



Scientific Paper

Evaluation of changes caused by HR-HPV infection in squamous cell carcinoma of the head and neck using Raman microspectroscopy in combination with multivariate statistical analysis

Agnieszka SZURKO ^{1,ACDEF,*}, Dorota ZYGADŁO ^{1,ABCDEF}, Sebastian STUDENT ^{2,3,CDEF},
Miroslaw ŚNIETURA ^{4,ABDEF}, Zbigniew LORENC ^{5,EF}, Agnieszka KIEŁBOŃ ^{1,DEF},
Dariusz WANICZEK ^{6,ABCDEF}

¹Faculty of Science and Technology, University of Silesia, 75 Pułku Piechoty 1A, 41-500, Chorzów, Poland

²Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, Gliwice, Poland

³Biotechnology Center, Silesian University of Technology, Poland

⁴Department of Pathomorphology and Molecular Diagnostics, Faculty of Medical Sciences in Katowice, Medical University of Silesia, Katowice, Poland

⁵Department of General, Colorectal and Polytrauma Surgery, Faculty of Health Sciences in Katowice, Medical University of Silesia, Katowice, Poland

⁶Department of Oncological Surgery, Faculty of Medical Sciences in Zabrze, Medical University of Silesia, Katowice, Poland

*Corresponding author: Agnieszka Szurko; agnieszka.szurko@us.edu.pl

(received 24 August 2023; revised 16 January 2024; accepted 21 April 2024)

Abstract

Introduction: Squamous cell carcinoma of the head and neck region (HNSCC), with a positive status for high oncogenic potential human papillomavirus (HR-HPV), represents a clinically distinct disease entity compared to HPV-independent cases. Patients exhibit variations in prognosis and proposed therapy regimens. A prompt and reliable diagnosis of the presence of HPV infection could optimize the treatment for these patients. Currently employed treatment methods are long-term, expensive, and lack specificity, especially when administered separately.

Material and methods: The research objective of this study is to explore significant differences in the Raman spectra of biological samples taken from patients with HNSCC, facilitating the confirmation of HPV virus presence. Study groups were delineated based on histopathological diagnosis and molecular biology tests, confirming the biological activity of the virus and the presence of the HR-HPV form with a diagnosis of a specific subtype.

Results: To identify high oncogenic potential human papillomavirus (HR-HPV) infection as a crucial factor in squamous cell carcinoma of the head and neck region, an effective automatic data analysis system was established, relying on Raman microspectroscopy and multivariate analysis. Our results showed clear ranges of the Raman spectrum that differentiated between HPV-associated and non-HPV-associated cancers.

Conclusions: In conclusion, our experience shows a great diagnostic potential of Raman confocal microscopy with multidimensional statistical analysis. In the future, the use of this method may allow for the creation of an effective and automated HR-HPV detection system in neoplastic tissue.

Keywords: Raman spectroscopy; human papilloma virus; multivariate statistical analysis; Squamous cell carcinoma of the head and neck region (HNSCC).

Introduction

Head and neck squamous cell carcinomas (HNSCCs) represent the sixth most prevalent type of cancer worldwide.¹ Risk factors for developing this cancer include smoking and alcohol abuse.^{2,3} Recently, the number of cases of HNSCC caused by infection with high oncogenic potential human papillomavirus (HR-HPV) has significantly increased. Among HNSCCs, a group of neoplasms with distinct molecular genetic changes caused by

HPV infection of high oncogenic potential has been distinguished. This group is characterized by a different clinical course, prognosis, and methods of treatment.⁴⁻⁸ It has been shown that HR-HPV-associated HNSCC is clinically distinct from non-HPV-associated HNSCC. Rapid and reliable diagnosis of patients with neoplastic lesions regarding the presence of HR-HPV infection would allow for treatment optimization.

© 2024 Authors

This is an open access article licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Authors' contribution:

A – Research concept and design, B – Collection and/or assembly of data, C – Data analysis and interpretation, D – Writing the article, E – Critical revision of the article, F – Final approval of the article.

Based on the measurement of scattered radiation (Raman scattering, i.e., inelastic photon scattering), Raman microspectroscopy has been a highly developed research technique in recent years. It is one of the most universal contemporary analytical techniques used for research in medical diagnosis. The position of the bands in the Raman spectrum is constant for a given substance and independent of the frequency of the incident beam, constituting a unique "chemical fingerprint" of the sample. At the same time, RS is a complementary method to Fourier transform infrared spectroscopy (FTIR), which also detects oscillations of chemical groups. The general difference between RS and FTIR concerns fewer restrictions on sample preparation (e.g. it does not require sample drying). However, the sample must be protected from damage before measurement and thus obtain signal amplification for both non-specific and specific substances. Raman spectroscopy is a method sensitive to amino acids, lipids, nucleic acids, as well as disulfides and cytochromes, but not sensitive to water content. However, infrared spectroscopy is sensitive to proteins and phosphates, but also to water.⁹⁻¹² Combining Raman spectroscopy with confocal microscopy offers additional advantages, as confocal microscopy is a type of light microscopy characterized by increased contrast and improved image quality. Classical techniques of multivariate data analysis used in supervised or unsupervised image analysis are applied to process complex image data, allowing for better interpretation and utilization.

The aim of the study was to search for significant differences in Raman spectra obtained for HNSCC using confocal Raman microspectroscopy in combination with multidimensional statistical analysis. Based on the immunohistochemical technique and the qRT-PCR method, the presence of high oncogenic potential human papillomavirus in the HPV-associated HNSCC group was also demonstrated. Therefore, we hope that our research will show that Raman microspectroscopy can be a useful tool, allowing for quick verification and confirmation of the results obtained with classical techniques in the case of HNSCC diagnosis.

Materials and methods

We utilized 11 paraffin blocks prepared from material collected intraoperatively from patients with HNSCC. Two distinct methods were employed concurrently to confirm the presence of high oncogenic potential human papillomavirus (HR-HPV) in the harvested HNSCC tissue. One method involved immunohistochemical staining, where the biological activity of HR-HPV in cells was indirectly confirmed by assessing the level of accumulation of P16 protein (INK4A), considered a surrogate marker of HR-HPV infection.¹³ In the subsequent stage, the genome of the 14 most common types of HR-HPV was directly detected using RT-PCR.

Immunohistochemical staining against P16 (INK4A)

To confirm the biological activity of HPV, we conducted an immunohistochemical assessment of the expression of the viral product, protein p16 (INK4A), in the tumor tissue.¹⁴ The tests were carried out on serial tissue sections using a commercial CINTech histology kit (mtm-Laboratories, Heidelberg, Germany) following the manufacturer's procedure. A uniform, intense staining of at least 80% of neoplastic cells was considered a positive result. Images captured for samples immunohistochemically stained for the presence of p16 protein (INK4A) were digitized using the 3DHitech Flash250 III histological slide scanner (Hungary). Examples of the obtained staining images are presented in **Figure 1**.

