promising evidence that Raman spectroscopy coupled with chemometric techniques can be used for distinguish different stages of oesophagus disease to adenocarcinoma.

In order to have a robust surveillance program for the early detection and improved detection rate for OAC, skilled endoscopists and pathologists are necessary for accurate mucosal sampling and histopathologic examination. However, the development of a quick, convenient, and inexpensive method for detecting early cancer or different stages from normal squamous epithelium through to OAC can be useful specifically to guide tissue biopsy thus increasing the yield of dysplasia detection. By using liquid samples the analysis would even be faster since no drying time would be necessary, and the level of discrimination could improve since relevant compounds might evaporate during the drying process. In addition, by using specially prepared slides for surface enhanced Raman spectroscopy the sensitivity of this method could be improved further due to the large magnification of the Raman signal for these samples [29]. This study shows the potential of Raman spectroscopy and chemometrics for detecting oesophageal stages of disease through to OAC based on biofluids with high accuracy, sensitivity and specificity.

5 CONCLUSION

Raman spectroscopy is a highly informative, non-destructive and robust technique that has been limitedly employed in the field of oesophageal disease. The results of this study

show that Raman spectroscopy coupled with multivariate classification algorithms (PCA-QDA, SPA-QDA and GA-QDA) result in a powerful alternative approach for detection of oesophageal stages of disease to OAC in biofluids with an excellent accuracy, specificity and sensitivity in saliva and urine. We present a new, rational and convenient approach to different biofluids (plasma, saliva, serum and urine) using a minimal sample volume and without special sample preparation and reagents. This pilot study is pioneer; but further work in this field including larger cohorts of oesophageal malignancies samples should be performed in the future to validate these encouraging results before clinical trials.

ACKNOWLEDGMENTS

This study was funded by Rosemere Cancer Foundation and the University of Central Lancashire start-up fund, UK. C.L.M.M. would like to thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Brazil (grant 88881.128982/2016-01) for financial support. Professor Kássio M.G. Lima would like to thank UFRN (process number: 23077.019698/2019-67) for his training license.

AUTHOR CONTRIBUTIONS

I.M. performed the sample collection, spectral data acquisition and analysis, and wrote the first manuscript draft. C.L.M.M. and K.M.G.L. performed the multivariate data analysis and aided writing the manuscript. K.M.A. helped with sample collection and preparation. F.L.M. conceived the study design. R.S.D. and F.L.M. supervised the project as principal investigators and finalised the manuscript.

CONFLICT OF INTEREST

The authors declare no financial or commercial conflict of interest.

REFERENCES

[1] S. Bhat, H. G. Coleman, F. Yousef, B. T. Johnston, D. T. McManus, A. T. Gavin, L. J. Murray. *J. Natl. Cancer Inst.* **2011**, *103*, 1049.

[2] A. Bhardwaj, D. B. Stairs, H. Mani, T. J. McGarrity. *Patholog. Res. Int.* **2012**, *814146*, 1.

[3] R. Milind, S. E. Attwood. *World J Gastroenterol.* **2012**, *18*, 3483.

[4] M. J. Baker, S. R. Hussain, L. Lovergne, V. Untereiner, C. Hughes, R. A. Lukaszewski, G. Thiéfin, G. D. Sockalingum. *Chem. Soc. Rev.* **2016**, *45*, 1803.

[5] H. J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, K. Esmonde-White, N. J. Fullwood, B. Gardner, P. L. Martin-Hirsch, M. J. Walsh, M. R. McAinsh, N. Stone, F. L. Martin. *Nat. Protoc.* **2016**, *11*, 664.

[6] F. S. Parker. *Applications of infrared, Raman and Resonance Raman spectroscopy in biochemistry*, vol. 9, Springer, New York, **1983**.

[7] X. Li, T. Yang, S. Li. *Appl. Opt.* **2012**, *51*, 5038.

[8] I. Taleb, G. Thiéfin, C. Gobinet, V. Untereiner, B. Bernard-Chabert, A. Heurgué, C. Truntzer, P. Hillon, M. Manfait, P. Ducoroy, G. D. Sockalingum. *Analyst* **2013**, *138*, 4006.

[9] J. L. González-Solís, J. C. Martínez-Espinosa, L. A. Torres-González, A. Aguilar-Lemarroy, L. F. Jave-Suárez, P. Palomares-Anda. *Lasers Med. Sci.* **2014**, *29*, 979.

