Vibrational spectroscopy – are we close to finding a solution for early pancreatic cancer diagnosis?

Krzysztof Szymoński, Łukasz Chmura, Ewelina Lipiec, Dariusz Adamek

Krzysztof Szymoński, Łukasz Chmura, Department of Pathomorphology, Jagiellonian University Medical College, Cracow 33-332, Poland

Krzysztof Szymoński, Łukasz Chmura, Dariusz Adamek, Department of Pathomorphology, University Hospital in Cracow, Cracow 31-501, Poland

Ewelina Lipiec, M. Smoluchowski Institute of Physics, Jagiellonian University, Cracow 30-348, Poland

Dariusz Adamek, Department of Neuropathology, Jagiellonian University Medical College, Cracow 33-332, Poland

Corresponding author: Krzysztof Szymoński, MD, Academic Research, Doctor, Researcher, Senior Researcher, Teaching Assistant, Department of Pathomorphology, Jagiellonian University Medical College, Grzegórzecka 16, Cracow 33-332, Poland. krzysztof.szymonski@uj.edu.pl

Abstract

Pancreatic cancer (PC) is an aggressive and lethal neoplasm, ranking seventh in the world for cancer deaths, with an overall 5-year survival rate of below 10%. The knowledge about PC pathogenesis is rapidly expanding. New aspects of tumor biology, including its molecular and morphological heterogeneity, have been reported to explain the complicated "cross-talk" that occurs between the cancer cells and the tumor stroma or the nature of pancreatic ductal adenocarcinoma-associated neural remodeling. Nevertheless, currently, there are no specific and sensitive diagnosis options for PC. Vibrational spectroscopy (VS) shows a promising role in the development of early diagnosis technology. In this review, we summarize recent reports about improvements in spectroscopic methodologies, briefly explain and highlight the drawbacks of each of them, and discuss available solutions. The important aspects of spectroscopic data evaluation with multivariate analysis and a convolutional neural network methodology are depicted. We conclude by presenting a study design for systemic verification of the VS-based methods in the diagnosis of PC.

Key Words: Spectroscopic cancer diagnosis; Raman spectroscopy; Pancreatic cancer diagnosis; DNA methylation; Liquid biopsy biomarkers; Convolutional neural networks

Core Tip: Vibrational spectroscopy (VS) may become a major player in the development of early diagnosis technology for pancreatic cancer. As with every technique, VS has promising attributes as well as drawbacks. We summarize recent reports about improvements in spectroscopic methodologies, briefly explain and highlight the drawbacks of each of them, and discuss available solutions. Additionally, the important aspects of spectroscopic data evaluation with multivariate analysis and a convolutional neural network methodology are depicted.

INTRODUCTION

Pancreatic cancer (PC) is a very aggressive and lethal neoplasm, ranking seventh in the world for cancer deaths[1]. In 2020 there was an estimated number of 466003 new PC-related deaths worldwide[2]. Despite a better understanding of the nature of pancreatic tumors, 5-year survival rates have not improved and remain below 10%. Late-stage disease at diagnosis is a significant issue that contributes to poor overall survival rates. Another prognostically relevant, yet complex factor is related to the morphological and molecular heterogeneity of the tumor cells and the surrounding stroma. On one hand, the need for proper pathological evaluation with a detailed prognostic assessment is required, on the other, more importantly, the variability of PC leads to chemoresistance[3]. Moreover, differentiating PC from large tumors of the ampulla of Vater remains challenging[4]. Our previous article highlighted recent trends in PC pathology and research[3]. Vibrational spectroscopy (VS)-based methods will play prominent roles in the early diagnosis of PC (Figure 1). Detailed studies regarding the molecular nature of PC are required to reveal novel early and precise diagnostic technologies, thus improving survival rates.

EARLY DIAGNOSIS OF PC

The lack of specific and sensitive early diagnostic options for PC screening results in late-stage disease at the time of diagnosis and is one of the reasons for the poor overall PC survival rates. Out of the available options, serum antigen levels, such as carbohydrate antigen 19-9, are insufficient because of poor specificity and sensitivity for malignancy detection[5,6]. Some studies highlighted the usefulness of measuring interleukin-6 serum levels to differentiate pancreatic ductal adenocarcinoma (PDAC) from chronic or acute pancreatitis[7-9]. Recently, leukemia inhibitory factor was reported to be a promising serum biomarker of pancreatic malignancy[10], a monitoring indicator of treatment response[11-14], and a predictor of metastatic disease in PDAC patients[15].

