Analytica Chimica Acta 1177 (2021) 338784

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Raman spectroscopy and surface-enhanced Raman spectroscopy (SERS) spectra of salivary glands carcinoma, tumor and healthy tissues and their homogenates analyzed by chemometry: Towards development of the novel tool for clinical diagnosis

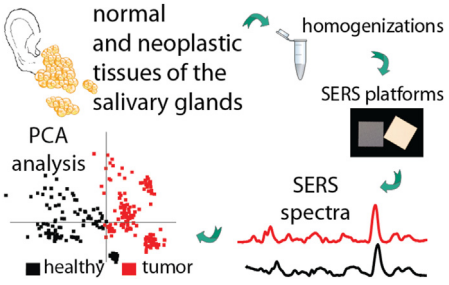
M. Czaplicka [a], A.A. Kowalska [a, **], A.B. Nowicka [a], D. Kurzydłowski [b], Z. Gronkiewicz [c], A. Machulak [c], W. Kukwa [c], A. Kamińska [a, *]

[a] Institute of Physical Chemistry, Polish Academy of Sciences, Kasprzaka 44/52, 01-224, Warsaw, Poland
[b] Cardinal Stefan Wyszyński University in Warsaw, Dewajtis 5, 01-815, Warsaw, Poland
[c] Szpital Czerniakowski, Medical University of Warsaw, Żwirki I Wigury 61, 02-091, Warsaw, Poland

HIGHLIGHTS

- A new method of tissue homogenization.
- Developed approach allows to register the spectra of tumor samples in similar spectral quality in comparison to healthy tissue.
- Salivary glands of healthy and neoplastic tissues analyzed by SERS technique combine with MVA method.
- PLS-DA analysis as a best methods among the other (PCA, PCR) for salivary glands tissues differentiation.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 12 April 2021
Received in revised form
25 May 2021
Accepted 17 June 2021
Available online 18 June 2021

Keywords:
Salivary glands
Raman spectroscopy
Surface-enhancement Raman spectroscopy
Multivariate analysis

ABSTRACT

In this study, two approaches to salivary glands studies are presented: Raman imaging (RI) of tissue cross-section and surface-enhanced Raman spectroscopy (SERS) of tissue homogenates prepared according to elaborated protocol. Collected and analyzed data demonstrate the significant potential of SERS combined with multivariate analysis for distinguishing carcinoma or tumor from the normal salivary gland tissues as a rapid, label-free tool in cancer detection in oncological diagnostics. Raman imaging allows a detailed analysis of the cell wall's chemical composition; thus, the compound's distribution can be semi-quantitatively analyzed, while SERS of tissue homogenates allow for detailed analysis of all moieties forming these tissues. In this sense, SERS is more sensitive and reliable to study any changes in the area of infected tissues. Principal component analysis (PCA), as an unsupervised pattern recognition method, was used to identify the differences in the SERS salivary glands homogenates. The partial least squares-discriminant analysis (PLS-DA), the supervised pattern classification technique, was also used to strengthen further the computed model based on the latent variables in the SERS spectra. Moreover, the chemometric quantification of obtained data was analyzed using principal component regression (PCR) multivariate calibration. The presented data prove that the PCA algorithm allows for 91% in seven following components and the determination between healthy and tumor salivary gland homogenates.

* Corresponding author.
** Corresponding author.
E-mail addresses: akowalska@ichf.edu.pl (A.A. Kowalska), akaminska@ichf.edu.pl (A. Kamińska).

https://doi.org/10.1016/j.aca.2021.338784

The PCR and PLS-DA methods predict 90% and 95% of the variance between the studied groups (in 6 components and 4 factors, respectively). Moreover, according to calculated RMSEC (RMSEP), R2C (R2P) values and correlation accuracy (based on the ROC curve), the PLS-DA model fits better for the studied data. Thus, SERS methods combined with PLS-DA analysis can be used to differentiate healthy, neoplastic, and mixed tissues as a competitive tool in relation to the commonly used method of histopathological staining of tumor tissue.

1. Introduction

Salivary gland carcinoma and tumor originate from any of the three major glands (parotid, submandibular and sublingual) and the minor glands (e.g. located in the palate, lips, buccal mucosa). This type of cancer constitutes about 3–11% of head and neck tumors [1] and 0.5% of all malignancies [2–7]. Mortality depends on the stage and type of the lesion; however, the 5-year survival rate is estimated at 72% [8]. Salivary gland cancers are histologically a highly heterogeneous group. In 2017, the World Health Organization (WHO) categorized tumors of salivary glands into 32 malignant and benign histologic subtypes [9].The most common histology types of salivary gland cancers included adenocarcinoma (68%) and epidermoid carcinoma (22%) [10]. Major salivary adenocarcinoma comprised acinic cell carcinoma, adenoid cystic carcinoma, salivary duct carcinoma, and polymorphous adenocarcinoma [9].

Fortunate, large number of salivary gland neoplasms are benign lesions, however in some cases the complex process of benign tumors dedifferentiation or malignant transformation leads also to salivary gland carcinomas [11]. The pleomorphic adenoma is the most common benign tumor of salivary glands, more frequent located in the parotid and submandibular glands and accounts for 90% of all salivary gland tumors [12,13].

Salivary glands neoplasms consist the diverse group of tumors with different structural and histologic characteristics. The great variety of histological types of salivary glands make theirs pathologic diagnosis and treatment still complex and challenging for otorhinolaryngologists and head and neck surgeons. The choice of the right treatment method determines the probability of survival for a patient [14].

In case of salivary gland carcinomas or tumors, the diagnosis is mainly based on histopathological analysis, but also on computer tomography CT scan, magnetic resonance imaging (MRI), positron emission tomography (PET), X-ray and ultrasound [15]. Histological analysis is the best method and standard procedure to diagnose neoplastic tissue in the salivary gland. However, this method can be time consuming, we receive the result after 2–3 weeks and sometimes it leads to ambiguity susceptible to human interpretations. A new intraoperative determination of the tumors can speed up diagnosis and implementation of appropriate treatment.

Raman spectroscopy is a light scattering method, used to measure the vibrational frequencies of samples and has applications in many areas, for example in medicine and industry [16]. Raman spectroscopy can provide a molecular fingerprint of a sample, but the signal is weak, due to the poor photon efficiency of the Raman scattering process, and often masked by strong background fluorescence. When the analyte is adsorbed on the rough surface of metal (usually made of silver and gold) we can observe an increase in the weak Raman signal, this phenomena is known as surface-enhancement Raman spectroscopy and was first reported by Fleishman et al., in 1974 [17]. This makes Raman spectroscopy and surface-enhanced Raman spectroscopy a very promising method to study the biological differences in tissues. These methods may facilitate distinguishing between healthy and abnormal tissue, and even distinguish between benign tumors and malignancy. In recent years these methods are widely used to distinguish between normal and tumors or carcinoma tissues in various body places, such as the brain [18], breast [19], skin [20], stomach [21], salivary glands [14], kidneys [22], colon [23].

In order to obtain an appropriately enhanced signal in surface-enhanced Raman spectroscopy, a rough substrate is required, usually covered with a layer of silver or gold, however, appropriate preparation of the test sample is also important. Tissue homogenization is the most common method for preparing cells and tissues. This process involves lysing the cells to release intracellular contents of interest, such as proteins and nuclear components. Tissue homogenization in solution is usually performed simultaneously with cell lysis. This method leads to the loss or weakening of the integrity of the cell wall and/or membrane, causing its disintegration and release of individual cell elements. Homogenization can be carried out using enzymes of cell wall components or using additional devices that cause direct damage to the membrane and/or the cell wall [24].

In the patent application “A tissue differentiation method based on surface enhanced Raman Scattering” [25] (number EP2260292B1), an innovative method of preparing tissue homogenate by immersing the tissue in nitrogen and mechanically disintegrating was proposed. The patent uses the SERS reinforcement method using silver nanoparticles. In our recent publication [18] we demonstrated the possibility of measuring SERS of tissue homogenates, made using the enzymatic method, on SERS media.

In this article we present a new method of sample preparation and tissue homogenization which enables distinguishing spectral differences between tumor or carcinoma and healthy gland tissues. The purpose of this work is to use Raman spectroscopy in a sense of Raman Imaging (RI) and surface-enhanced Raman spectroscopy (SERS). SERS method open up new possibilities of diagnostics, allows fast, noninvasive, sensitive diagnosis of benign and malignant tumors of the salivary glands and to distinguish them from healthy tissues. To demonstrate the possibility of distinguishing between two types of salivary gland tissues, namely healthy and neoplastic, samples were examined using two different approaches. First, tissues were studied using RI and recorded data were analyzed by components analysis, implemented in the WITec software. Then salivary gland homogenates were studied using SERS technique. The obtained SERS data were analyzed using the multi-dimensional approach, as first the PCA was used as initial data processing of the multivariate analysis. Then, in order to maximize the variations between the analyzed groups the principal component regression (PCR) and the partial least squares - discriminant analysis (PLS-DA) were used. The aim of this approach was to compare the potential of various techniques and methods of analysis of the obtained spectral data with the algorithms to reach the maximum group separation of the homogenates salivary glands. The PLS-DA model combine with Raman spectroscopy was used before to analyze

colon tissues [26] and colorectal cancer diagnostics [27]. Other modification of chemometric analysis such as PCR combined with Raman spectra of the teeth was used in forensic application to age determination model [28] while combined with ATR-FTIR to detection of breast cancer in blood serum [29]. In our case such approach (PCA, PCR and PLS-DA) enables to select the best method of healthy and tumor salivary glands differentiation.

2. Experimental section

In the presented research Raman Imaging and surface-enhanced Raman spectroscopy have been employed to analyze human salivary tumors. The protocol of study was approved by the Ethics and Bioethics Committee of the Cardinal Stefan Wyszyński University in Warsaw. Informed consent was obtained from all patients.

2.1. Clinical samples preparation

Total number of patients included in this study are 11. Total number of collected samples 22 (11 healthy and 11 tumor and carcinoma). Within these samples 5 are adenoma polymorphum, 2 are adenolymphoma, 1 sample with both tumors (total 8 neoplastic samples) and 2 salivary duct carcinoma G3, 1 carcinoma andenoides cysticum G-3 (3 carcinoma samples). The histopathological images are gathered on Fig. 1.

Samples tissues were obtained from the safety margin (healthy samples) and the tissues from the tumor mass (neoplastic samples). Tissues samples were snap frozen and stored at −80 °C. The samples were divided into two parts, for testing with Raman spectroscopy and surface-enhanced Raman spectroscopy.