[10] J. L. Pichardo-Molina, C. Frausto-Reyes, O. Barbosa-García, R. Huerta-Franco, J. L. González-Trujillo, C. A. Ramírez-Alvarado, G. Gutiérrez-Juárez, C. Medina-Gutiérrez. *Lasers Med. Sci.* **2007**, *22*, 229.

[11] A. Sahu, S. Sawant, H. Mamgain, C. M. Krishna. *Analyst* **2013**, *138*, 4161.

[12] P. H. C. Eilers. *Anal. Chem.* **2004**, *76*, 404.

[13] R. Kennard, L. Stone. *Technometrics* **1969**, *11*, 137.

[14] K. T. Cheung, J. Trevisan, J. G. Kelly, K. M. Ashton, H. F. Stringfellow, S. E. Taylor,

M. N. Singh, P. L. Martin-Hirsch, F. L. Martin. *Analyst* **2011**, *136*, 2047.

[15] Z. Movasaghi, S. Rehman, I. U. Rehman. *Appl. Spectrosc. Rev.* **2007**, *42*, 493.

[16] S. Stewart, R. J. Priore, M. P. Nelson, P. J. Treado. *Annu. Rev. Anal. Chem.* **2012**, *5*, 337.

[17] K. Virkler, I. K. Lednev. *Analyst* **2010**, *135*, 512.

[18] J. A. Norton, C. M. Gorschboth, R. A. Wesley, M. E. Burt, M. F. Brennan. *Cancer* **1985**, *56*, 1181.

[19] A. Fadaka, B. Ajiboye, O. Ojo, O. Adewale, I. Olayide, R. Emuowhochere. *Journal of Oncology Sciences* **2017**, *3*, 45.

[20] H. Abramczyk, A. Imiela. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2018**, *188*, 8.

[21] A. C. S. Talari, S. Rehman, I. U. Rehman. *Expert. Rev. Mol. Diagn.* **2019**, *19*, 929.

[22] G. W. Auner, S. K. Koya, C. Huang, B. Broadbent, M. Trexler, Z. Auner, A. Elias, K. C. Mehne, M. A. Brusatori. *Cancer Metastasis Rev.* **2018**, *37*, 691.

[23] S. A. Wajed, P. W. Laird, T. R. DeMeester. *Ann. Surg.* **2001**, *234*, 10.

[24] N. Kuhar, S. Sil, T. Verma, S. Umapathy. *RSC Adv.* **2018**, *8*, 25888.

[25] A. J. Berger, T. W. Koo, I. Itzkan, G. Horowitz, M. S. Feld. *Appl Opt.* **1999**, *38*, 2916.

[26] T. J. Harvey, E. C. Faria, A. Henderson, E. Gazi, A. D. Ward, N. W. Clarke, M. D. Brown, R. D. Snook, P. Gardner. *J. Biomed. Opt.* **2008**, *13*, 234.

[27] J. W. Chan, D. S. Taylor, T. Zwerdling, S. M. Lane, K. Ihara, T. Huser. *Biophys. J.* **2006**, *90*, 648.

[28] C. L. M. Morais, K. M. G. Lima. *J. Braz. Chem. Soc.* **2018**, *29*, 472.

[29] D. Graham, K. Faulds. *Expert. Rev. Mol. Diagn.* **2009**, *9*, 537.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Table S1: Category-distinguishing wavenumbers for Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC using plasma samples obtained for SPA-QDA and GA-QDA models.

Table S2: Category-distinguishing wavenumbers for Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC using saliva samples obtained for SPA-QDA and GA-QDA models.

Table S3: Category-distinguishing wavenumbers for Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC using serum samples obtained for SPA-QDA and GA-QDA models.

Table S4: Category-distinguishing wavenumbers for Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC using urine samples obtained for SPA-QDA and GA-QDA models.