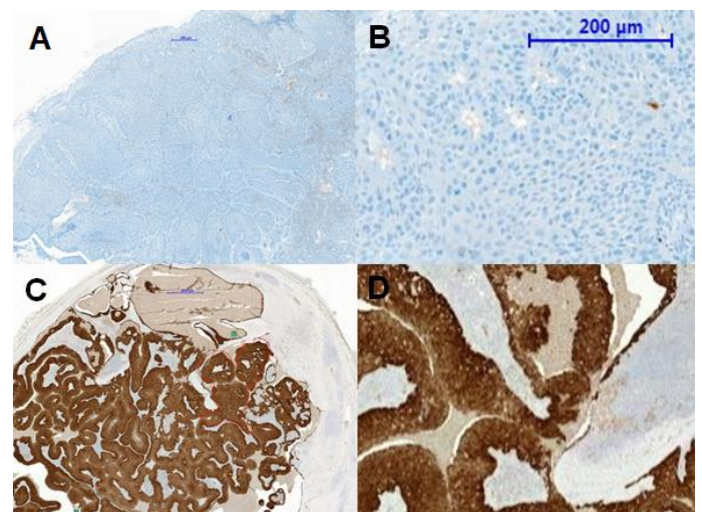


Figure 1. Images of p16 (INK4A) immunohistochemical staining in HPV-negative tumor tissue (a; b), which was the control group and HPV-positive tumor tissue (c; d).

Detection of HPV DNA by RT-PCR

The first step involved DNA isolation from tumor samples, which were formalin-fixed and paraffin-embedded. The procedure for DNA isolation from archival paraffin blocks entailed preparing serial sections of five 10-µm thick slices under sterile conditions. This process was meticulously performed to prevent contamination of the material with other DNA, particularly DNA obtained from blocks that could contain the HPV genome. The prepared tissues were suspended in mineral oil with the addition of a proteinase K solution for the proteolytic digestion of the samples. The isolated DNA was absorbed using magnetic nanospheres of iron oxide to magnetically separate it from impurities. The entire process was automated in the MagCore HF16 Plus device, using the MagCore® Genomic DNA FFPE One-Step Kit (RBC-Bioscience, Taiwan).

The subsequent step involved analyzing the purity and concentration of the obtained DNA using the Nanodrop ND-1000 spectrophotometer (ThermoFisher Scientific, Waltham, USA). The presence of 14 types of HR-HPV viruses (Abbott Molecular, Abbott Park, USA) was assessed by real-time PCR. The test recognized each type of HPV16 and HPV18 separately, as well as less common types of HPV (31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68) in total.¹⁵

For virus type analysis using commercial tests, a reaction mixture containing Taq polymerase in an appropriate buffer medium and a sample dilution containing 10 ng genomic DNA in each case were prepared. Product detection was performed with a ViiA 7 Real Time instrument (ThermoFisher Scientific, Waltham, USA). According to the procedure, negative and positive controls were analyzed to confirm that the reaction was conducted correctly. DNA with the sequence of β -globulin, poly(dA:dT), and HPV16, HPV18, and HPV58 sequences were used as negative and positive controls, respectively.¹⁶ The cut-off point was the CT value of 32. The method's effectiveness in assessing HPV status was demonstrated in previous studies by Śnietura et al.^{17,18} As a result, the group was divided into an HPV-negative control (six patients) and an HPV-positive study group (five patients).

Regions of interest (ROI), corresponding to the tumor tissue, were marked by a histopathologist on a previously scanned virtual slide using Bezier curves and made available remotely for the measurement and analysis of Raman spectra in the serial sections. The spectra obtained from confocal Raman microspectroscopy were compared using multivariate statistical analysis. Images of unstained tissue subjected to Raman spectroscopy measurements and the corresponding stained tissue with the tumor site marked by a histopathologist are depicted in **Figure 2**.

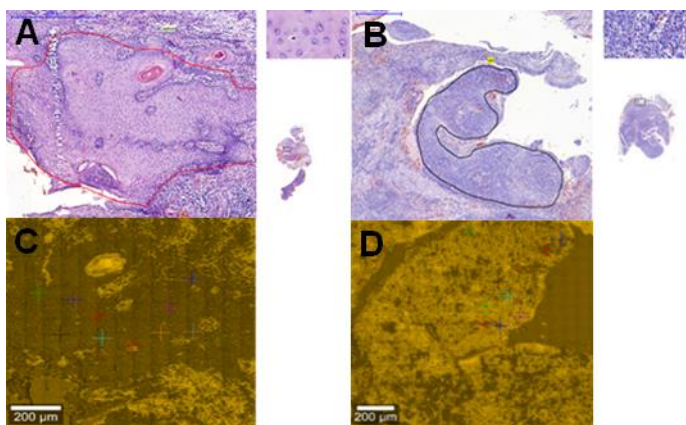


Figure 2. Selected virtual hematoxylin and eosin images with ROI areas with tumor tissue; a) HR HPV-negative; b) HR HPV-positive and the corresponding images of unstained samples tissues recorded in a confocal microscope, in which single spectra were measured (marked with the crosses); c) and d) HR HPV-negative and HR HPV-positive, respectively.

Determination of molecular changes in HNSCC by Raman confocal microscopy

For spectroscopic measurements, HNSCC tissue samples were prepared by dewaxing with xylene followed by hydration in decreasing concentrations of ethanol. All Raman spectra measurements were conducted using an alpha 300 R microscope (WITec, Ulm, Germany). The apparatus comprises a UHTS spectrometer and a CCD camera (cooled to -60°C), coupled with a confocal Zeiss microscope equipped with an automatic mapping table. The laser line with a wavelength of 532 nm (Nd:YAG laser) was employed for measurements by focusing on the sample, with a $50\times$ magnification objective (NA = 0.75). A surface resolution of 650 nm was achieved. The average laser intensity was 10 mW. The spectra were recorded using the WITEC Control software and pre-processed using the WITEC Project Plus. Point measurements were made with the Olympus LMPlanFL 50x NA = 0.76 objective. The integration time of one spectrum was 2 s, and the number of repetitions of a single spectrum in a given place was 20. To create virtual slides of a microscopic image of unstained tissue, the "Image Stitching" function and Olympus MPlan lenses 10x NA = 0.3 (scan of the entire tissue) and Olympus 50x NA = 0.76 (scan before measuring the area) were utilized.

Raman scattering spectra were determined for all 11 HNSCC samples. The spectra were acquired from various areas of the sample, covering 7-14 points, depending on its size. The examined areas were identified by the pathologist as the regions of neoplastic infiltration. All spectra underwent preliminary processing by removing backscatter peaks using a dedicated algorithm.