2.2. Preparation of tissues for histopathological measurements

Tissues from the patient were gathered, fixed in 4% formaldehyde, dehydrated and embedded in paraffin. Next, the paraffin bloc was sectioned into 5 μm thick slices. After deparaffinization and rehydration the slides were stained with hematoxylin for 30 s. Next, the slides were rinsed in running water for the duration of 1 min. After this, 30 s of staining with 1% eosin Y solution occurred. The tissues were then dehydrated in absolute alcohols and the alcohol was extracted with two changes of xylene. Lastly, a mounting medium was applied and the tissues were covered with a coverslip and analyzed and photographed under a microscope.

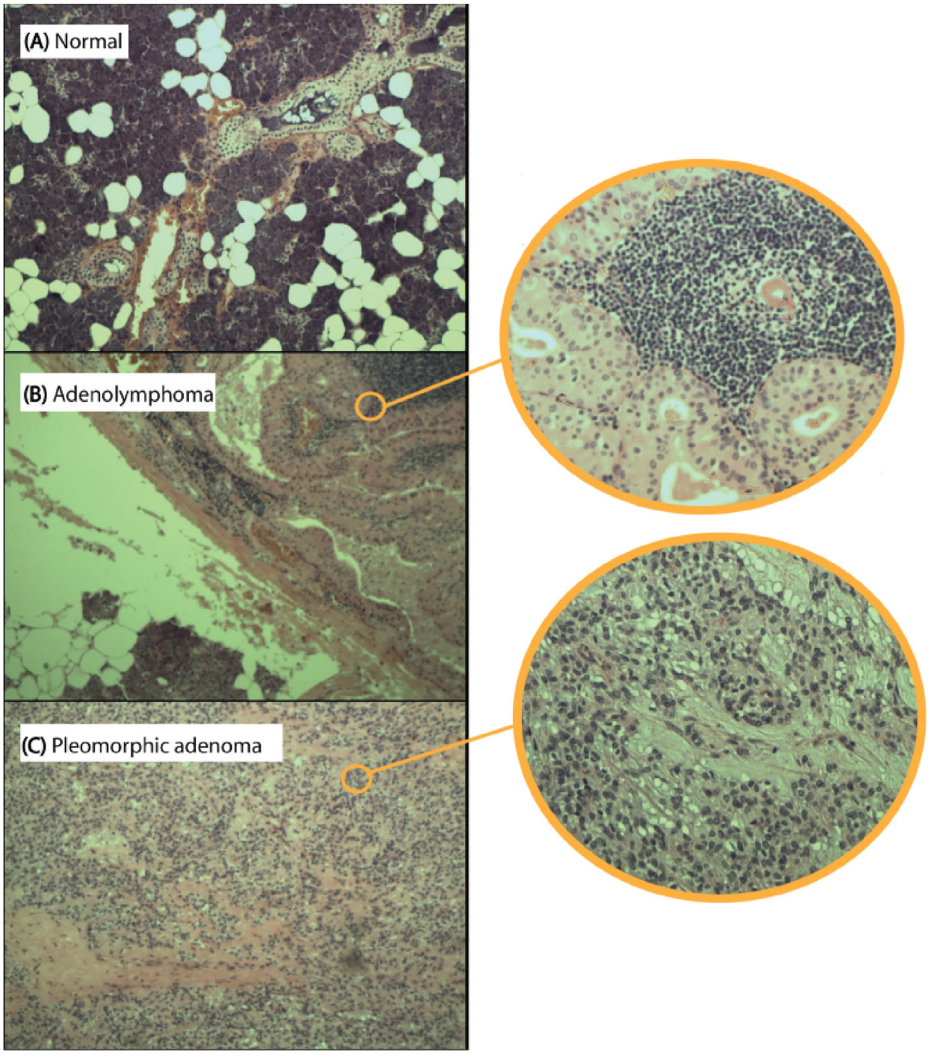


Fig. 1. Photomicrographs of histological slides of H&E stained salivary glands. Normal histological appearance of the salivary gland in the healthy tissues (**A**), adenolymphoma (**B**), and pleomorphic adenoma (**C**) (hematoxylin and eosin staining; magnification, x200 (right), x100 (left)).

2.3. Preparation of tissue samples for Raman imaging

Frozen tissues (−80 °C) were removed from the low-temperature freezer, then were immersed in liquid nitrogen and cut with a scalpel. This piece of tissue was placed directly on a microscope slide and subjected to Raman mapping measurements.

2.4. SERS platform preparation

The SERS-active silicon substrate was prepared according already to published procedure [30]. Firstly, the physical modification of the silicon surface was performed using femtosecond laser ($\lambda = 1030$ nm) with the active medium of the oscillator doped with ytterbium crystal Yb:KYW, repetition rate 300 kHz, pulse width 300 fs, distance between scanning lines 30 μm and scanning rate of the laser beam on the surface of silicon was 1.5 m/s. Then to completed the SERS substrate preparation the Ag layer was sputtered on the surface of modified silicon using the PVD device (Quorum, Q150T ES, Laughton, UK). The thickness of the silver layer, measured during the process with quartz microbalance, was set to 100 nm for all analyzed samples. The sputtering layer was applied to the silicon with current 25 mA and under the pressure 10^{-2} mbar. Such freshly prepared SERS platforms were use throughout all presented experiments.

2.5. Raman measurements and component analysis

The spectra were acquired with an Alpha300 M + confocal microscope (Witec Gmbh, Ulm, Germany) equipped with a motorized stage. The confocal set-up enabled acquiring the signal from a small portion of the sample (approximately 2 × 2 × 2 μm) A 633 nm laser line was used, delivered to the microscope through a single-mode optical fiber. The laser power at the sample did not exceed 2 mW. The backscattered Raman signal was collected through a 20 × long working distance objective (numerical aperture, N.A., equal to 0.4), and passed through a multi-mode optical fiber (50 μm core diameter) to a lens based spectrometer (Witec UHTS 300, f/4 aperture, focal length 300 mm) coupled with a back-illuminated Andor iDUS 401 detector (Oxford Instruments, Abingdon-on-Thames, UK) thermoelectrically cooled to −60 °C. The spectra were collected with the use of an 600 lines/mm grating in the 600–3250 cm^{-1} range. Area scans (80 × 80 μm, 40 x 40 = 1 600 points) were conducted with an integration time per point equal to 0.5 s Data preprocessing of gathered Raman spectra was performed using the WITec Project FIVE software (WITec, Germany). This operation consists of the following steps: cosmic ray remover (CRR), background subtraction, and normalization. Cosmic ray removal was carried out before any further analysis based on an intensity threshold set by taking into account spectral and spatial pixels adjacent to the pixel of interest.

For the analysis of the area scans obtained with the WITec spectrometer we used a built-in function of the WITec Project Plus software, which allows to decompose the spectrum registered at each pixel of the area scan into a linear combination of a given set of basis spectra (components). The spectra of components 1 and 2 (see text) were selected manually from the regions were strong Raman signal was observed. Component 1 was taken from a region were strong bands at 1159 and 1523 cm^{-1} were observed, while component 2 from a region were these bands were absent or weak.

After averaging the spectrum of components 1 and 2 over the respective regions a difference map (difference between the registered spectrum and the linear combination of components 1 and 2) was drawn. After that components 3 and 4 were chosen as spectra from two regions characterized by large intensity of the difference map. These components are characterized by a very weak Raman signal.

The four bases functions (components) obtained in the procedure described above were used to draw the Raman map shown in Fig. 2B – the intensity of each color in this map is proportional to the coefficient of linear combination.

2.6. SERS measurements

The measurements were performed using Bruker's BRAVO spectrometer equipped with Duo LASER™(700–1100 nm) and CCD camera. The laser power was 100 mW for both LASERs and the spectral resolution was 2–4 cm^{-1}. Typically, 20 SERS spectra for each cell type were acquired. Each spectrum was measured for 30 s.

2.7. Multivariate analysis (MVA)

Firstly, to reduce dimensionality of gathered data, the unsupervised MVA method, such as principal component analysis (PCA) was performed using the commercial Unscrambler® software

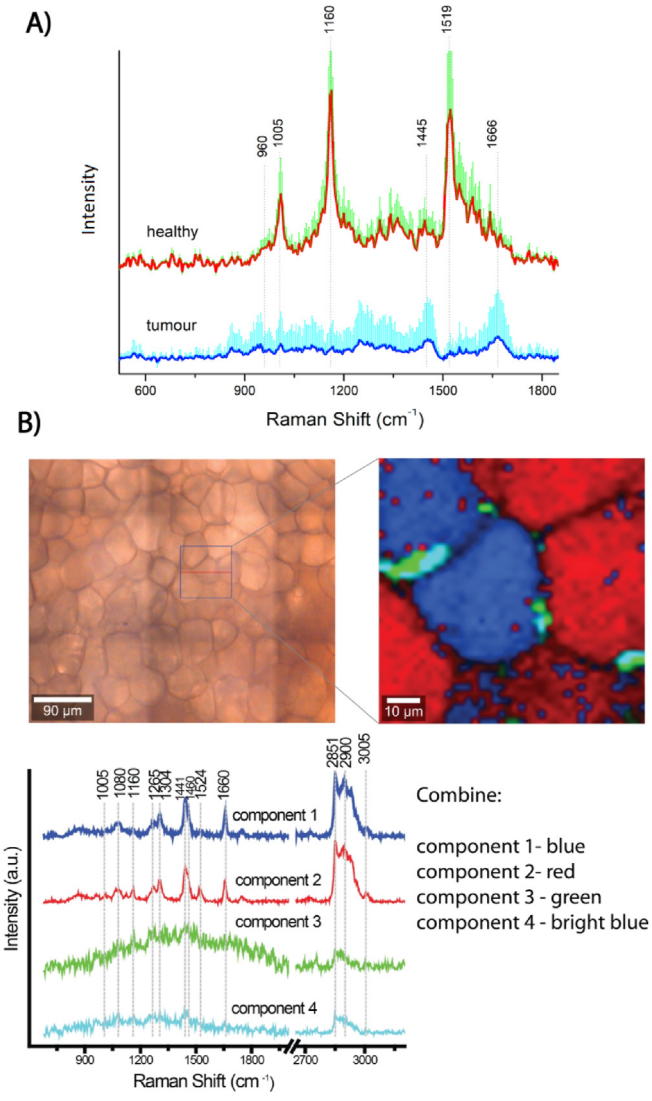


Fig. 2. Representative Raman mean spectra from RI of neoplastic and healthy salivary glands tissues (**A**) presented together with the tissues images and components analysis based on the gathered data (**B**). Presented Raman spectra averages, for both healthy and tumor tissues, are based onto 40 spectra (within total 1600 spectra for each region). The standard error deviation is visualized on the plots in a plus direction. Data presented for the 600-1800 cm^{-1} region.