FIGURES AND TABLES

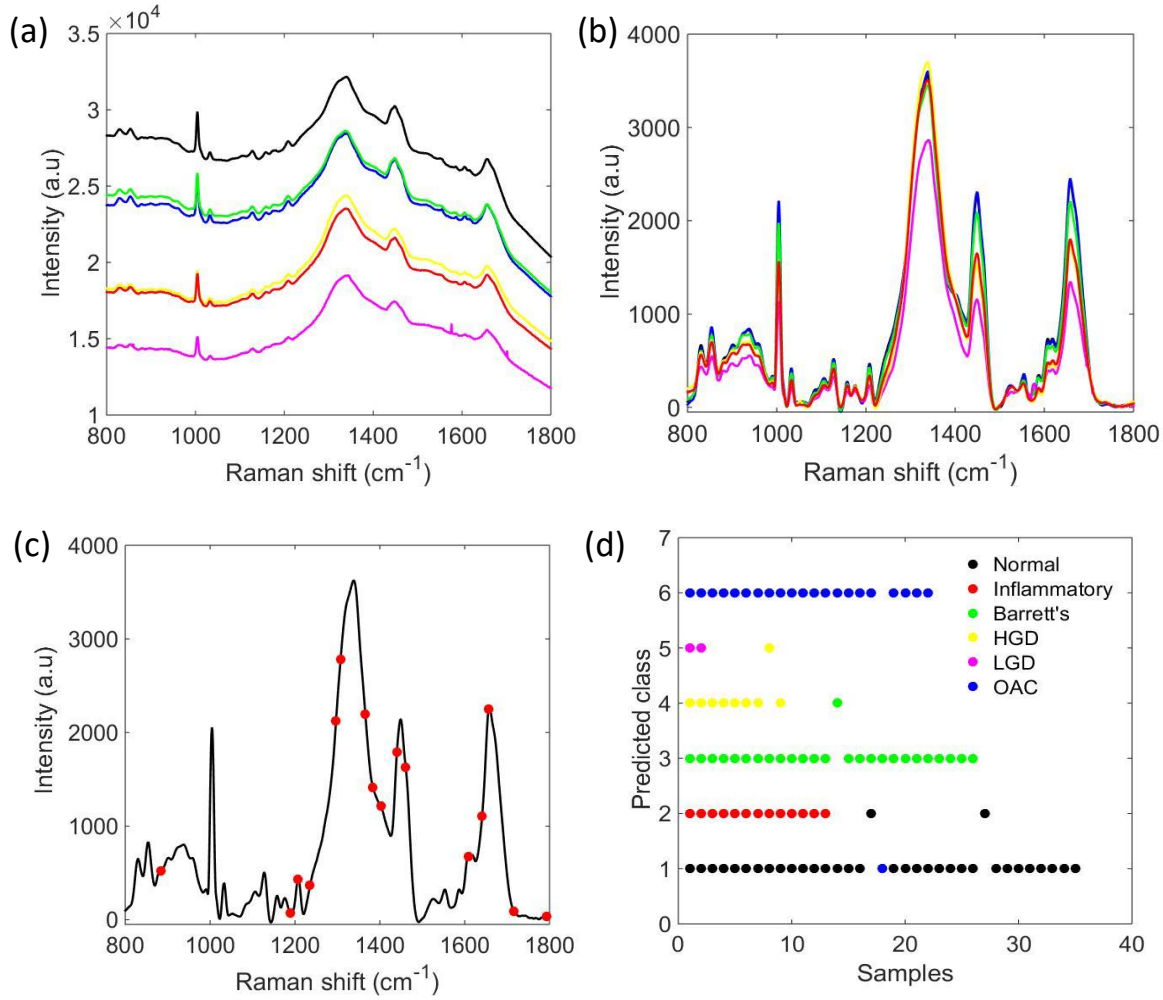


FIGURE 1. Comparison of Normal/Inflammatory/Barrett's/LGD/HGD/OAC oesophageal stages using plasma samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1}; (B) Average pre-processed Raman spectrum obtained from all stages segregated into Normal (black colour) *vs.* Inflammatory (blue colour) *vs.* Barrett's (green colour) *vs.* LGD (yellow colour) *vs.* HGD (magenta colour) *vs.* OAC (red colour); (C) 16 selected variables used by the GA-QDA model; (D) Samples *vs.* predicted class for training and test sets. The spectrum colour for each class in (A) and (B) are the same depicted in (D) inset.