Detailed data processing was carried out using the R environment¹⁹ and the hyperSpec package.²⁰ All spectra were limited to the range of $500\text{-}1800\text{ cm}^{-1}$ and shifted along the wavelength axis. Baseline correction was performed using a polynomial baseline correction²¹ with the automatic detection of spectral regions for a polynomial fitting of the spectral baseline. Before commencing the normalization process, the spectra were smoothed using the Locally Weighted Scatterplot Smoothing (LOWESS) algorithm. The spectra were normalized using the methods for the area under the curve (AUC). Additionally, offset and baseline corrections were applied after normalization.

To determine statistically significant differences in the spectra between the study groups, each spectrum was divided into 65 segments ($20\text{ cm}^{-1}/\text{segment}$). The values of each segment for each tissue separately were averaged and compared using the non-parametric Wilcoxon test with a cutoff value of 0.05 (FDR-corrected).²² Next, to visualize the differences in the analyzed groups, the results were subjected to further multidimensional analysis for the regions differentiating the study groups.

Multivariate analysis of the spectroscopic data

In the study, principal component analysis (PCA) was conducted, and hierarchical clustering analysis (HCA) was applied using the Ward clustering criterion.^{23,24} It was assessed whether the selected spectral ranges exhibited sufficient characteristics to serve as HR-HPV markers in Raman spectroscopy analysis. To conduct discriminant analysis for selected regions (i.e., those that statistically differentiated the spectra the most), partial least squares discriminant analysis (PLSDA) was employed with the technique of assessing the accuracy of cross-validation.²⁵ Following this, a supervised analysis of the spectroscopic data for the study groups was undertaken. The results were compared between PLS, ada, AdaBag, and AdaBoost to identify the most effective classifier. Based on this comparison, the PLS classifier was selected as the most effective tool for material assessment. 10-fold cross-validation was utilized in the assessment using the PLS classifier.

Results

Determination of the presence and type of HPV

Immunohistochemical staining aimed to determine the presence of HPV by assessing the expression of p16 protein (INK4A). In six samples from the control group, i.e., non-HPV-associated HNSCC, tumor cells did not exhibit a massive diffuse color reaction of mixed nature, characteristic of HR-HPV infection.²⁶ Positive immunohistochemical staining was observed only in normal parts of the analyzed section, including isolated basal and basic squamous epithelial cells or stromal cell elements, serving as the positive internal control. In a group of patients with HR-HPV infection, the presence of the virus was confirmed based on the accumulation of p16 protein (INK4A) in the neoplastic cells. This confirmation was supported by the results of RT-PCR. Tests in the control group did not reveal the presence of DNA of any of the 14 types of HR-HPV viruses detected by the test. In the study group, an increase in the amplification curve of HPV16 was observed before the cut-off cycle for all samples, confirming the presence of DNA of the virus with high oncogenic potential.

The comparison of the consistent results of p16 protein (INK4A) expression and the determination of the virus's presence by RT-PCR confirmed the division of the study material into two groups that differed in terms of the etiology of the disease.

Basic analysis of Raman spectra for the biological tissues

The results of the Raman spectra measurements for the samples are presented in **Figure 3**. To enhance the visualization of differences between HPV-positive and negative tissues of tonsillar squamous cell carcinoma, averaged spectroscopic

spectra were prepared and statistically significant ranges were highlighted (in red) after preliminary data processing. Following the initial data processing, a statistical analysis was conducted for all spectral components using the non-parametric Wilcoxon test, along with correction for multiple testing, to identify these ranges. The averaged Raman spectra for HNSCC are depicted in **Figure 4**.

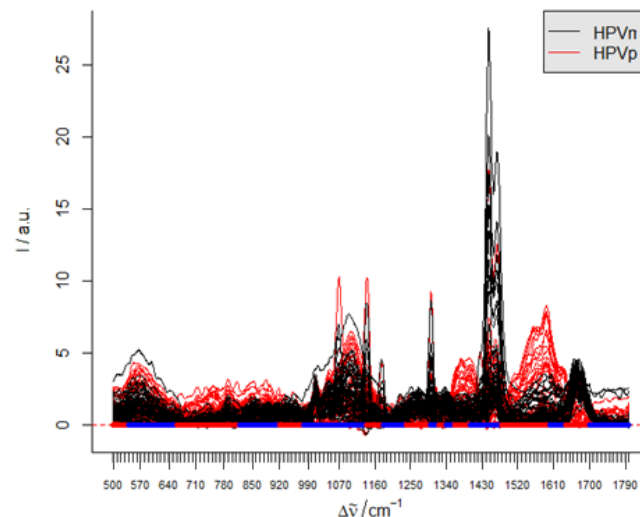


Figure 3. Raman spectra for all tissue sections of tonsillar squamous cell carcinoma, HPV-positive samples (red) and HPV-negative samples (black) after the preliminary analysis and determination of regions that statistically significantly differentiated the study groups. Assignment of bands appearing in the recorded spectra: carotinoids (Raman bands 1490-1580 cm^{-1}), proteins (Raman bands 2900-3010 cm^{-1}), lipids (Raman bands 2850-2950 cm^{-1}), lipids (Raman bands C=C stretch 1652-660 cm^{-1}), lipids (Raman bands CH₂ bending/stretching/scis/soring 1442-1660 cm^{-1}), nucleic acids (Raman bands 1077-1095 cm^{-1}), DNA and RNA (Raman bands 676 and 1400 cm^{-1}), protein (Raman bands 1650 cm^{-1}), amide I (Raman bands (1628-1662 cm^{-1}), amide III (Raman bands (1223-1280 cm^{-1}).

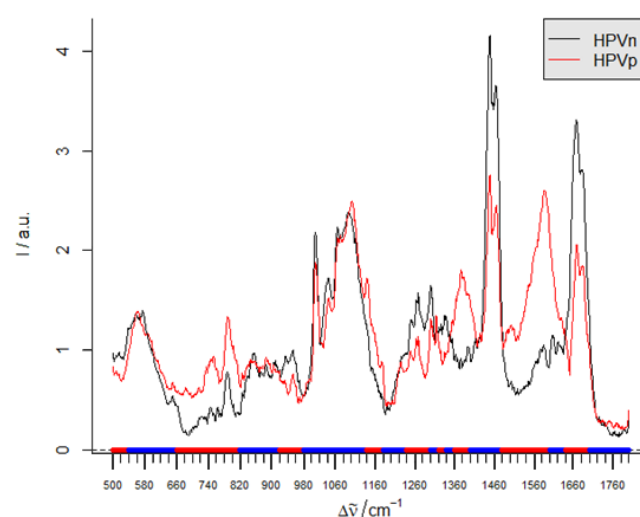


Figure 4. Averaged Raman spectra for HPV-positive and HPV-negative HNSCC, after the initial analysis with the determination of statistically significant (red) and non-significant differentiating regions (blue).