(CAMO software AS, version 10.3, Oslo, Norway). Before PCA analysis SERS data were optimized for PCA using the following steps: (*i*) smoothing with a Savitzky–Golay filter (Oslo, Norway), (*ii*) background correction (concave rubber band correction; the number of baseline points was 34 and the number of iterations was 10), and (*iii*) normalization using OPUS software (Bruker Optic GmbH, 2012 version, Ettlingen, Germany). The PCA was completed based on the NIPALS algorithm, cross-validation validation (random with 20 segments), significance 0.05 (statistical significance of the contribution of variables to the PCA solution based on independent T-test), and a SERS spectra number of 120. In other words, the selected PC scores from significantly different PCs ($p < 0.05$) was used to build the PCA model in an unbiased manner the leave-one-out and random cross validation method.

To classify new spectra of salivary glands homogenates, with unknown origin, two methods were applied, such as the supervised classification algorithm partial least squares - discriminant analysis (PLS-DA) and principal component regression (PCR). For that, we have used original SERS data and preprocesses them using both manners: first mentioned above implemented in the OPUS software and second using the standard normal variate (SNV) method implemented in the Unscrambler. The second method turned out to be better for analyzed data. Therefore whole data presented for PCR and PLS-DA are preprocessed using SNV.

Based on the PCR and PLS-DA the RMSE (the root mean square error) and R2 (the coefficient of determination of the model fit) are presented in the manuscript. Additionally, the ROC (the receiver operating characteristic) and AUC (the area under the curve) allows to determinate the correlation accuracy, sensitivity, and specificity of presented multivariate methods of differentiation between homogenates of neoplastic and healthy salivary glands tissues.

The PCR and PLS-DA calculations were performed for 8 maps from healthy patients and 9 maps from sick patients. The training set used 5 maps from healthy patients and 6 maps from sick patients. The test set includes maps that were not included in the training set, 3 maps from healthy patients samples and 3 maps from sick patients samples.

3. Results and discussion

The most common benign tumor of the salivary glands is the pleomorphic adenoma. This is an epithelial tumor of complex morphology, possessing epithelial and myoepithelial elements intermingled with mucoid, myxoid, or chondroid tissue arranged in a variety of patterns and embedded in a mucopolysaccharide stroma [31]. Finding a spectroscopic technique to quickly analyze and distinguish between tissues of healthy and neoplastic or carcinoma may contribute to a faster medical diagnosis, which in the future may enable to be used intraoperatively and consequently reduce requiring wide safety margins only to cancerous tissues. That conception of used SERS intraoperatively in brain tumor of mouse shown H. Karabebern et al. [32]. For this purpose, a method of distinguishing different types of salivary gland tumors between themselves as well as differentiating them from healthy tissues using the MVA technique was presented.

3.1. Histopathological method

Histopathological examination is essential for tumors and neoplastic tissues recognition in medical diagnosis. The topographic overview staining, allows to assess the whole structure of the tissue, by contrasting the staining of the cytoplasm and cell nuclei. Fig. 1 presents a comparison of histology images of the healthy (Fig. 1A) salivary gland tissue with two neoplastic lesions (Fig. 1B, Fi.g 1C). Adenolymphona (Fig. 1B) shows a typical strong Lymphoid infiltration and solid epithelial sockets, but the second tumor lesion, adenoma pleomorphum (Fig. 1C) consists of mixed epithelial and mesenchymal cell components.

3.2. Raman imaging and component analysis

The Raman imaging (RI) allows for the spectral characteristics of different tissues. Variance between the spectra are expected in relation to the tissues origin, as can be specific for a given organ. In that sense, in the recorded Raman data for the healthy and cancerous section of salivary gland, only the difference coming from the state of samples – are expected to be visualized. Raman spectra presented on Fig. 2A, indicate that good quality of data are gathered only for healthy samples. In a case of neoplastic sample the efficiency of Raman signal do not allows for its sufficient chemical analysis. The Raman data of healthy tissues shows two the most intense bands at 1160 cm^{-1} due to vibration of carotenoids and at 1519 cm^{-1} assigned to carotenoids or ring breathing DNA bases. Other, less intensive bands are gathered in Table 1. In Fig. 1SA, showing 2800–3020 cm^{-1} region, revealed bands can be assigned to $\upsilon(CH_2)$ and υ_{as} (CH_2) modes vibration of lipids and proteins (2888 cm^{-1}), υ_{as} (CH_2) vibrations protein aromatic and aliphatic amino acids (2934, 2932 cm^{-1}) and vibration of υ(CH) in the nucleic acids and proteins (3074 cm^{-1}) (Table 1).

Additionally, on the basis of the component analysis performed with the WITec software we created a Raman map shown in Fig. 2B. The same color is maintained for images and for spectra indicated using built-in component analysis. As can be seen on, within gathered data four different component of average spectrum are found. The main difference among them are due to modes appeared in the fingerprint region (600-1800 cm^{-1}). The most intensive bands at 1304, 1460, 1524 cm^{-1} and 1600 cm^{-1} are calculated for the component 1 and 2. However, the main differences between component 1 and component 2 are due to 1159 and 1523 cm^{-1} modes. The intensity of those two bands are increased in component 2. Presented green (component 3) and bright blue spectrum (component 4) do not allows for appropriate assignment of observed bands due to its very small Raman respond. Similarly, for the region 2800–3020 cm^{-1} calculated component 1 and 2 have the greatest influence on the observed difference (Fig. 1SB). In this case, it should be noted, that to visualize main difference, that is intensity of obtained bands and the variety of obtained spectra data are presented as average plots with the standard error deviations presented in the positive direction calculated for analyzed tissues samples. This is because already only one direction indicates that the deviation is too great to consider the averaged spectra of the neoplastic tissues as flawless. Thus, it should be concluded, the analysis of the differentiation between healthy and neoplastic tissues in this region (2800–3020 cm^{-1}), based on the collected tumor spectra, is not reliable. Moreover, in the case of RI, the collected spectra come from the surface of the tissues, i.e. from the surface and inside of the cells (sample prepared as tissue cross-section). Therefore, it should be in mind that the presented picture averaged the differences between the compounds that build cell membranes, and not even more specific differences between cells in the studied area, which could be used to distinguish neoplastic tissues from healthy ones. Therefore, despite the expected spectral differences between the cell membrane and the cell content, we mainly see vibrations attributed to the components of the cell membrane. Therefore, these data could not be used to distinguish neoplastic from healthy tissues.

While Raman imaging, especially in combination with component analysis is a sensitive method that can provide detailed information about the chemical composition (proteins, lipids, carbohydrates) that builds the cell wall due to its location, it is also

Table 1
Assignments of the bands observed in RI and SERS spectroscopy [14,33–41].

Raman bands	SERS bands	Compound/Assignments
	623	Phenylalanine (skeletal)
	650	C–C twisting mode in tyrosine
	680	Guanine (DNA)
	726 (t), 725	DNA, Lipids, Tryptophan
	758, 856(t),	Tryptophan
	888	Protein (collagen)
960	959	CH_3 deformation: lipids, proteins or carotenoids
1005	1001	Phenylalanine (ring breathing mode)
	1063 (t), 1067	ν(CC/CN) proteins or C–C skeletal stretching of lipids, Fatty acids
1080	1085, 1095 (t)	ν(PO_2): DNA, phospholipids
	1127	ν (C–N) of proteins or ν (C–C) lipids
1160		Carotenoids
	1244	Amide III, RNA, keratin
	1286 (t)	L-Tryptophan, Amide III (α-helis)
	1330	Phospholipids or Nucleic acids
	1381 (t)	δ(CH_3) of lipids
1445	1450 (t), 1460	CH_2 bending in proteins and lipids, keratin, fatty acids, triglycerides, CH_2, CH_3 deformation/lipids/proteins C–H wag.
	1468	Lipids
1519		Carotenoids or ring breathing DNA bases
1552		ν(C=C) tryptophan vibrations
	1589	Nucleic acids (guanine, adenine), tryptophan
	1608 (t)	Phenylalanine, tyrosine, cytosine,
1666	1691	C=O stretching Amide I,
2888	2888	Fatty acids, triglycerides υ(CH_2), υ_{as} (CH_2) of lipids and proteins, υ(CH_3), Amide I lipids,
2944	2954	υ_{as} (CH_2) fatty acids or lipids/protein, (CH_3) protein aromatic and aliphatic amino acids,
	3070	υ(CH), nucleic acids, proteins, lipids υ(CH),

t – indicate vibration shifting or bands observed only in spectra of tumor homogenates.

limited to healthy tissues. Therefore, it is not an appropriate tool to distinguish between healthy and cancerous or infected samples. The SERS method can fill this gap. Until now, the great challenge was to prepare a sample for SERS measurements. The newly developed method of sample tissues preparation in the form of a homogenate, shown here, significantly improves the intensity of the observed SERS bands, influencing the sensitivity of the method and allows for spectral differentiation between studied samples based on Raman analysis.

3.3. Development of the homogenization protocol for SERS measurements

Briefly, three methods of tissues homogenization, with appropriate modifications, serve as a single sample preparation process for measurement on SERS platforms. First step, the temperature homogenization is used (−80 °C), in the isolation of proteins resistant to denaturation, and is based on the increase in the volume of water freezing in the cytoplasm, which finally causes damage to the cell wall and membrane as well as cell organelles. Second step is homogenization based on the enzymatic digestion process using solvents and/or detergents and protease inhibitors. And third step is ultrasonic homogenization based on the tissues residues grinding.

The protocol of the tissues homogenization process [42] presented on Fig. 2S (Supplementary Materials) is performed as follows:

I) **temperature homogenization.** The collected tissue preparations are placed in small 1.5 mL aliquots and after precooling to 0 °C, frozen and kept at −80 °C. The use of temperature homogenization enables the destruction of tissue cells and the release of cell DNA and their individual components. This method of tissue homogenization also causes the release, mainly from lysosomes, of proteolytic enzymes. Thus, there is a need to protect tissue proteins from the action of these enzymes by the use of protease inhibitors. Therefore, the second phase of homogenization, which includes the addition of inhibitors, should take place as soon as possible - immediately after thawing the tissues.