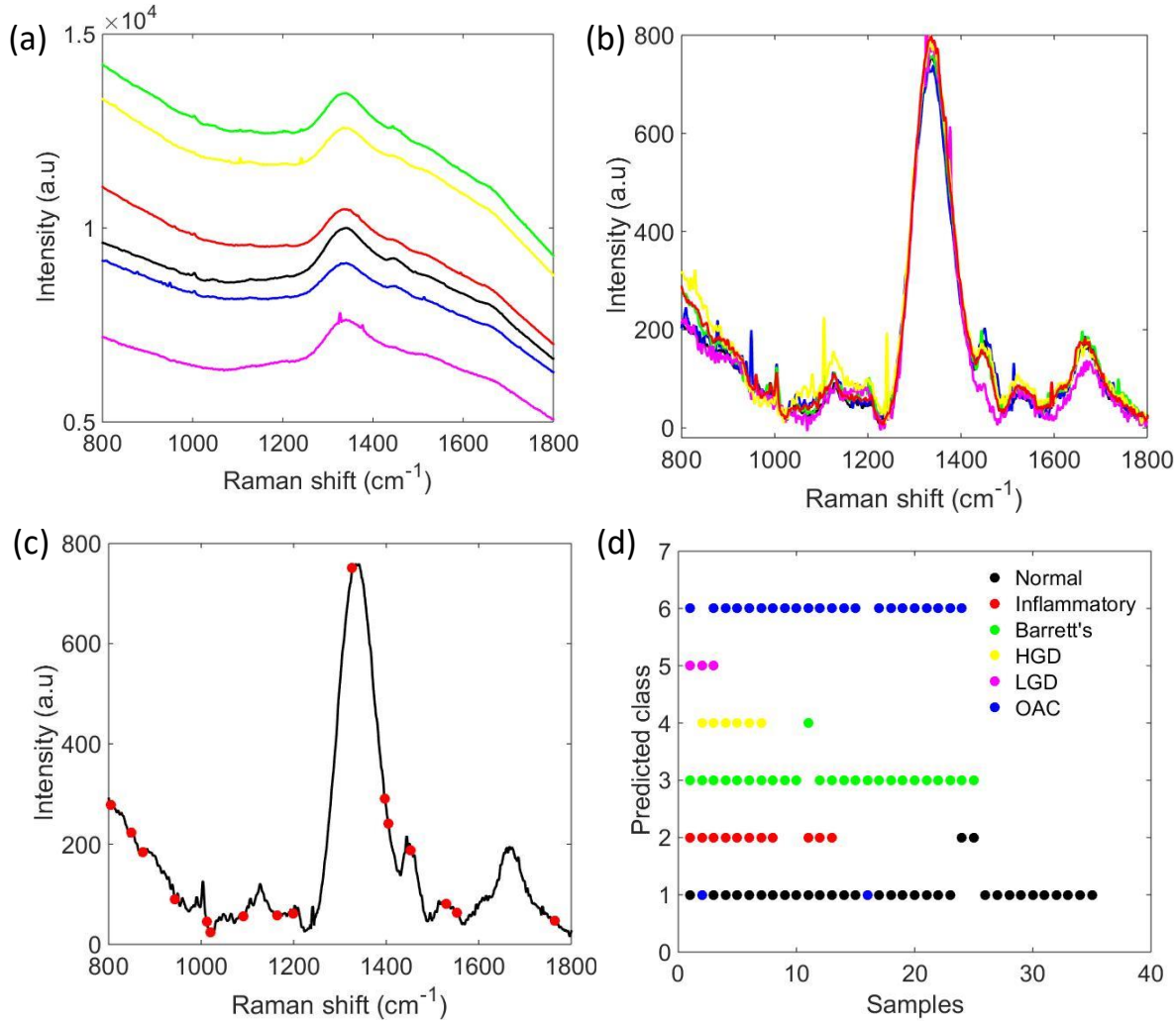


FIGURE 2. Comparison of Normal/Inflammatory/Barrett's/LGD/HGD/OAC oesophageal stages using saliva samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1}; (B) Average pre-processed Raman spectrum obtained from all stages segregated into Normal (black colour) *vs.* Inflammatory (blue colour) *vs.* Barrett's (green colour) *vs.* LGD (yellow colour) *vs.* HGD (magenta colour) *vs.* OAC (red colour); (C) 16 selected variables used by the GA-QDA model; (D) Samples *vs.* predicted class for training and test sets. The spectrum colour for each class in (A) and (B) are the same depicted in (D) inset.