Through the spectral analysis, four ranges were identified in which the bands most effectively differentiated HPV-positive from HPV-negative HNSCC. The selected differentiation ranges include 900 - 980 [rel. 1/cm], 1260 - 1290 [rel. 1/cm], 1360 - 1400 [rel. 1/cm], and 1520 - 1600 [rel. 1/cm].

Multidimensional statistical analysis of the spectroscopic data

Unsupervised multivariate analysis

In the initial step, the selected spectral ranges were evaluated using PCA and HCA. This process aimed to determine whether these regions exhibited sufficient distinctiveness to serve as markers for highly oncogenic HPV. The results obtained for individual ranges are presented below.

PCA for selected ranges (500 – 1800; rel. 1/cm)

In our study, PCA was employed to determine whether the classification of samples based on Raman spectra aligned with histological classification.

The selected differentiating ranges, for which the best results of unsupervised analysis were achieved, were described in detail: 1360 – 1420 [rel. 1/cm], 1540 – 1620 [rel. 1/cm], and 1355 – 1385 [rel. 1/cm]. The graphs presented in **Figure 5** show

the results of PCA for two main components and the dependence of the eigenvalues on the number of factors for the best ranges. The ranges of 1360 - 1400 [rel. 1/cm], 1540 - 1620 [rel. 1/cm], and 1355 - 1385 [rel. 1/cm] were selected. The remaining ranges of 1260 - 1290 [rel. 1/cm] and 900 - 980 [rel. 1/cm] yielded unsatisfactory results in the unsupervised classification.

In the Raman spectra recorded for both groups (HPV positive and negative), differences were observed in the range of 1355 - 1385 cm^{-1} . This range is related to vibrations characteristic of CH₃ and C-CH₃ bonds and the differences in spectra observed in this range may indicate changes in DNA, e.g. methylation. Also, the range of 1540 to 1620 cm^{-1} can most likely be attributed to a change in nucleic acids satisfactory results in the unsupervised classification.

HCA for the ranges from 500 - 1800 [rel. 1/cm]

The results of the HCA are presented in the form of a dendrogram, where the clusters in the diagram are indicated by colored branches. **Figures 6, 7, and 8** display the results of HCA for the selected ranges. The goal of this analysis was to assess whether the classification of the samples based on the Raman spectra aligned with the histological classification.

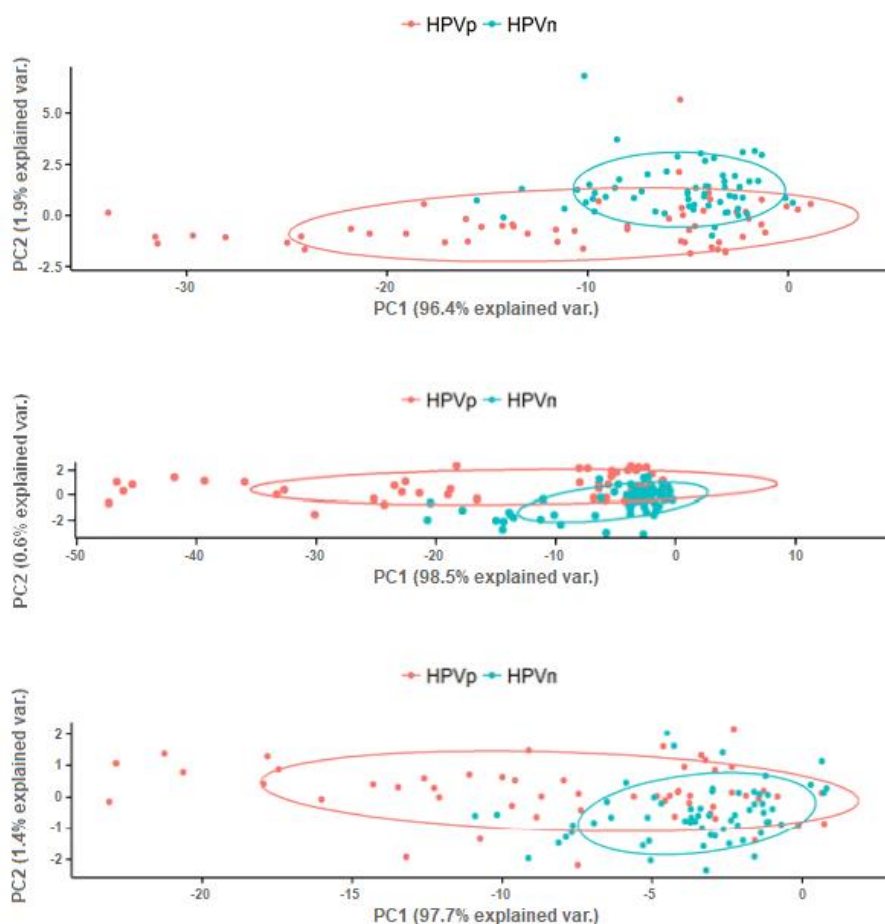


Figure 5. PCA. Dependence of PC2 on PC1 for selected ranges: a) 1360 - 1420 [rel. 1/cm], b) 1540 - 1620 [rel. 1/cm] and c) 1355 - 1385 [rel. 1/cm].

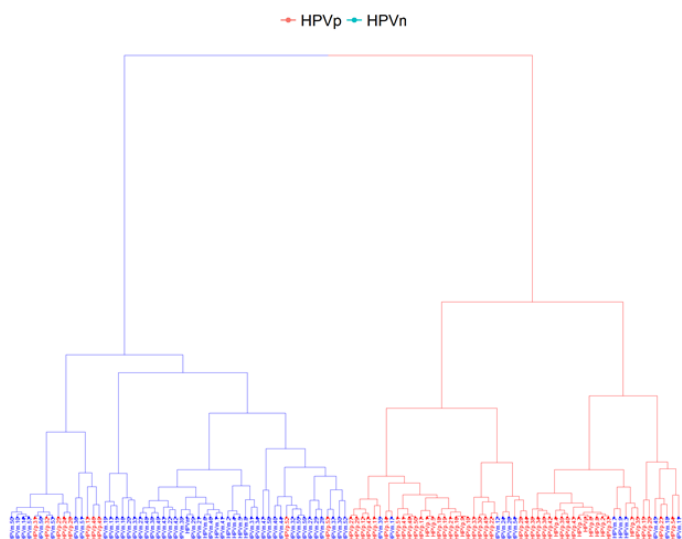


Figure 6. Hierarchical clustering analysis for the range of 1360 – 1420 [rel. 1/cm]. The samples were color-coded due to their belonging to the actual class.