II) **enzymatic homogenization.** The solution for dissolving the cell walls (lysis) is prepared just before the homogenization process. The homogenization/lysis solution contains:
- RIPA buffer (radioimmunoprecipitation test) consisting of ionic and non-ionic organic solvents (50 mM Tris-HCl, 150 mM NaCl, 0.1% SDS, 1% by weight octylphenoxypolyethoxyethanol and 0.5% by weight sodium deoxycholate, pH 8.0).
- enzymatic inhibitor of serine and cysteine proteases, 1 mM phenylmethyl sulfonate fluoride PMSF.
- additionally contains O-Complet Mini serine and cysteine protease inhibitors (inhibitors: 0.02 mg mL^{-1} of pancreatic digestive enzyme, 0.0005 mg mL^{-1} thermolysin, 0.002 mg mL^{-1} and 0.02 mg mL^{-1} chymotrypsin, 0.02 mg mL^{-1} and 0.002 mg mL^{-1} trypsin, 0.33 mg mL^{-1} papain enzyme). The best concentration according to the recipe on the package - 1 tablet of O-Complet Mini per 10 mL of RIPA buffer.

All tissue operations are performed at reduced temperature, most preferably in ice water at 0 °C. The solid tissue should weigh about 20 mg. The tissues were placed in 1.5 mL disposable containers. The lysis solution is used in an amount of 1 mL and added to the tissue container. The enzymatic homogenization process - step two - is carried out for 15–40 min and most preferably for 30 min.

III) **ultrasonic (sonication) homogenization.** For the sonication process the Qsonica sonicator were used. The sonication process is carried out by immersing the containers with the tissue preparations in an ice-water solution, most preferably at 0 °C, for 2 min. The medium (tissues and a buffer) can absorb energy and during the sonication process it can generate heat which will cause the temperature to rise. If the

heat generation is greater than the free heat, the elevated temperature depends on destruction. Therefore, it is usually necessary to determine the minimum amplitude (or amount of power applied) to pierce the cuvette without overheating the sample. The duration of the process is also important. Hence, for the purposes of the SERS test, optimal conditions favoring the determination of the lowest possible power (by adjusting the amplitude) that prolongs the decay (and not its destruction) and the penetration of the buffers used, in the case of obtaining insufficient Raman price with possible extension of time. The use of an amplitude of 50% (180 ± 30 W) indicated a correction of the tissue homogenate after continuous 2 min sonication, without repeating step III, in order to find the term of the homogenate giving an intense Raman signal. To complete the preparation of tissue homogenates, the samples were centrifuged in order to separate the obtained homogenate (supernatant) solutions from the tissue debris. Centrifugation was performed at 10,000×*g* for 20 min. It is most preferred to repeat the centrifugation process on the obtained supernatant. During the first centrifugation, the supernatant is separated from the tissue debris, the separated supernatant is centrifuged again and the supernatant is separated from the tissue debris again. The second supernatant is used for SERS measurements. The obtained homogenates can be frozen for further use. It is recommended to freeze the homogenates at two temperatures of −20 °C (for measuring samples in the shortest possible time) and −80 °C (for later use).

In order to use tissue homogenates and obtain maximum enhancement of the Raman signal on the SERS substrates, the SERS analysis should be performed immediately after preparing homogenates.

Additionally, the efficiency of the enzymatic homogenization process (step **II**) was tested in 20 min, 30 min and 40 min. The obtained SERS spectra were collected for the obtained homogenates prepared with different use of this process time (Fig. 3S). In the examples of SERS spectra, the number of cps (counts per second) of the Raman signal for the most intense band 1450 cm^{-1} is marked. Gained intensity of the observed vibrations affects the resolution of the obtained spectral data (SERS spectra). Higher intensity of the bands allows more precise analysis and identification, and thus has a positive effect on the sensitivity of the method used. The calculated cps (counts per second) values are shown in Table 1S. Therefore, for further enzymatic homogenization process, as the optimal time was selected 30 min. The enzymatic homogenization time of 20 min resulted in a relatively lower intensity of the observed band, and the extension of the time to 40 min did not improve the intensity of the 1450 cm^{-1} band. Then, the SERS spectra were recorded for tissues homogenates of both, healthy and tumor samples prepared according the protocol (Fig. 3).

3.4. SERS measurement of salivary gland homogenates

As can be seen, by using tissue homogenates prepared in accordance with the protocol (Fig. 2S), good quality SERS data were recorded for homogenates prepared from neoplastic tissues (Fig. 3, Fig. 4S) in comparison to the data recorded with RI (Fig. 2A, Fig. 1SA). Generally, any difference in the intensity of all bands observed in SERS of healthy and neoplastic tissues homogenates may serve as tool to distinguish tumor from healthy samples. However, at first glance, in the fingerprint 600-1800 cm^{-1}, spectral region the SERS spectra from both types of homogenates are almost identical (Fig. 3A). The spectrum representative for neoplastic homogenates shows shifting of two bands 1095 cm^{-1} (ν(PO_2): DNA,

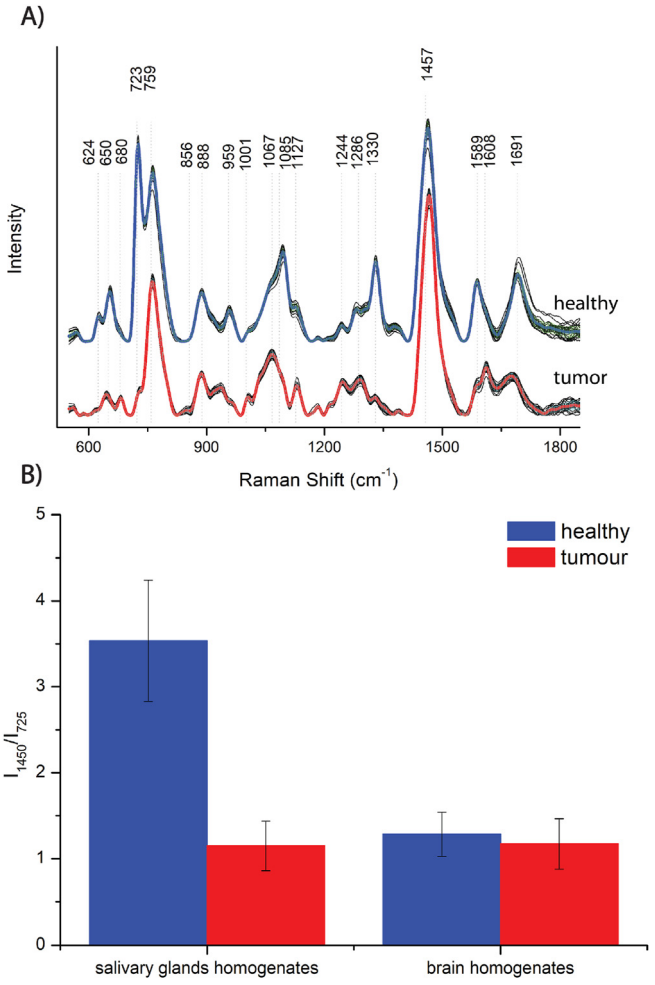


Fig. 3. The representative SERS spectra of homogenates prepared from healthy and tumor salivary gland tissues (**A**) and the column chart showing the calculated I_{1450}/I_{725} ratio for representative SERS spectra of homogenates prepared according to presented protocol and without step I and III (**B**). The error bars (green for healthy and cyan for tumor homogenates) represent the standard deviation. Data presented for the 600-1800 cm^{-1} region. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

phospholipids) and 1063 cm^{-1} (ν(CC/CN) proteins or C–C skeletal stretching of lipids and fatty acids). Moreover, spectrum of tumor homogenates, in comparison to spectrum of healthy homogenates, reveals two additional bands at 1286 cm^{-1} (L-Tryptophan, Amide III of α-helis and 1381 cm^{-1} δ(CH3) of lipids). In Fig. 4S, that shows the data of 2800–3020 cm^{-1} region, the revealed bands can be assignment to fatty acids, triglycerides υ(CH_2), $υ_{as}$ (CH_2) of lipids and proteins (2888 cm^{-1}), $υ_{as}$ (CH_2) vibrations of fatty acids or lipids/protein (2934, 2954 cm^{-1}) and vibration of υ(CH) in lipids (3070 cm^{-1}).

It should be mentioned that the bands intensity observed in this area is quite large compared to the intensity of the bands in the fingerprint area and more, this area concerns only a maximum of four vibrations. Fingerprint region is characterized by many vibrations that help identifying potential marker bands. Therefore, and due to the fact that, to understand the differences in biochemical relationships between tumor tissues and healthy tissues is crucial, further MVA analysis will be performed only for fingerprint area. Assignments of all observed in SERS spectra modes are gathered in Table 1.

The most important thing to note is the observation of the changes in the intensity bands in the SERS data recorded for the

prepared homogenates (Fig. 3) according to the protocol (Fig. 2S) in relation to the corresponding intensity bands of the homogenates of brain tissues prepared without steps I and II [18]. As we have already published the brain tissues homogenates SERS data herein we compare the intensity of recorded SERS signal in relation to the homogenates preparation for both samples – brain and salivary gland tissues. Therefore to extend the presented analysis, two bands were selected for further deduction, including 725 cm^{-1} band of L-tyrosine and 1450 cm^{-1} band characteristic for L-tryptophan, which at the same time have the highest intensity in the obtained SERS spectra. The calculated ratio of the intensity of the two selected bands, 725 cm^{-1} (CH_2 oscillation, symmetrical breathing, L-tyrosine) and bands 1450 cm^{-1} (CH deformation vibration, L-tryptophan) depending on the method of tissue sample preparation is summarized in Table 2S (Supplementary Materials).

The calculated ratios of the intensity of the bands (I_{1450}/I_{726}) corresponding to the vibrations of L-tyrosine and L-tryptophan in the examined tissues (Tab. 2S) show that in the case of tissue preparation according to the homogenization method used herein (Fig. 3B), high coefficients were obtained of the dependence of the intensity of the bands on the type of homogenate tested. On the other hand, for homogenates prepared without Steps I and III (Fig. 5S), the calculated band intensity ratio is nearly equal for both healthy and neoplastic tissue homogenates (Tab. 2S).

The obtained data show that the tissues of the homogenates prepared in accordance with the method give SERS spectra characterized by increased intensity and resolution of spectral images, which in turn increases the sensitivity of the SERS method to distinguishing between healthy and neoplastic tissue. The increase in resolution is reflected in the calculated ratios (healthy tissue vs cancer 3:1, 1:1), presented in Table 2S. In other words - the preparation of homogenates in accordance with the method allows to obtain such a resolution observed vibration bands that the usually observed dependencies of the intensity of selected bands 1:1 differ in the ratio of 3:1. Thus, when preparing the tissues of homogenates in accordance with the method, we will obtain SERS spectra, allowing for the assessment and unambiguous differentiation of the tested homogenates, and thus the identification of healthy and neoplastic tissues.