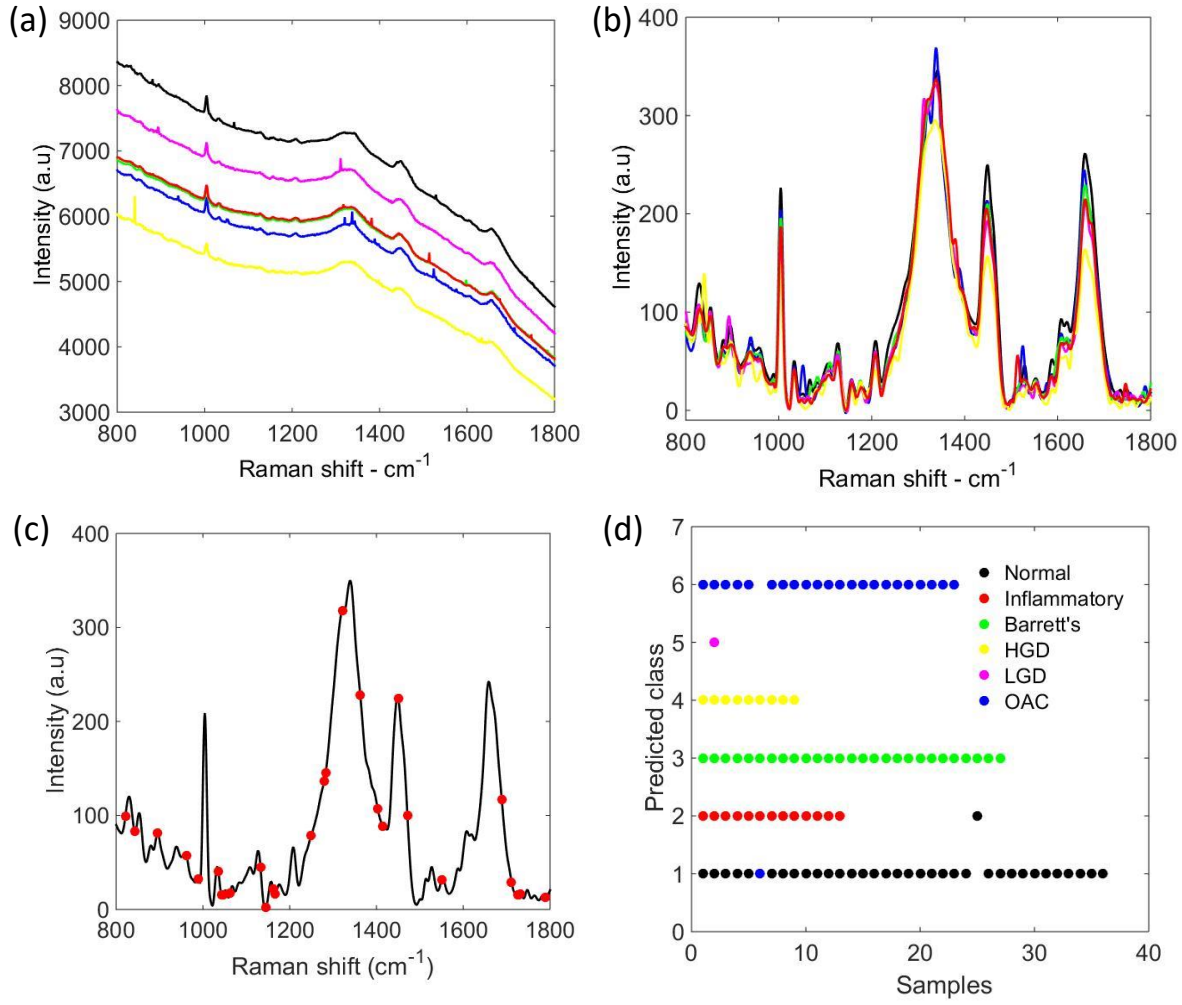


FIGURE 3. Comparison of Normal/Inflammatory/Barrett's/LGD/HGD/OAC oesophageal stages using serum samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1}; (B) Average pre-processed Raman spectrum obtained from all stages segregated into Normal (black colour) *vs.* Inflammatory (blue colour) *vs.* Barrett's (green colour) *vs.* LGD (yellow colour) *vs.* HGD (magenta colour) *vs.* OAC (red colour); (C) 30 selected variables used by the GA-QDA model; (D) Samples *vs.* predicted class for training and test sets. The spectrum colour for each class in (A) and (B) are the same depicted in (D) inset.