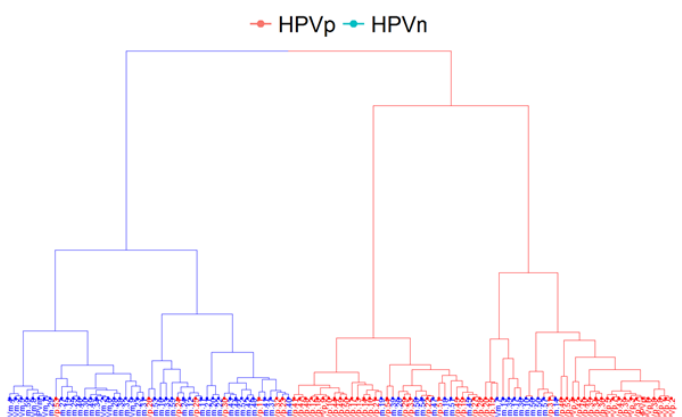


Figure 7. Hierarchical clustering analysis for the range of 1540 – 1620 [rel. 1/cm]. The samples were color-coded due to their belonging to the actual class.

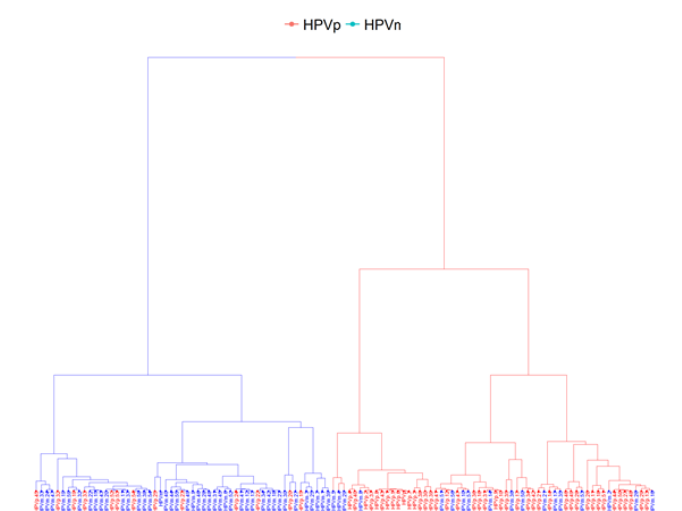


Figure 8. Hierarchical clustering analysis for the range of 1355 – 1385 [rel. 1/cm]. The samples were color-coded due to their belonging to the actual class.

In the dendrograms, samples histologically assessed as HPV-negative were marked with blue, while HPV-positive samples were indicated in red. The division of the samples into two clusters is the outcome of the cluster analysis. Each cluster was designated with a different color (of the dendrogram branch). It can be inferred that the clusters created based on histological assessment partially overlapped with the results of the cluster analysis for the specified ranges. The coverage quality, calculated based on the number of well-matched groups in relation to the total number of trials, is presented in **Table 1**.

Table 1. Comparison of the quality of the HCA classification of the samples based on the Raman spectra for selected ranges in relation to the histopathological diagnosis, n = 115 (HPV-positive n1 = 55, HPV-negative n2 = 60).

Region [rel. 1/cm]	The quality of the HCA classification		
	HPV positive [%]	HPV negative [%]	Total [%]
1360 – 1420	82 %	80 %	81 %
1540 – 1620	86 %	68 %	77 %
1355 – 1385	76 %	70 %	73 %

Supervised analysis of spectroscopic data of the study groups (PLS classifier)

To determine the most effective classifier, the results were compared among PLS, ada, AdaBag, and AdaBoost. Following the comparison, the PLS classifier was chosen as the most effective tool for material assessment. The results of the analyses for each spectral range are provided in **Table 2**.

The optimal matching results of the PLS classifier, at a level of 92%, were observed for the range of 1390 - 1420 [rel. 1/cm]. However, the results of unsupervised clustering remained at a similar level or improved for the ranges of 1380 - 1420, 1390 - 1420, and 1400 - 1420.

Table 2. Summary of the quality of the supervised classification of samples based on Raman spectra for the selected ranges in relation to the histopathological diagnosis.

Range [rel. 1/cm]	Number of components	Quality of the classification	
		Value obtained [%]	SD:
1360 – 1420	7 components of PLS	87 %	10 %
1380 – 1420	7 components of PLS	91 %	9 %
1390 – 1420	8 components of PLS	92 %	6 %
1400 – 1420	5 components of PLS	85 %	9 %
1360 – 1440	10 components of PLS	89 %	19 %
1355 – 1385	8 components of PLS	76 %	12 %
1540 – 1620	8 components of PLS	89 %	9 %
1520 – 1600	10 components of PLS	79 %	10 %
1260 – 1290	8 components of PLS	74 %	16 %
900 – 980	8 components of PLS	74 %	16 %
900 – 980	3 components of PLS	75 %	8 %

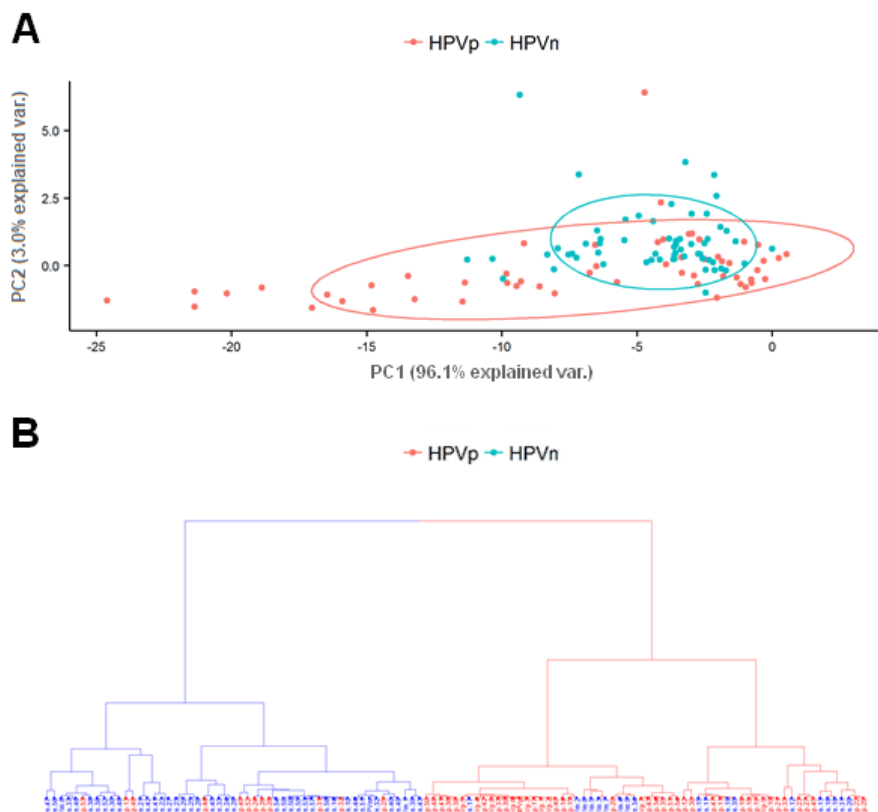


Figure 9. PCA and HCA results for the range of 1390 - 1420 [rel. 1/cm].