To conclude, it is not possible to gain any reasonable Raman signal from neoplastic tissues using Raman spectroscopy imaging. Therefore, for tumor tissues analyzes only the SERS techniques are possible. To perform SERS measurements of tissues samples and to gain satisfactory intensity of observed signal the protocol of tissues homogenates preparation was elaborated. The recorded SERS spectra for the prepared salivary gland homogenates allow for a detailed analysis of the observed spectra, which ultimately allows to find a relationship between the two bands (I_{1450}/I_{726}), and thus to distinguish between homogenates prepared from healthy and neoplastic tissues. Such effect was possible to observe due to both the signal enhancement in SERS method and homogeneity of used for study materials, i.e. lack of decreasing the laser intensity during passing through thicker cellular components.

Additionally, to find more adequate tool for medical purposes, based onto gathered SERS data the chemometric analysis in a form of principal component analysis (PCA), principal component regression (PCR), and partial least square – discriminant analysis (PLS-DA) were applied.

3.5. Multivariate analysis (MVA)

In the beginning it should be mentioned, that for some of the samples, we received very weak Raman signals, which made impossible to take them for full analysis. Therefore, presented MVA are based on 9 samples from sick patients and 8 samples from healthy patients.

Firstly, to reduce the dimensionality of the original raw data to several principal components (PCs) and to visualize the existing differences among the spectra of different homogenates, the principal component analysis (PCA) were applied. PCA is not generally regarded as a clustering method, but can be used for “indirect” and visual clustering. It is an effective technique that gives the possibility to categorize SERS spectra that are readily distinguishable via visual empirical analysis. The calculated PCs contain the most significant information from the whole introduced data set.

The PCA data calculated for healthy and carcinoma or neoplastic salivary glands homogenates in the fingerprint 500-1800 cm^{-1} region is presented on Fig. 4. The analyzed spectral data are presented on the basis of the first three principal components. As it is presented on the PC1 vs PC2 plots (Fig. 4A) the results obtained from both analyzed groups, data based on the spectra of healthy and neoplastic homogenates, are slightly mixed with each other. The first component (PC1 = 37%) has the greatest influence on the obtained calculations and presented results, followed by PC2 = 21%, and finally PC3 is responsible for 13% of the difference between these groups. Other following four important components, which are necessary for accurate differentiation among samples are presented in Fig. 6S (Supplementary Materials; PC4 = 7%, PC5 = 6%, PC6 = 4% and PC7 = 3%). Therefore, based on the PCA calculation performed for the spectral fingerprint region, and using following seven PCs it is possible to determine 91% of spectral variance between the healthy and tumor homogenates. Similarly, PCA calculation based on the spectral SERS data of carcinoma and healthy homogenates (Fig. 4B) indicate importance of three first components, as PC1 = 59%, PC2 = 24% and PC3 = 6% respectively, explain variance between the analyzed groups (89% in total). Good differentiation is obtained also between the scores calculated for pleomorphic adenoma(tumor) and salivary duct carcinoma (according WHO, G3) (Fig. 4C). The most important variance is already obtained along X axis of PC1 (49%). Thus three PCs explain 81% of variance between the tumor (adenoma polymorphum) and salivary duct carcinoma (G3) samples.

Additionally, to extend this analysis it is important to observe the loadings plots of those first three components. It is possible to determine which variables contributed the most to the presented determination. All those data are gathered in Fig. 7S and Table 3S, Supplementary Materials (subsection **A** - healthy and tumor, **B** - pleomorphic adenoma (tumor) and salivary duct carcinoma, **C** - pleomorphic adenoma (tumor) and salivary duct carcinoma).

Additionally, we were lucky to get one sample with co-existence two different tumor, thus we analyzed the whole SERS data as two sets:

1) samples with pleomorphic adenoma and adenolymphoma and with both tumors (pleomorphic adenoma and adenolymphoma), Fig. 5A;
2) sample with adenolymphoma and with co-existence of two tumors (adenolymphoma and adenoma polymorphum), Fig. 5B.

Presented data explain in total 79% of variance between the pleomorphic adenoma and adenolymphoma samples, including the sample with both tumors (Fig. 5A). Scores calculated for this particular tumor-mixed sample appears between the scores of other samples, what is especially nicely visualized on 3D projection of PCs scores. To strengthen of this analysis the scores calculated for homogenates of adenolymphoma and for sample with co-existence of two tumors are presented on Fig. 5B. The PCA analysis allows for determination of 96% of variance between the analyzed groups in three PCs. Moreover, the scores of sample with two tumors (adenolymphona and adenoma pleomorphum) are divided in two groups by X-axis of PC1, while the results calculated for the tumor

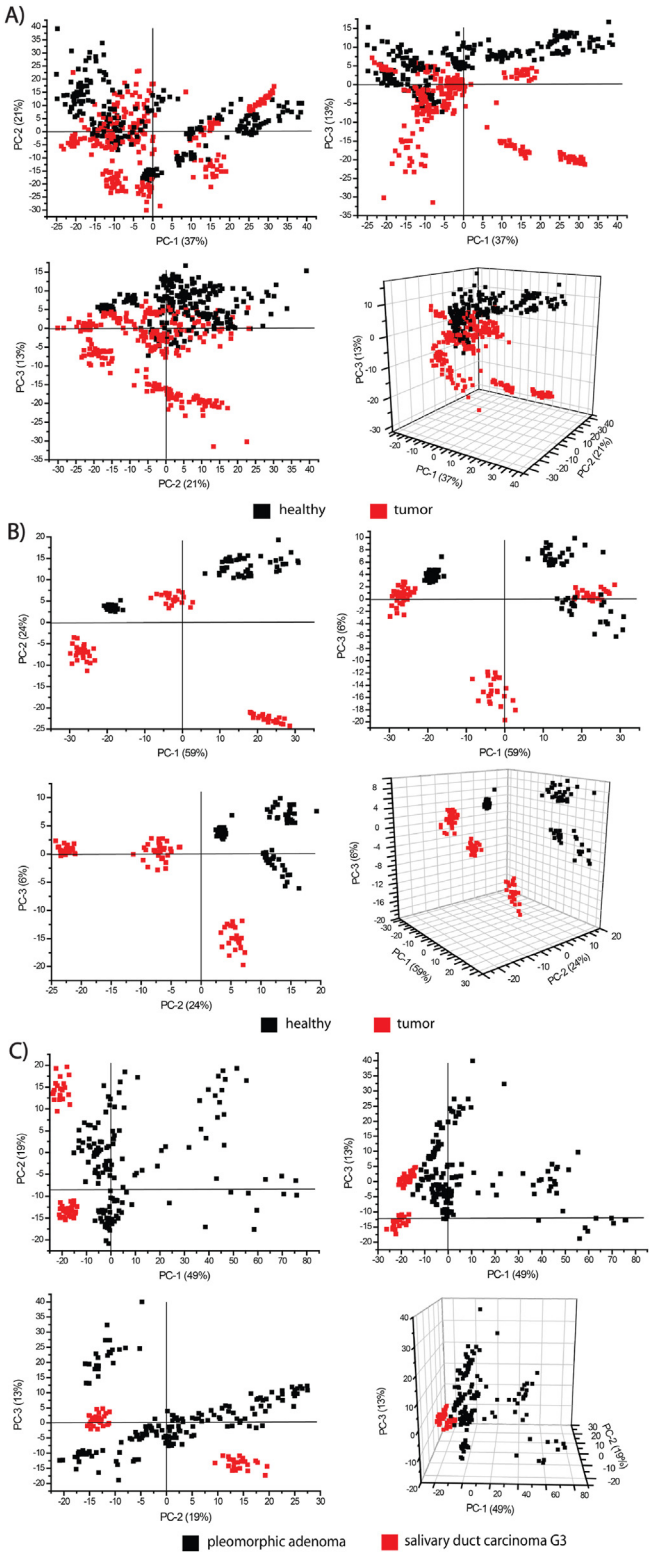


Fig. 4. Calculated scores, in 2D and in 3D projection, are presented as plots: PC1 vs. PC2, PC1 vs PC3 and PC2 vs PC3 for healthy and tumor (**A**) and healthy and carcinoma samples (**B**) but also pleomorphic adenoma (tumor) and salivary duct carcinoma (according WHO, G3) (**C**). PCA analysis based on the SERS data gathered for the fingerprint region.

adenolymphoma homogenates are grouped on left side of this axis. Therefore, taking into account the origin of the samples and the position of the points on the graphs, it can be concluded that the graph (Fig. 5) visualizes two different states of the tested mixed sample. And indeed for that samples according to histopathological measurement two types of tumor co-exist - first adenolymphoma and the second pleomorphic adenoma. Therefore, the presented SERS with PCA can be used as a method of differentiating healthy, neoplastic and samples with co-existing tumors in one tissues as a competitive tool in relation to the commonly used method of histopathological staining of tumor tissue. The loadings of the first three components for pleomorphic adenoma (tumor) adenolymphoma and sample with both these tumors are shown Fig. 7SD, and Table 2SD, for the sample with co-existence of two different tumors: adenolymphoma and poorly differentiated cancer in the lymph nodes (Fig. 7SE, and Table 2SE).

Then, to categorize and quantify the unknown spectrum of the salivary gland sample, as healthy or neoplastic one, the additional MVA methods, such as partial last square – discriminant analysis (PLS-DA) and principal component regression (PCR) were applied over the SERS data. Both methods PLS-DA and PCR are regression methods which predict a sample as belonging to given group of category based on prediction values close to zero or one and can lead to fruitful prediction of the outcome on the basis of a complex model by appropriately selecting the major components used for regression. Therefore, the SERS spectra were modelled by both PLS-DA and PCR algorithm to classify the obtained SERS spectra, as coming from healthy or neoplastic salivary gland homogenates. A major difference between PCR and PLS; are related to the scores of the PCA used in PCR regression, that best explain the data X (the projection space only depends on X). In PLS the projection space of X is explains both values X and Y. For this reasons the PLS usually achieves the same result as PCR with lower number of latent variables [43,44]. It should be pointed, that in PLS-DA method, the observed separation of any groups is usually optimistic in comparison to scatter plots calculated using PCA method. That is because, for example, random noise is sometimes used to make any additional determination between groups. For this reasons after applying a supervised method of analysis, such as PLS, there is always a validation to make sure that the observed differences are significant and generalize well to new data. The simplest form of validation consists in dividing the data table into a calibration set and a prediction set. Herein, all SERS data, from both tumor and healthy homogenates, were divided into a calibration set and a prediction set in approximately 4:1 ratio (a calibration set- 11 samples from 6 sick and 5 healthy patients and a validation set- 6 samples from 3 healthy and 3 sick patients; 20 spectra for each sample, total 340 spectra). The training set was used to build the model, and a calibration set was used to evaluated obtained model by the root mean square errors of calibration and prediction (RMSEC, RMSEP), the coefficients of determination of calibration and prediction (R2C, R2P). Based on those data the receiver operating characteristic (ROC), the correlation accuracy (AUC the area under the ROC curve), sensitivity, and specificity were also calculated for presented PCR and PLS-DA models.