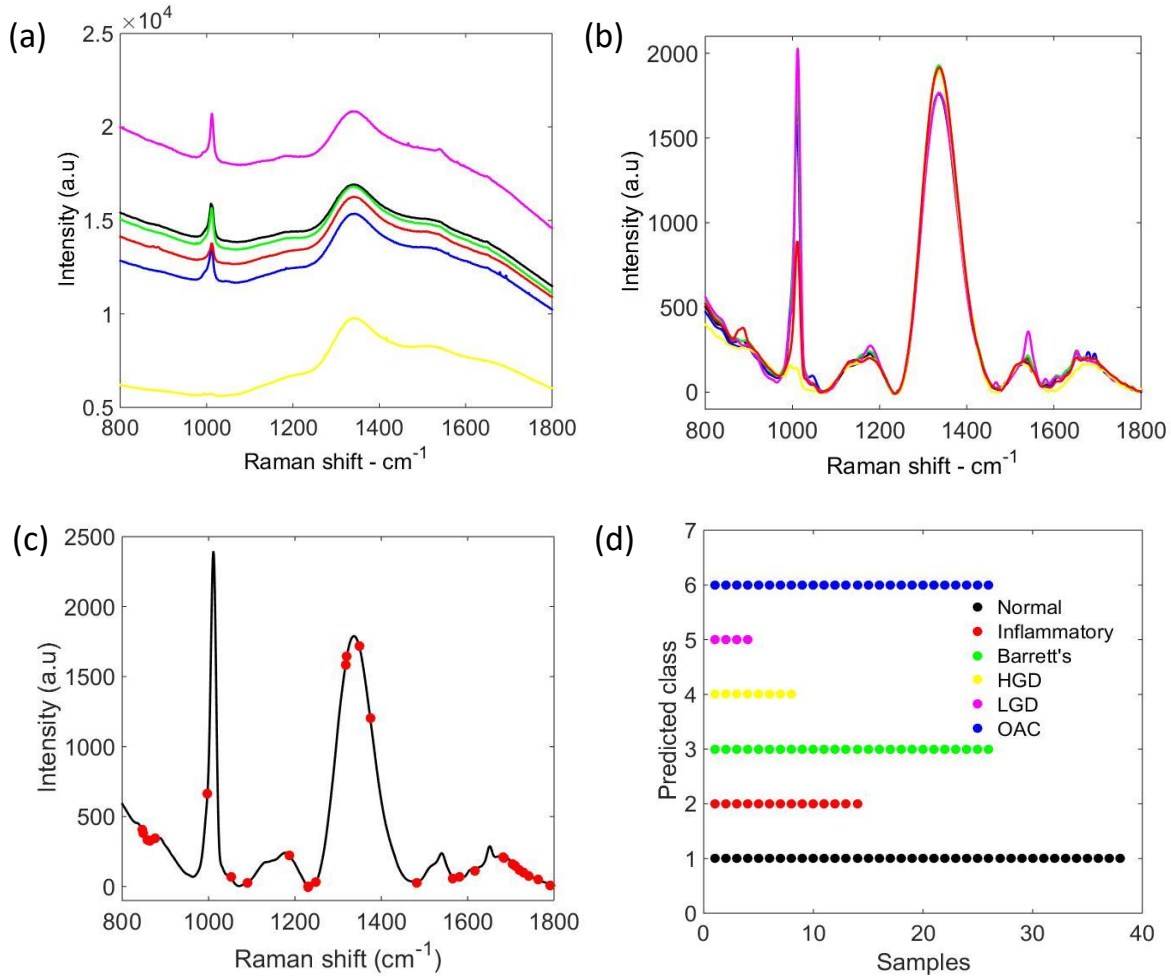


FIGURE 4. Comparison of Normal/Inflammatory/Barrett's/LGD/HGD/OAC oesophageal stages using urine samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1}; (B) Average pre-processed Raman spectrum obtained from all stages segregated into Normal (black colour) *vs.* Inflammatory (blue colour) *vs.* Barrett's (green colour) *vs.* LGD (yellow colour) *vs.* HGD (magenta colour) *vs.* OAC (red colour); (C) 29 selected variables used by the GA-QDA model; (D) Samples *vs.* predicted class for training and test sets. The spectrum colour for each class in (A) and (B) are the same depicted in (D) inset.

TABLE 1. Number of training, validation and prediction samples in each category of biofluids.

Category – biofluids	Training	Validation	Test
Normal – plasma	21	7	7
Normal – serum	22	7	7
Normal – saliva	21	7	7
Normal – urine	23	7	8
Inflammatory – plasma	11	3	4
Inflammatory – serum	12	3	4
Inflammatory – saliva	11	3	4
Inflammatory – urine	12	3	4
Barrett's – plasma	16	5	6
Barrett's – serum	17	5	6
Barrett's – saliva	16	5	5
Barrett's – urine	17	5	5
LGD – plasma	3	1	2
LGD – serum	3	1	2
LGD – saliva	2	1	2
LGD – urine	3	1	2
HGD – plasma	7	2	3
HGD – serum	7	2	3
HGD – saliva	7	2	3
HGD – urine	7	3	3
OAC – plasma	13	4	5
OAC – serum	13	5	5
OAC – saliva	14	5	5
OAC – urine	15	5	6

TABLE 2. Figures of merit (accuracy, sensitivity, specificity and F-scores) for Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC using plasma samples.