Figure 9 presents the results of PCA and HCA for the range of 1390 - 1420 [rel. 1/cm]. The range of 1360 - 1420 [rel. 1/cm] corresponds to the vibrations of CH₃ and C-CH₃ groups. This range may be related to DNA methylation, specifically the conversion of cytosine to C5-methylcytosine, and the range of CHS and C-CH₃ vibrations.

Discussion

The study explored the potential use of Raman microspectroscopy and multidimensional statistical analysis to assess the presence of HR-HPV in HNSCC. This technique, employed for the analysis of biological material, involves combining the capabilities of Raman spectroscopy and confocal microscopy. The integration of Raman scattering spectroscopy with confocal microscopy, utilizing a laser beam with a wavelength of 532 nm and a lens with a magnification of $\times 100$, enables the measurement of individual points and surface mapping (around 300 nm) or depth mapping ($< 1 \mu\text{m}$) with high resolution. The confocal Raman microscope facilitates the registration of a spectrum with high spatial resolution from a very small, precisely selected fragment (approx. $1 \mu\text{m}$), allowing for microstructural level analysis. Additionally, it enables the detection of differences resulting from changes in the chemical structure, such as genetic changes. The study demonstrated that the molecular structure variances in HPV-infected HNSCC cancer cells are significant and detectable by Raman spectroscopy, indicating the potential use of Raman microscopy

in developing promising diagnostic methods. The unique combination of vibrational states of a particle and its environment serves as a chemical fingerprint that can identify and characterize biomolecules, cells, and tissues. Numerous reports have emphasized the use of spectroscopy in assessing molecular changes in DNA structure.²⁷⁻²⁹ In 2010, Kelly et al.³⁰ demonstrated the potential use of Raman spectroscopy, combined with multivariate statistical analysis, to study DNA methylation. DNA methylation significantly influences gene expression regulation and the mechanical properties of DNA. Oligonucleotides with different methylation structures were used as a model system. Hence, spectroscopic methods can effectively study biomarkers related to mutations and diseases caused by epigenetic factors. Vibration spectroscopy techniques have been utilized for years in diagnosing various pathological changes, particularly neoplastic changes.³¹⁻³⁴ A 2015 review by Ramos et al.³⁵ summarized the achievements of Raman spectroscopy in diagnosing cervical cancer, emphasizing its potential in detecting precancerous conditions, early-stage cancer, HPV infection, and monitoring treatment response. These tests are conducted *ex vivo* (organs, cells, or tissues are taken from a living body for treatment or some procedure, and then returned to the living body), *in vivo* (in the body), and *in vitro* (living cells isolated from the parent organism and kept alive under laboratory conditions).

A significant challenge in detecting biochemical changes and conducting immunohistochemical tests on fixed tissues is

associated with the presence of paraffin residues in the samples. Nevertheless, effective methods for removing paraffin from samples and preparing the material for spectroscopic measurements have been reported.³⁶⁻³⁹ The utilization of archival material in the form of formalin-fixed paraffin-preserved (FFPP) samples enables retrospective research. This preservation method, involving formaldehyde fixation and paraffin encapsulation, ensures stability and protection against oxidation for biomolecules. Structurally fixed and embedded tissues remain resilient and suitable for research over many years. However, the stability of nucleic acids and proteins diminishes over time, limiting the utility of such preserved biological material for IHC or molecular tests, especially for samples not older than several years. While double-stranded DNA remains incredibly stable in FFPE blocks, less stable biomolecules, such as RNA, can degrade rapidly, particularly in slice samples.³⁶⁻³⁹

Krishna and Lyng demonstrated the ability to differentiate neoplastic changes (on paraffin sections) based on differences in the concentrations of lipids, nucleic acids, glycogen, and proteins.^{40,41} In *in vitro* studies, Raman spectroscopy primarily allows for differentiation between healthy and neoplastic tissues based on the increasing concentration of DNA in relation to proteins.⁴²

Ostrowska et al. conducted a study to differentiate HPV-positive from HPV-negative cancer, revealing significant differences in LR-HPV cells from HR-HPV cells due to varying levels of proteins, nucleic acids, and lipids.⁴³ In the present study, the assessment focused on tissue changes caused by high oncogenic risk virus types, which are crucial in oncology. This is justified by Śnietura's findings⁴⁴, indicating that there were no infections with subtypes from the LR-HPV group in malignant lesions, suggesting mutual exclusivity between HNSCC and LR-HPV infection.

Raman spectroscopy, employed for detecting various pathological changes, enables the detection of viral infections that often lead to diseases.⁴⁵⁻⁴⁷ This technique allows for non-invasive detection of viruses and their components by assessing the spatial structure of tissue-building macromolecules at the molecular level. Viruses successfully detected using this method include HPV^{43,46}, Ad-CMV adenovirus⁴⁸, and herpes simplex virus type 1 (HSV-1).⁴⁹ Molecular changes visualized by Raman spectroscopy mainly pertain to different proportions of nucleic acids in relation to proteins, particularly involving the PO₃-group⁴⁸ and amide vibrations from amide III proteins.⁴⁹ The impact of external factors on the spatial structure of viruses,

determining methods of infection treatment, has also been examined for echovirus 1 (EV1)⁵⁰, avipoxvirus (APV), and adeno-associated virus (AAV).⁵¹

In the study, two groups were defined based on histological assessment and the evaluation of the biological activity of HR-HPV. An indirect immunohistochemical method of P16 protein (INK4A) accumulation was used for this purpose, and the presence of high oncogenic potential viruses was directly confirmed by RT-PCR. The specific type of the virus was also determined. Additionally, Raman spectroscopy successfully differentiated between HR-HPV-positive and HR-HPV-negative samples in HNSCC.

Conclusions

The research demonstrated that Raman confocal microspectroscopy enables the examination of tissue section samples from paraffin blocks after dewaxing. The utilization of virtual slides of stained tissues with marked tumor structures facilitated measurements in well-defined tissue areas. The basic analysis of Raman scattering spectra from heterogeneous tissue, based on characteristic bands of individual biological substances, was insufficient for effectively classifying the presence of active HPV. However, applying the Wilcoxon test and correction for multiple testing allowed the identification of regions that statistically differentiated the groups. Unsupervised multivariate statistical analysis (PCA, HCA) for evaluating spectroscopic data enabled the accurate classification of active HR-HPV presence. Similarly, supervised multivariate statistical analysis for evaluating spectroscopic data allowed for high-specificity classification of active HR-HPV.