The number of components was selected on the basis of empirical analysis of the obtained PCR and PLS-DA data. It is possible to detect overfitting of variables by analyzing the load plots of each PC or factors, as well as the course of the characteristic curve of the explained variance vs. variance Y. Choosing too few components gives a not good enough model, but selecting too many components is also not good due to providing a noise-sensitive model [45]. In this light, for the PCR method, six PCs, while for the PLS-DA, four factors were chosen and presented. The calculated PCR data are shown in Fig. 8S (Supplementary Materials) in 2D and 3D scatter plots of six consecutive PCs. In presented data scores PC-1 explains 36% of variance in block X, while PC-2 explain

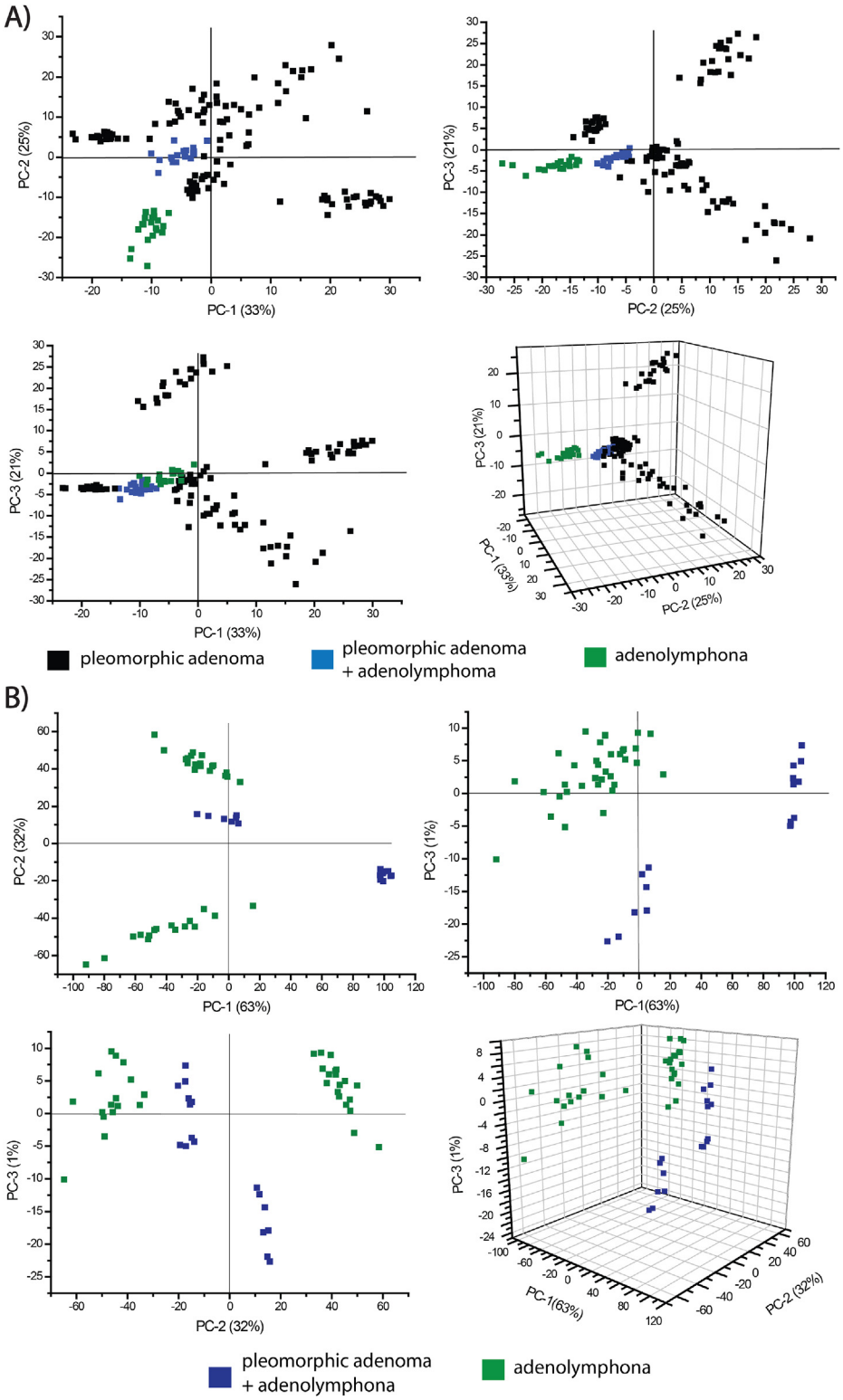


Fig. 5. Calculated scores, in 2D and in 3D projection, are presented as plots: PC1 vs. PC2, PC1 vs PC3 and PC2 vs PC3 for pleomorphic adenoma (tumor) adenolymphoma and sample with both these tumors (**A**). Scores PC1 vs PC2 for the sample with co-existence of two different tumors: adenolymphoma and poorly differentiated cancer in the lymph nodes (**B**). PCA analysis based on the SERS data gathered for the fingerprint region.

22%, PC-3 explain only 13%, PC-4 explain 10%, PC-5 explain 5%, and the last PC-6 explain 4%. To summarize, the total of 90% variance is explained by the first six components identified by PCR. Then, the presented calibration model was evaluated by the root mean square errors of calibration and prediction. It should be in mind, that the smaller the value of RMSE, the quantitative model is the higher quality. Similarly, with the smaller the value of RMSEP, the stronger the predictive ability of model is achieved. In the

calculated and presented model RMSEC = 0.156 and RMSEP = 0.160. The coefficients of determination of calibration and prediction in square roots are R2C = 0.90 and R2P = 0.94. As can be seen from the predictive values are very close to 1 and a small value of standard ensure that the model is robust which provide a correct classification. All those data are gathered in Table 2.

The calculated PLS-DA data are shown in Fig. 6 in 2D and 3D scatter plots of three consecutive factors. This model based onto three latent variables shows, for F-1 70% of variance in block Y with 24% of the spectral data (X matrix), while F-2 explain 19% of variance in block Y with 23% of analyzed data, F-3 explain 3% with 21% of X, and the last F-4 explain only 3%, with 6% of the data within X matrix. The first four factors identified by PLS-DA explain the total of 95% variance. In the calculated and presented model RMSEC = 0.113 and RMSEP = 0.119. The coefficients of determination of calibration and prediction in square roots are R2C = 0.95 and R2P = 0.94.

The data presented in Table 2 reveal that the total variance for six following PCs (PCR) or four factors (PLS-DA) have values of 90% and 95%, respectively. Therefore, to select the appropriate model to distinguish between healthy and neoplastic or carcinoma salivary glands homogenates it is necessary to take into account values by the root mean square errors of calibration and prediction (smaller values of RMSEC, RMSEP), the coefficients of determination of calibration and prediction (values of R2C, R2P closer to 1). From the presented data, the PLS-DA model is more robustness, feasible and efficient in categorizing data of unknown origin as spectra of healthy or neoplastic salivary glands based on SERS.

It must be concluded that SERS spectroscopy together with MVA analysis could be successfully used to discriminate neoplastic among healthy salivary glands samples. At this point it should be added – as to comparing three types of MVA methods (PCA, PLS-DA, PCR), evidently the PCA model indicates its applicability with a small number of samples, but due to the fact that PCR enables mathematical classification, prediction and quantitative analysis, it should find wider application while using a larger number of biological or clinical samples.

However, since PCR and PLS-DA give an additional matrix containing the predicted and correlated values, it is possible to build a ROC curve. The area under the curve (AUC) makes it possible to determine the correlation accuracy, specificity, and sensitivity to every possible level of probability of significance. Therefore, the ROC for all cutoff values and the correlation accuracy (Fig. 9S; Supplementary Materials) are determined and presented for both PCR and PLS-DA methods. Based on the presented data the calculated correlation accuracy for PCR is 0.71 (71%), while for PLS-DA method is 0.98 (98%). The optimal cut-point identified from the ROC curve had a sensitivity of 0.70 (70%) or 0.97 (97%) and specificity of 0.53 (53%) or 0.89 (89%) (for PCR and PLS-DA, respectively; Fig. 9SA and Fig. 9SB).

The most weighted and with the greatest influence on the obtained discrimination loadings of the first three components for both PCR and PLS-DA methods are presented on Fig. 7SF and Fig. 7SG in Supplementary Materials. The gathered data prove that in case of PLS-DA variables at 1452 cm^{-1}, 1604 cm^{-1} and at 758 cm^{-1}, 1075 cm^{-1}, 1592 cm^{-1} and 770 cm^{-1}, 1478 cm^{-1}, 1606 cm^{-1} for Factor-1, Factor-2 and Factor-3, respectively (Table 3SF) are the most important. Those variables have direct correlation with marker bands observed in SERS spectra.

Table 2
The data obtained from PCR and PLS-DA multivariate analysis.

MVA method	RMSEC	RMSEP	R2C	R2P	Total variance [%]
PCR (6 PCs)	0.156	0.160	0.90	0.90	90
PLS-DA (4 Fs)	0.113	0.119	0.95	0.94	95

Additionally, the PCA, PLS-DA and PCR analysis of gathered SERS spectra for the presented herein protocol of homogenates preparation, proved that unmistakeable determination of the marker bands and the differentiation of healthy and neoplastic salivary glands sample is possible even when considering only a thorough analysis of the characteristics of the spectral features observed in SERS. Those marker bands appeared as most weighted variables in MVA and have corresponding shifts of Raman bands in relation to the studied sample (neoplastic or healthy). As some of them can be observed with shifting or are observed only for tumor samples, can be treated as marker bands. Those bands revealed at 1608 cm^{-1} (t), 1468 cm^{-1}, 1127 cm^{-1}, 1063 cm^{-1} (t)/1067 cm^{-1}, 1085 cm^{-1}/1095 cm^{-1} (t), 856 cm^{-1} (t), and at 758 cm^{-1}. These bands are due to modes vibration of tryptophan (856 cm^{-1} (t), 758 cm^{-1}), protein or lipids (1063 cm^{-1} (t), 1067 cm^{-1} and 1127 cm^{-1}), lipids (1468 cm^{-1}), DNA, phospholipids (1085 cm^{-1}, 1095 cm^{-1} (t)) and phenylalanine or DNA (1608 cm^{-1} (t)).