PCA-QDA	**Oesophageal stages**					
	Normal	*Inflammatory*	*Barrett's*	*LGD*	*HGD*	*OAC*
Accuracy (%)	91.3	91.3	95.6	100	95.6	100
Sensitivity (%)	100	90	100	100	95.2	100
Specificity (%)	71.4	100	80	100	100	100
F-Scores (%)	83.3	94.7	88.8	100	97.5	100
SPA-QDA	**Oesophageal stages**					
	Normal	*Inflammatory*	*Barrett's*	*LGD*	*HGD*	*OAC*
Accuracy (%)	82.6	86.9	95.6	95.6	100	86.4
Sensitivity (%)	87.5	90	94.4	100	100	94.4
Specificity (%)	71.4	66.6	100	100	100	60
F-Scores (%)	78.6	76.6	97.1	73.3	100	73.3
GA-QDA	**Oesophageal stages**					
	Normal	*Inflammatory*	*Barrett's*	*LGD*	*HGD*	*OAC*
Accuracy (%)	95.2	100	100	91.3	95.6	91.3
Sensitivity (%)	93.7	100	100	95.4	100	94.4
Specificity (%)	100	100	100	100	70	80
F-Scores (%)	96.7	100	100	83.3	66.6	86.6

TABLE 3. Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC using saliva samples.

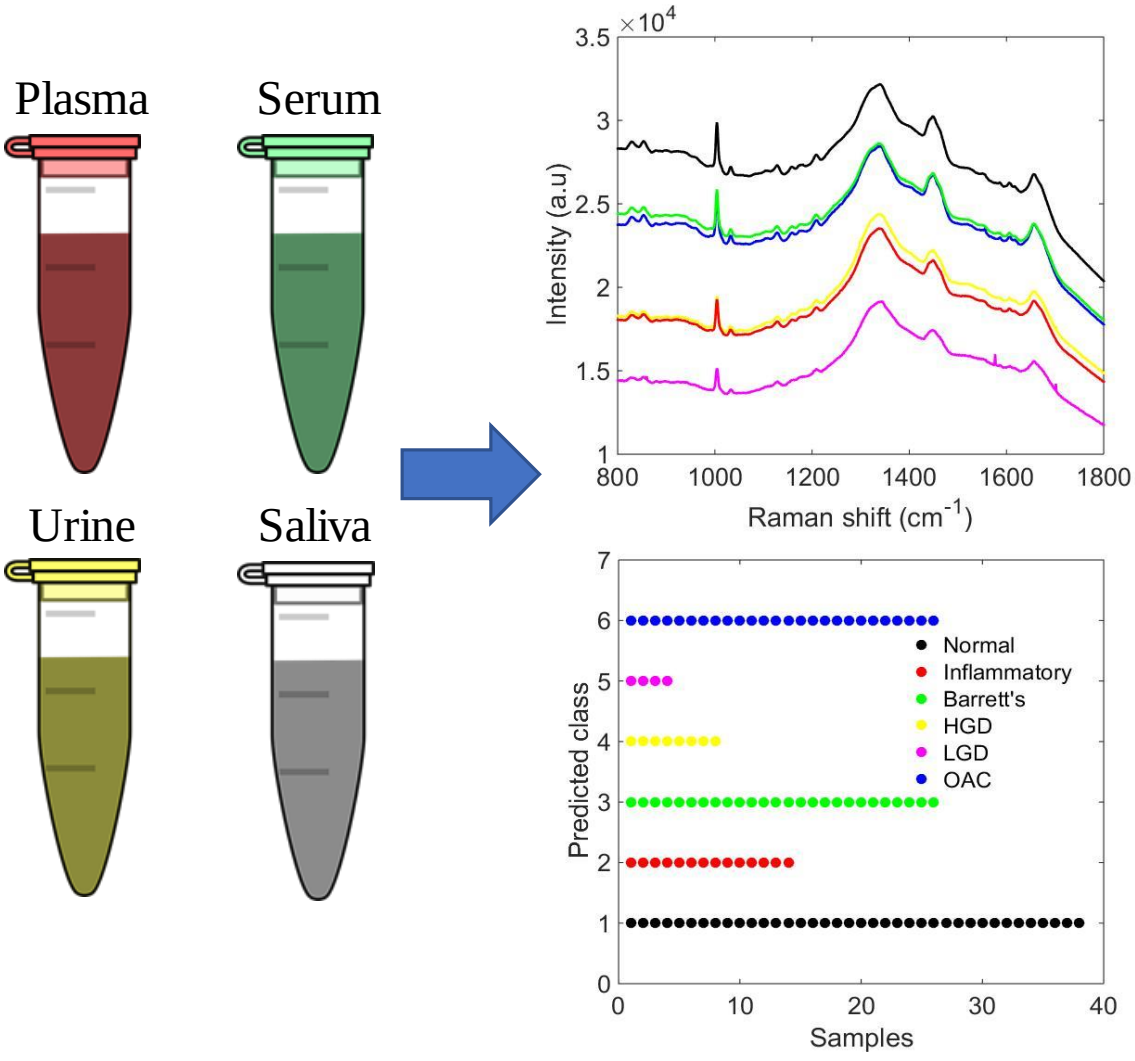
PCA-QDA	**Oesophageal stages**					
	Normal	*Inflammatory*	*Barrett's*	*LGD*	*HGD*	*OAC*
Accuracy (%)	82.6	91.3	95.6	95.6	95.6	86.9
Sensitivity (%)	87.5	90	100	100	95.2	94.4
Specificity (%)	71.4	100	80	0	100	60
F-Scores (%)	78.6	94.7	88.8	0	97.5	73.3
SPA-QDA	**Oesophageal stages**					
	Normal	*Inflammatory*	*Barrett's*	*LGD*	*HGD*	*OAC*
Accuracy (%)	95.6	95.6	95.6	95.6	91.3	100
Sensitivity (%)	100	95	100	95.4	95.2	100
Specificity (%)	85.7	100	80	100	50	100
F-Scores (%)	92.3	97.4	88.8	97.6	65.5	100