In conclusion, our experience highlights the significant diagnostic potential of Raman confocal microscopy combined with multidimensional statistical analysis. This method facilitates testing biological material obtained from paraffin blocks. Our results revealed distinct ranges of the Raman spectrum that differentiated between HPV-associated and non-HPV-associated cancers. In the future, this method's use may lead to the development of an effective and automated HR-HPV detection system in neoplastic tissue. It also holds promise for creating a rapid and cost-effective screening tool for intraoperative examinations. Considering that the positions of Raman bands are characteristic of specific chemical groups, Raman spectra can be employed for the chemical identification and quantification of tissues.

References

1. Ferlay J, Shin H R, Bray F, Forman D, Mathers C, Parkin D M. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;127(12):2893-2917. <https://doi.org/10.1002/ijc.25516>

2. Decker J, Goldstein JC. Risk factors in head and neck cancer. *N Engl J Med.* 1982;306:1151-1155. <https://doi.org/10.1056/NEJM198205133061905>
3. Licitra L, Locati LD, Bossi P. Head and neck cancer. *Ann Oncol.* 2004;15(Suppl4):iv267-273. <https://doi.org/10.1093/annonc/mdh937>
4. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med.* 2010;363:24-35. <https://doi.org/10.1056/NEJMoa0912217>
5. Ragin CC, Taioli E. Survival of squamous cell carcinoma of the head and neck in relation to human papillomavirus infection: review and meta-analysis. *Int J Cancer.* 2007;121:1813-1820. <https://doi.org/10.1002/ijc.22851>
6. Fakhry C, Westra WH, Li S, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *J Natl Cancer Inst.* 2008;100:261-269. <https://doi.org/10.1093/jnci/djn011>
7. Vidal L, Gillison LV. Human papillomavirus in HNSCC: recognition of a distinct disease type. *Hematol Oncol Clin N Am.* 2008;22:1125-1142. <https://doi.org/10.1016/j.hoc.2008.08.006>
8. Snietura M, Piglowski W., Jaworska M, Mucha-Malecka A, Wozniak G, Lange D, Suwinski R. Impact of HPV infection on the clinical outcome of p-CAIR trial in head and neck cancer. *Eur Arch Otorhinolaryngol.* 2011;268:721-726. <https://doi.org/10.1007/s00405-010-1396-7>
9. Kujdowicz M, Januś D, Taczanowska-Niemczuk A, Lankosz MW, Adamek D. Raman Spectroscopy as a Potential Adjunct of Thyroid Nodule Evaluation: A Systematic Review. *Int J Mol Sci.* 2023;24(20):15131. <https://doi.org/10.3390/ijms242015131>
10. Kujdowicz M, Placha W, Mech B, Chrabaszcz K, Okoń K, Malek K. In Vitro Spectroscopy-Based Profiling of Urothelial Carcinoma: A Fourier Transform Infrared and Raman Imaging Study. *Cancers (Basel).* 2021;13(1):123. <https://doi.org/10.3390/cancers13010123>
11. Neto V, Esteves-Ferreira S, Inácio I, et al. Metabolic Profile Characterization of Different Thyroid Nodules Using FTIR Spectroscopy: A Review. *Metabolites.* 2022;12(1):53. <https://doi.org/10.3390/metabo12010053>
12. Wahadoszamen M, Rahaman A, Hoque NM, I Talukder A, Abedin KM, Haider AF. Laser raman spectroscopy with different excitation sources and extension to surface enhanced raman spectroscopy. *J Spectrosc.* 2015;2015:1-8. <https://doi.org/10.1155/2015/895317>
13. Śnietura M, Brewczynski A, Waniczek D, Kopec A, Stanek-Widera A, Muc-Wierzoń M, Rutkowski T. Is p16 expression still a surrogate marker for human papillomavirus infection in oral squamous cell carcinomas? *J Biol Regul Homeost Agents.* 2020;34(3):1153-1156. <https://doi.org/10.23812/19-423-L-58>
14. Smeets SJ, Hesselink AT, Speel E-JM, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer.* 2007;121(11):2465-2472. <https://doi.org/10.1002/ijc.22980>
15. Huang S, Erickson B, Tang N, et al. Clinical performance of Abbott RealTime High Risk HPV test for detection of high-grade cervical intraepithelial neoplasia in women with abnormal cytology. *J Clin Virol.* 2009;45(Suppl 1):S19-23. [https://doi.org/10.1016/S1386-6532\(09\)70004-6](https://doi.org/10.1016/S1386-6532(09)70004-6)
16. Huang S, Tang N, Mak W-B, et al. Principles and analytical performance of Abbott RealTime High Risk HPV test. *J Clin Virol.* 2009;45(Suppl 1):S13-17. [https://doi.org/10.1016/S1386-6532\(09\)70003-4](https://doi.org/10.1016/S1386-6532(09)70003-4)
17. Snietura M, Piglowski W, Jaworska M, et al. Impact of HPV infection on the clinical outcome of p-CAIR trial in head and neck cancer. *Eur Arch Otorhinolaryngol.* 2011;268(5):721-726. <https://doi.org/10.1007/s00405-010-1396-7>
18. Kocjan BJ, Maver PJ, Hosnjak L, et al. Comparative evaluation of the Abbott RealTime High Risk HPV test and INNO-LiPA HPV Genotyping Extra test for detecting and identifying human papillomaviruses in archival tissue specimens of head and neck cancers. *Acta Dermatovenerol Alp PannonicaAdriat.* 2012;21(4):73-75.
19. R Core Team. R: A language and environment for statistical computing. R 2.5.2 Foundation for Statistical Computing, Vienna, Austria. 2016; URL <https://www.R-project.org/>.
20. Beleites C, Sergio V.: hyperSpec: a package to handle hyperspectral data sets in R, R package version 0.98-20161118.
21. HovdeLiland, Helge Mevik B. Baseline: Baseline Correction of Spectra, 2015; URL <https://CRAN.R-project.org/package=baseline>. R package version 1.2-1.
22. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B.* 1995;57:289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
23. Mardia KV, Kent JT, Bibby JM. *Multivariate Analysis*, London: Academic Press.1979.
24. Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *Journal of Classification.* 2014;31:274-295. <https://doi.org/10.1007/s00357-014-9161-z>
25. Perez-Enciso M, Tenenhaus M. Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach. *Human Genetics.* 2003;112:581-592. <https://doi.org/10.1007/s00439-003-0921-9>
26. Smith EM, Wang D, Kim Y, et al. p16INK4a Expression, human papillomavirus, and survival in head and neck cancer. *Oral Oncology.* 2008;44(2):133-142. <https://doi.org/10.1016/j.oraloncology.2007.01.010>
27. AW Auner, JC Thomas, Double-Stranded DNA Damage Assessed with Raman Spectroscopy. *Biochemistry & Analytical Biochemistry.* 2016;5(3). <https://doi.org/10.4172/2161-1009.1000284>