To summarize, combing the SERS with PCA analysis the seven diagnostic bands were found. It should be highlighted, the more diagnostic bands, the better understanding the biochemical relationships in tumor tissues. Spectral images are directly related to biochemical changes found in the neoplastic cells. It is obvious, that the area with CH vibrations does not give such opportunities, they are vibrations that can come from a number of compounds present in the tissues. For better understanding the issues how the biochemical pathway in cancer cells is different and how it translates into a spectral image in relation to healthy cells, the most promising is an in-depth fingerprint area analysis.

4. Conclusions

In conclusion, it is not possible to obtain any reasonable Raman signal from tumor tissues by using Raman Imaging spectroscopy. As was presented, only SERS techniques are feasible for the analysis of neoplastic tissues. In order to perform SERS measurements of tissue samples and obtain a satisfactory intensity of the observed signal, a protocol for the preparation of tissue homogenates was developed. The recorded SERS spectra in the fingerprint region (600-1800 cm^{-1}) allow for a detailed analysis of the observed modes, which help to establish a relationship between the two bands (I_{1450}/I_{726}), and to distinguish between homogenates prepared from healthy and neoplastic tissues.

Additionally, SERS data were analyzed with the PCA algorithm, and in results showed that explain 91% of variance among the data, while the PCR and the PLS-DA methods explain 90% and 95% of the variance between the study groups (identified by the first seven components by PCA, six components by PCR and four factors of PLS-DA, respectively). The obtained model of PCR on the basis of the mean square errors of calibration and prediction (RMSEC = 0.156 and RMSEP = 0.160) and the coefficients of determining the calibration and prediction (R2C = 0.90 and R2P = 0.90) show the less accuracy when applied to the studied data. While the resultant model of PLS-DA gives values of RMSEC = 0.113 and RMSEP = 0.119 and R2C = 0.95 and R2P = 0.94, *ipso facto* prove the robustness, convenience, feasibility, and the efficiency, of the applied model, in categorizing data of unknown origin as a spectrum of healthy or neoplastic salivary glands based on SERS. Additionally, the ROC analysis performed based on PCA and PLS-DA data show correlation accuracy as 0.71 with sensitivity 0.70, specificity 0.53, and as 0.98 with sensitivity 0.97, and specificity 0.89, respectively. Based on that, the PCA model clearly shows its applicability with a small number of samples, but the PLS-DA enables mathematical classification and prediction, with good accuracy, sensitivity, and specificity. To summarize, PLS-DA analysis excellently works for homogenates of salivary gland tissues, and may especially work with a larger number of samples biological or clinical.

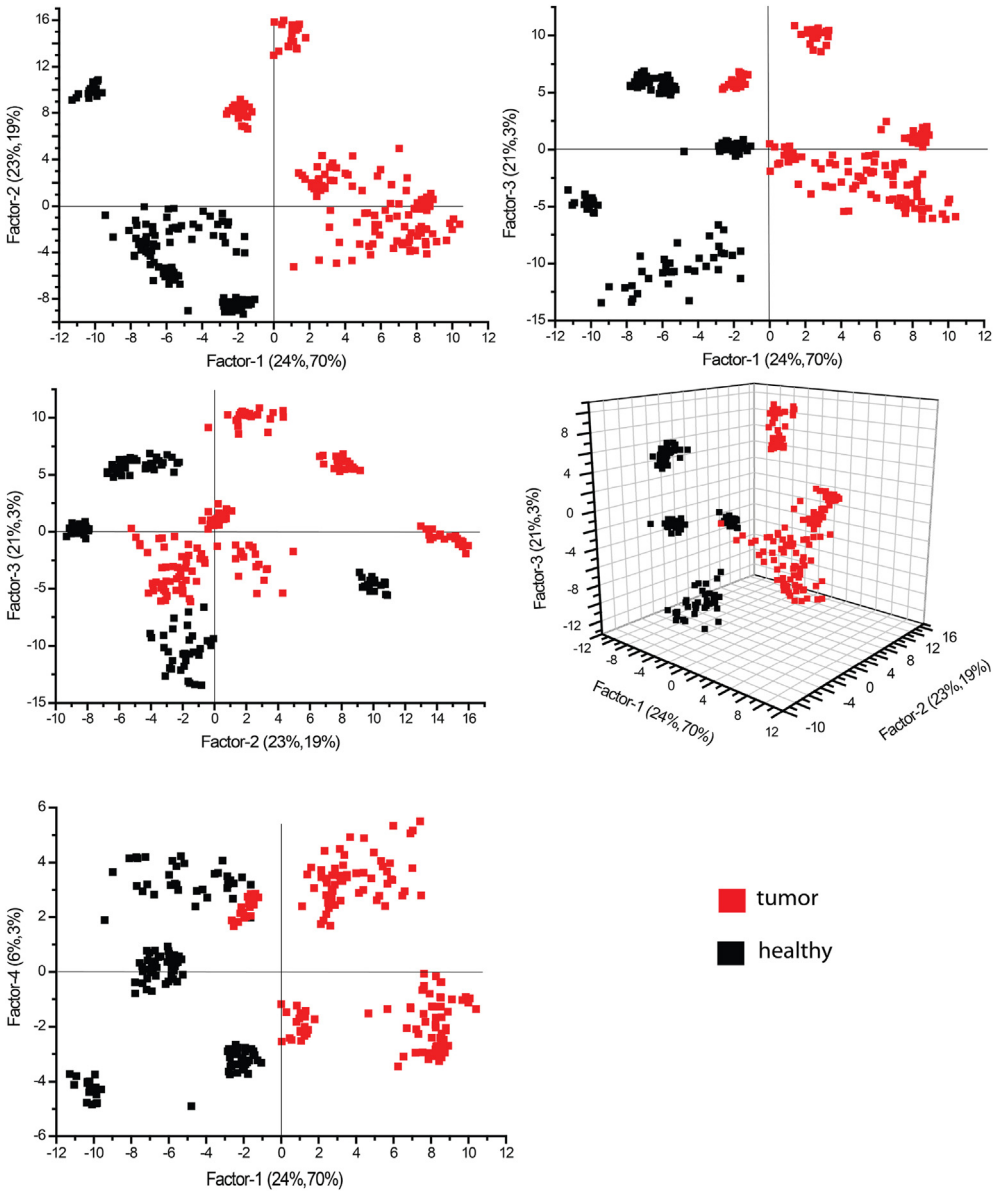


Fig. 6. Correlation between the predicted values and reference values by PLS-DA model in four following factors.

It should be highlighted, that presented herein combined methods (homogenates preparation, SERS and MVA) made possible to define, in the fingerprint region, seven diagnostic bands revealed in the SERS spectra at 1608 (t), 1468, 1127, 1063 (t)/1067, 1085/1095 (t), 856 (t) and 758, cm^{-1}. The bands observed in the spectra of the tumor homogenates (t) are shifted in relation to the corresponding bands in the spectra of healthy homogenates. Others only appeared for normal or tumor homogenates. Therefore, despite of the variety of cells forming the salivary gland tissues, the presented method has great potential in determining a clinical salivary gland sample and final recognition between healthy and neoplastic tissues.

CRediT authorship contribution statement

M. Czaplicka: Methodology, Software, Investigation. **A.A. Kowalska:** Writing – original draft. **A.B. Nowicka:** Software, Investigation. **D. Kurzydłowski:** Data curation, Methodology. **Z. Gronkiewicz:** Investigation, Validation. **A. Machulak:** Resources, Investigation. **W. Kukwa:** Conceptualization. **A. Kamińska:** Supervision, Reviewing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Authors would like to acknowledge the support from National Science Centre under grant UMO-2017/25/B/ST4/01109.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2021.338784.

References

[1] M. Stenner, J.P. Klussmann, Current update on established and novel biomarkers in salivary gland carcinoma pathology and the molecular pathways involved, Eur. Arch. Oto-Rhino-Laryngol. 266 (2009) 333–341, https://doi.org/10.1007/s00405-008-0882-7.

[2] P.M. Speight, A.W. Barrett, Salivary gland tumours, Oral Dis. 8 (2002) 229–240, https://doi.org/10.1034/j.1601-0825.2002.02870.x.

[3] L. Li, Y. Li, Y. Wen, H. Liu, H. Zhao, Clinical analysis of salivary gland tumor cases in West China in past 50 years, Oral Oncol. 44 (2008) 187–192, https://doi.org/10.1016/j.oraloncology.2007.01.016.

[4] Z. Tian, L. Li, L. Wang, Y. Hu, J. Li, Salivary gland neoplasms in oral and maxillofacial regions: a 23-year retrospective study of 6982 cases in an eastern Chinese population, Int. J. Oral Maxillofac. Surg. 39 (2010) 235–242, https://doi.org/10.1016/j.ijom.2009.10.016.

[5] M. Ziółkowska, S. Bień, S. Okła, S. Zyłka, Epidemiological and clinical characteristics of 705 salivary glands neoplasms, Otolaryngol. Pol. 67 (2013) 154–163, https://doi.org/10.1016/j.otpol.2013.03.002.

[6] I. Kordzińska-Cisek, L. Grzybowska-Szatkowska, Salivary gland cancer — epidemiology, Nowotw. J Oncol. 68 (2018) 22–27, https://doi.org/10.1007/978-3-540-85516-3_59.

[7] A.L. Carvalho, I.N. Nishimoto, J.A. Califano, L.P. Kowalski, Trends in incidence and prognosis for head and neck cancer in the United States: a site-specific analysis of the SEER database, Int. J. Canc. 114 (2005) 806–816, https://doi.org/10.1002/ijc.20740.

[8] https://www.cancer.org/cancer/salivary-gland-cancer/about/what-is-key-statistics.html. (Accessed 20 October 2020).

[9] Tumours of salivary glands, in: A. El-Naggar, J. Chan, J. Grandis, T. Takata, S. PJ (Eds.), WHO Classif. Head Neck Tumours, fourth ed., International Agency for Research on Cancer, Lyon, 2017, pp. 159–202.

[10] F.A. de Oliveira, E.C.B. Duarte, C.T. Taveira, A.A. Máximo, É.C. de Aquino, R. de Cássia Alencar, E.F. Vencio, Salivary gland tumor: a review of 599 cases in a Brazilian population, Head Neck Pathol 3 (2009) 271–275, https://doi.org/10.1007/s12105-009-0139-9.

[11] P. Sowa, K. Goroszkiewicz, J. Szydelko, J. Chechlinska, K. Pluta, W. Domka, M. Misiolek, W. Scierski, A review of selected factors of salivary gland tumour formation and malignant transformation, BioMed Res. Int. 2018 (2018), https://doi.org/10.1155/2018/2897827.