TABLE 4. Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC using serum samples.

PCA-QDA	**Oesophageal stages**					
	Normal	*Inflammatory*	*Barrett's*	*LGD*	*HGD*	*OAC*
Accuracy (%)	86.9	91.3	91.3	95.6	91.3	91.3
Sensitivity (%)	93.7	90.0	100	100	90.4	94.4
Specificity (%)	71.4	100	60	0	100	80.0
F-Scores (%)	81.0	94.7	75	0	95	86.6
SPA-QDA	**Oesophageal stages**					
	Normal	*Inflammatory*	*Barrett's*	*LGD*	*HGD*	*OAC*
Accuracy (%)	95.6	95.6	91.3	95.6	86.9	100
Sensitivity (%)	100	95.0	100	95.4	90.4	100
Specificity (%)	85.7	100	60	100	50	100
F-Scores (%)	92.3	97.4	75	97.6	64.4	100
GA-QDA	**Oesophageal stages**					
	Normal	*Inflammatory*	*Barrett's*	*LGD*	*HGD*	*OAC*
Accuracy (%)	100	95.6	95.6	95.6	100	95.6
Sensitivity (%)	100	100	94.4	100	100	94.4
Specificity (%)	100	66.6	100	60	100	100
F-Scores (%)	100	80.0	97.1	70	100	97.2

TABLE 5. Figures of merit (accuracy, sensitivity, specificity and F-scores) for Normal *vs.* Inflammatory *vs.* Barrett's *vs.* LGD *vs.* HGD *vs.* OAC using urine samples.

PCA-QDA	**Oesophageal stages**					
	Normal	*Inflammatory*	*Barrett's*	*LGD*	*HGD*	*OAC*
Accuracy (%)	62.5	66.7	83.3	95.8	87.5	87.5
Sensitivity (%)	87.5	66.7	94.7	100	86.3	94.7
Specificity (%)	12.5	66.7	40	0	100	60
F-Scores (%)	21.8	66.7	56.2	0	92.7	73.4
SPA-QDA	**Oesophageal stages**					
	Normal	*Inflammatory*	*Barrett's*	*LGD*	*HGD*	*OAC*
Accuracy (%)	100	100	100	95.8	100	95.8
Sensitivity (%)	100	100	100	100	100	94.7
Specificity (%)	100	100	100	50	100	100
F-Scores (%)	100	100	100	50	100	97.3

Graphical Abstract

Four biofluids (plasma, serum, urine and saliva) were measured through Raman spectroscopy in order to detect oesophageal transformation stages to adenocarcinoma.