28. Lipiec E, Bambery KR, Heraud P, et al. Monitoring UVR induced damage in single cells and isolated nuclei using SR-FTIR microspectroscopy and 3D confocal Raman imaging. *Analyst*. 2014;139(17):4200-4209. <https://doi.org/10.1039/C4AN00838C>
29. Sánchez V, Redmann K, Wistuba J, et al. Oxidative DNA damage in human sperm can be detected by Raman microspectroscopy. *FertilSteril*. 2012;98(5):1124-1129.e1-3. <https://doi.org/10.1016/j.fertnstert.2012.07.1059>
30. Kelly JG, Najand GM, Martin FL. Characterisation of DNA methylation status using spectroscopy (mid-IR versus Raman) with multivariate analysis. *J Biophotonics*. 2011;4(5):345-354. <https://doi.org/10.1002/jbio.201000085>
31. Krafft C. Bioanalytical applications of Raman spectroscopy. *Anal Bioanal Chem*. 2004;378(1):60-62. <https://doi.org/10.1007/s00216-003-2266-6>
32. Nijssen A, Koljenović S, Bakker Schut TC, Caspers PJ, Puppels GJ. Towards oncological application of Raman spectroscopy. *J Biophotonics*. 2009;2(1-2):29-36. <https://doi.org/10.1002/jbio.200810055>
33. Kendall C, Isabelle M, Bazant-Hegemark F, et al. Vibrational spectroscopy: a clinical tool for cancer diagnostics. *Analyst*. 2009;134(6):1029-1045. <https://doi.org/10.1039/b822130h>
34. Kondepati VR, Heise HM, Backhaus J. Recent applications of near-infrared spectroscopy in cancer diagnosis and therapy. *Anal Bioanal Chem*. 2008;390(1):125-139. <https://doi.org/10.1007/s00216-007-1651-y>
35. Ramos IRM, Malkin A, Lyng FM. Current Advances in the Application of Raman Spectroscopy for Molecular Diagnosis of Cervical Cancer. *Biomed Res Int*. 2015;2015:561242. <https://doi.org/10.1155/2015/561242>
36. Hewitt SM, Lewis FA, Cao Y, et al. Tissue Handling and Specimen Preparation in Surgical Pathology: Issues Concerning the Recovery of Nucleic Acids From Formalin-Fixed, Paraffin-Embedded Tissue. *Archives of Pathology & Laboratory Medicine*. 2008;132(12):1929-1935. <https://doi.org/10.5858/132.12.1929>
37. Howat WJ, Wilson BA. Tissue fixation and the effect of molecular fixatives on downstream staining procedures. *Methods*. 2014;70(1):12-19. <https://doi.org/10.1016/j.ymeth.2014.01.022>
38. Shrivastava A, Aggarwal LM, Murali Krishna C, et al. Diagnostic and prognostic application of Raman spectroscopy in carcinoma cervix: A biomolecular approach. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 2021;250:119356. <https://doi.org/10.1016/j.saa.2020.119356>
39. Man Y, Moifar F, Bratthauer GL, Kuhls EA, Tavassoli FA. An Improved Method for DNA Extraction from Paraffin Sections. *Pathology - Research and Practice*. 2001;197(9):635-642. <https://doi.org/10.1078/0344-0338-00138>
40. Krishna CM, Sockalingum GD, Vadhiraja BM, et al. Vibrational spectroscopy studies of formalin-fixed cervix tissues. *Biopolymers*. 2007;85(3):214-221. <https://doi.org/10.1002/bip.20631>
41. Lyng FM, Traynor D, Ramos IRM, Bonnier F, Byrne HJ. Raman spectroscopy for screening and diagnosis of cervical cancer. *Anal Bioanal Chem*. 2015;407(27):8279-8289. <https://doi.org/10.1007/s00216-015-8946-1>
42. Yazdi Y, Ramanujam N, Lotan R, Mitchell MF, Hittelman W, Richards-Kortum R. Resonance Raman Spectroscopy at 257 nm Excitation of Normal and Malignant Cultured Breast and Cervical Cells. *Appl Spectrosc, AS*. 1999;53(1):82-85. <https://doi.org/10.1366/0003702991945254>
43. Ostrowska KM, Malkin A, Meade A, et al. Investigation of the influence of high-risk human papillomavirus on the biochemical composition of cervical cancer cells using vibrational spectroscopy. *Analyst*. 2010;135(12):3087-3093. <https://doi.org/10.1039/c0an00571a>
44. Sniectura M, Lamch R, Kopec A, Waniczek D, Likus W, Lange D, Markowski J Oral and oropharyngeal papillomas are not associated with high-risk human papillomavirus infection. *Eur Arch Otorhinolaryngol*. 2017 Sep;274(9):3477-3483. <https://doi.org/10.1007/s00405-017-4649-x>
45. Schäfer A, Lengenfelder D, Grillhösl C, Wieser C, Fleckenstein B, Ensser A. The latency-associated nuclear antigen homolog of herpesvirus saimiri inhibits lytic virus replication. *J Virol*. 2003;77(10):5911-5925. <https://doi.org/10.1128/JVI.77.10.5911-5925.2003>
46. Vargis E, Tang Y-W, Khabele D, Mahadevan-Jansen A. Near-infrared Raman Microspectroscopy Detects High-risk Human Papillomaviruses. *Transl Oncol*. 2012;5(3):172-179. <https://doi.org/10.1593/tlo.12106>
47. Ringelhan M, Heikenwalder M, Protzer U. Direct effects of hepatitis B virus-encoded proteins and chronic infection in liver cancer development. *Dig Dis*. 2013;31(1):138-151. <https://doi.org/10.1159/000347209>
48. Moor K, Ohtani K, Myrzakozha D, Zhanserkenova O, Andriana BB, Sato H. Noninvasive and label-free determination of virus infected cells by Raman spectroscopy. *J Biomed Opt*. 2014;19(6):067003. <https://doi.org/10.1117/1.JBO.19.6.067003>
49. Salman A, Shufan E, Zeiri L, Huleihel M. Characterization and detection of Vero cells infected with Herpes Simplex Virus type 1 using Raman spectroscopy and advanced statistical methods. *Methods*. 2014;68(2):364-370. <https://doi.org/10.1016/j.ymeth.2014.02.022>
50. Ruokola P, Dadu E, Kazmertsuk A, Häkkänen H, Marjomäki V, Ihalainen JA. Raman Spectroscopic Signatures of Echovirus 1 Uncoating. *J Virol*. 2014;88(15):8504-8513. <https://doi.org/10.1128/JVI.03398-13>
51. Hermann P, Hermelink A, Lausch V, et al. Evaluation of tip-enhanced Raman spectroscopy for characterizing different virus strains. *Analyst*. 2011;136(6):1148-1152. <https://doi.org/10.1039/c0an00531b>