[12] R.H. Spiro, Salivary neoplasms: overview of a 35-year experience with 2,807 patients, Head Neck Surg. 8 (1986) 177–184, https://doi.org/10.1002/hed.2890080309.

[13] R.W. Illes, M.B. Brian, A review of the tumors of the salivary gland, Surg. Gynecol. Obstet. 163 (1986) 399–404.

[14] B. Brozek-Pluska, M. Kopec, I. Niedzwiecka, A. Morawiec-Sztandera, Label-free determination of lipid composition and secondary protein structure of human salivary noncancerous and cancerous tissues by Raman microspectroscopy, Analyst 140 (2015) 2107–2113, https://doi.org/10.1039/c4an01394h.

[15] F. Salaffi, M. Carotti, A. Iagnocco, F. Luccioli, R. Ramonda, E. Sabatini, M. De Nicola, M. Maggi, R. Priori, G. Valesini, R. Gerli, L. Punzi, G.M. Giuseppetti, U. Salvolini, W. Grassi, Ultrasonography of salivary glands in primary Sjögren's syndrome: a comparison with contrast sialography and scintigraphy, Rheumatology 47 (2008) 1244–1249, https://doi.org/10.1093/rheumatology/ken222.

[16] Y.S. Li, J.S. Church, Raman spectroscopy in the analysis of food and pharmaceutical nanomaterials, J. Food Drug Anal. 22 (2014) 29–48, https://doi.org/10.1016/j.jfda.2014.01.003.

[17] M. Fleischmann, Raman spectra of pyridine adsorbed at silver electrode, Chem. Phys. Lett. 26 (1974) 2–5, https://doi.org/10.1016/0009-2614(74)85388-1.

[18] A.A. Kowalska, S. Berus, Ł. Szleszkowski, A. Kamińska, A. Kmiecik, K. Ratajczak-Wielgomas, T. Jurek, Ł. Zadka, Brain tumour homogenates analysed by surface-enhanced Raman spectroscopy: discrimination among healthy and cancer cells, Spectrochim, Acta Part A Mol. Biomol. Spectrosc. (2019) 117769, https://doi.org/10.1016/j.saa.2019.117769.

[19] N. Stone, P. Matousek, Advanced transmission Raman spectroscopy: a promising tool for breast disease diagnosis, Cancer Res 68 (2008) 4424–4430, https://doi.org/10.1158/0008-5472.CAN-07-6557.

[20] A. Nijssen, T.C. Bakker Schut, F. Heule, P.J. Caspers, D.P. Hayes, M.H.A. Neumann, G.J. Puppels, Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy, J. Invest. Dermatol. 119 (2002) 64–69, https://doi.org/10.1046/j.1523-1747.2002.01807.x.

[21] S. Feng, R. Chen, J. Lin, J. Pan, Y. Wu, Y. Li, J. Chen, H. Zeng, Gastric cancer detection based on blood plasma surface-enhanced Raman spectroscopy excited by polarized laser light, Biosens. Bioelectron. 26 (2011) 3167–3174, https://doi.org/10.1016/j.bios.2010.12.020.

[22] J. Dybas, K.M. Marzec, M.Z. Pacia, K. Kochan, K. Czamara, K. Chrabaszcz, E. Staniszewska-Slezak, K. Malek, M. Baranska, A. Kaczor, Raman spectroscopy as a sensitive probe of soft tissue composition – imaging of cross-sections of various organs vs. single spectra of tissue homogenates, TrAC - Trends Anal. Chem. 85 (2016) 117–127, https://doi.org/10.1016/j.trac.2016.08.014.

[23] M. Heine, B. Freund, P. Nielsen, C. Jung, R. Reimer, H. Hohenberg, U. Zangemeister-Wittke, H.J. Wester, G.H. Lüers, U. Schumacher, High interstitial fluid pressure is associated with low tumour penetration of diagnostic monoclonal antibodies applied for molecular imaging purposes, PloS One 7 (2012), https://doi.org/10.1371/journal.pone.0036258.

[24] D. Cioranescu, P. Donato, An Introduction to Homogenization, 1999.

[25] M. Culha, A TISSUE DIFFERENTIATION METHOD BASED ON SURFACE ENHANCED RAMAN SCATTERING, vol. 72, 2009. WO 2009/122363 Al.

[26] M.S. Bergholt, W. Zheng, K. Lin, J. Wang, H. Xu, J.L. Ren, K.Y. Ho, M. Teh, K.G. Yeoh, Z. Huang, Characterizing variability of in vivo Raman spectroscopic properties of different anatomical sites of normal colorectal tissue towards cancer diagnosis at colonoscopy, Anal. Chem. 87 (2015) 960–966, https://doi.org/10.1021/ac503287u.

[27] W. Liu, Z. Sun, J. Chen, C. Jing, Raman spectroscopy in colorectal cancer diagnostics: comparison of PCA-LDA and PLS-DA models, J. Spectrosc. 2016 (2016) 1603609, https://doi.org/10.1155/2016/1603609.

[28] A. Osmani, M. Par, M. Škrabić, M. Vodanović, O. Gamulin, Principal component regression for forensic age determination using the Raman spectra of teeth, Appl. Spectrosc. 74 (2020) 1473–1485, https://doi.org/10.1177/0003702820905903.

[29] V.E. Sitnikova, M.A. Kotkova, T.N. Nosenko, T.N. Kotkova, D.M. Martynova, M.V. Uspenskaya, Breast cancer detection by ATR-FTIR spectroscopy of blood serum and multivariate data-analysis, Talanta 214 (2020) 120857, https://doi.org/10.1016/j.talanta.2020.120857.

[30] T. Szymborski, Y. Stepanenko, P. Piecyk, K. Nicinski, A. Kamińska, Sposób Wytwarzania Platformy Krzemowej Do Pomiarów Metodą SERS, Platforma Krzemowa Do Pomiarów Metodą SERS Oraz Zastosowanie Platformy Krzemowej Do Pomiarów Metodą SERS Próbek Biologicznych Albo Związków Chemicznych, 2020, https://doi.org/10.1016/j.solener.2019.02.027, 434300.

[31] A.G. Lewis, T. Tong, E. Maghami, Diagnosis and management of malignant salivary gland tumors of the parotid gland, Otolaryngol. Clin. North Am. 49 (2016) 343–380, https://doi.org/10.1016/j.otc.2015.11.001.

[32] H. Karabeber, R. Huang, P. Iacono, J.M. Samii, K. Pitter, E.C. Holland, M.F. Kircher, Guiding brain tumor resection using surface-enhanced Raman scattering nanoparticles and a hand-held Raman scanner, ACS Nano 8 (2014) 9755–9766, https://doi.org/10.1021/nn503948b.

[33] E. Staniszewska-Slezak, K. Malek, M. Baranska, Complementary analysis of tissue homogenates composition obtained by Vis and NIR laser excitations and Raman spectroscopy, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 147 (2015) 245–256, https://doi.org/10.1016/j.saa.2015.03.086.

[34] C. Zong, M. Xu, L.J. Xu, T. Wei, X. Ma, X.S. Zheng, R. Hu, B. Ren, Surface-enhanced Raman spectroscopy for bioanalysis: reliability and challenges, Chem. Rev. 118 (2018) 4946–4980, https://doi.org/10.1021/acs.chemrev.7b00668.

[35] Z. Huang, A. McWilliams, H. Lui, D.I. McLean, S. Lam, H. Zeng, Near-infrared Raman spectroscopy for optical diagnosis of lung cancer, Int. J. Canc. 107 (2003) 1047–1052, https://doi.org/10.1002/ijc.11500.

[36] C. Krafft, L. Neudert, T. Simat, R. Salzer, Near infrared Raman spectra of human brain lipids, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 61 (2005) 1529–1535, https://doi.org/10.1016/j.saa.2004.11.017.

[37] C.M. Krishna, G.D. Sockalingum, J. Kurien, L. Rao, L. Venteo, M. Pluot, M. Manfait, V.B. Kartha, Micro-Raman spectroscopy for optical pathology of oral squamous cell carcinoma, Appl. Spectrosc. 58 (2004) 1128–1135, https://doi.org/10.1366/0003702041959460.

[38] Y. Li, Z.N. Wen, L.J. Li, M.L. Li, N. Gao, Y.Z. Guo, Research on the Raman spectral character and diagnostic value of squamous cell carcinoma of oral mucosa, J. Raman Spectrosc. 41 (2010) 142–147, https://doi.org/10.1002/jrs.2421.

[39] S. Koljenović, T.B. Schut, A. Vincent, J.M. Kros, G.J. Puppels, Detection of meningioma in dura mater by Raman spectroscopy, Anal. Chem. 77 (2005) 7958–7965, https://doi.org/10.1021/ac0512599.

[40] N. Stone, C. Kendall, J. Smith, P. Crow, H. Barr, Raman spectroscopy for identification of epithelial cancers, Faraday Discuss 126 (2004) 141–157, https://doi.org/10.1039/b304992b.

[41] G. Chundayil Madathil, S. Iyer, K. Thankappan, G.S. Gowd, S. Nair, M. Koyakutty, A novel surface enhanced Raman catheter for rapid detection, classification, and grading of oral cancer, Adv. Healthc. Mater. 8 (2019) 1–9, https://doi.org/10.1002/adhm.201801557.

[42] M. Czaplicka, A.A. Kowalska, A.B. Nowicka, W. Kukwa, Z. Gronkiewicz, D. Kurzydłowski, A. Kamińska, Sposób Przygotowania Homogenatów, Zwłaszcza Tkanek Stałych, W Tym Nowotworowych Do Wykorzystania W Technice Powierzchniowo Wzmocnionego Efektu Ramana (SERS), Sposób Ich Wykonania, Oraz Ich Zastosowanie Z Użyciem Platformy SERS, 2020, p. 434777.

[43] B. Hemmateenejad, M. Akhond, F. Samari, A comparative study between PCR and PLS in simultaneous spectrophotometric determination of diphenylamine, aniline, and phenol: effect of wavelength selection, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 67 (2007) 958–965, https://doi.org/10.1016/j.saa.2006.09.014.

[44] H. Khajehsharifi, E. Pourbasheer, H. Tavallali, S. Sarvi, M. Sadeghi, The comparison of partial least squares and principal component regression in simultaneous spectrophotometric determination of ascorbic acid, dopamine and uric acid in real samples, Arab. J. Chem. 10 (2017) S3451–S3458, https://doi.org/10.1016/j.arabjc.2014.02.006.

[45] P. Geladi, E. Dåbakk, in: Encyclopedia of Spectroscopy and Spectrometry, second ed., 1999. Computational methods and chemometry in near infrared spectroscopy.