LIQUID BIOPSY BIOMARKERS

Circulating tumor DNA

Recently, some new circulating biomarkers, particularly those associated with early-stage disease, have been intensely studied. These include a group of tumor nucleic acids in the patient's blood serum, such as circulating tumor DNA (ctDNA), cell-free DNA, cell-free RNA, circulating tumor cells, or extracellular vesicles (EVs) (Figure 2)[16]. Aberrant methylation markers specific for certain malignancies, including PDAC, were previously detected using genetic methods, such as next-generation sequencing and droplet digital polymerase chain reaction[17-20], but the sensitivity of these methods is still low[20].

EVs

Both tumor and healthy cells release lipid-membrane vesicles, termed "extracellular vesicles", into the bloodstream. EVs include oncosomes, apoptotic bodies, microvesicles, and exosomes of different sizes and biogenesis[21,22]. They have an established role in cancer cell communication and metastasis[23,24]. EVs are released in a passive mode (apoptotic bodies) from dying cells (apoptosis, necrosis) or in an active mode (microvesicles, exosomes) from living cells[25]. They make up a large part of the ctDNA (single-strand DNA and double-strand DNA) in a liquid biopsy, but also carry other biomarkers, such as tumor protein antigens[26,27], or microRNA[28,29]. The cargo within an EV creates a unique spectro-
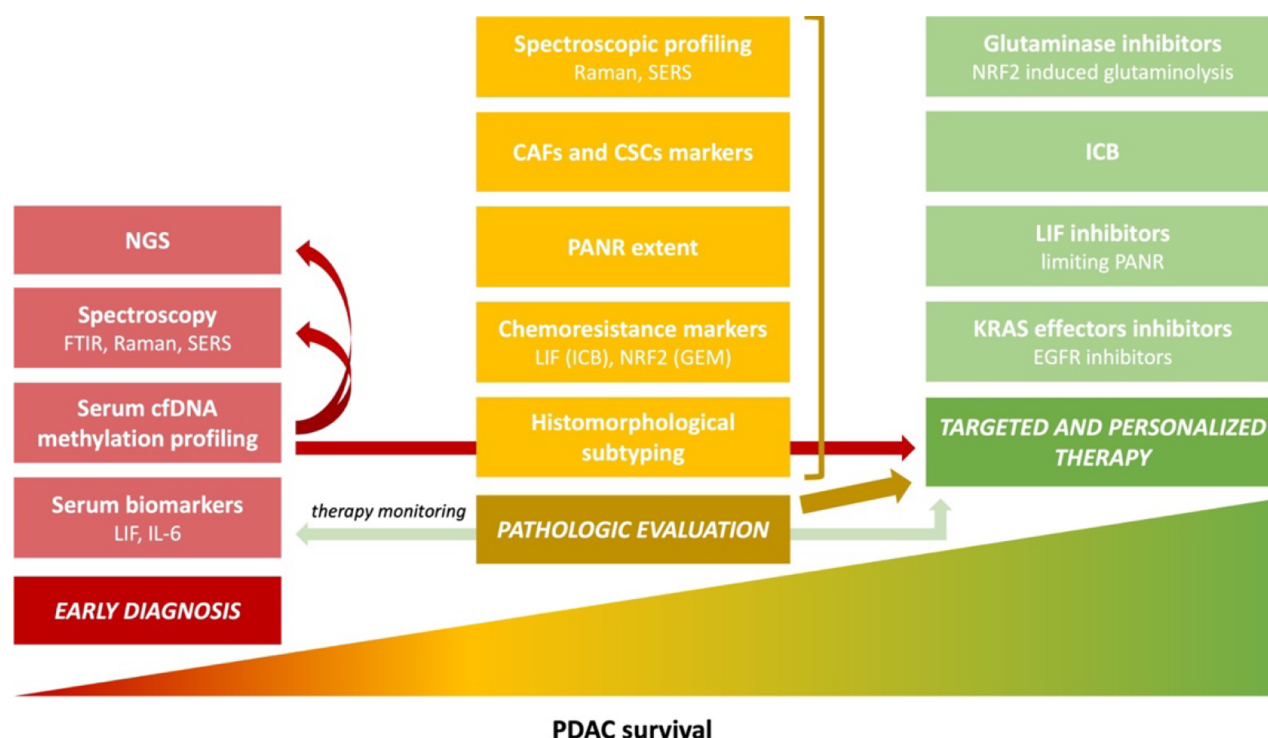
Figure 1 Main trends in pancreatic cancer pathology and research aim to improve survival[3]. Poor pancreatic ductal adenocarcinoma (PDAC) prognosis for a patient is multifactorial, and the lack of sensitive and specific early diagnostic methods is one of the reasons. Another is the resistance to available therapeutic options, which is caused, among others, by the tumor's molecular and morphological heterogeneity. Detailed pathological reporting is crucial for targeted and personalized therapy. The development of new diagnostic methods, combined with a proper pathologic evaluation and spectroscopic profiling, lead to effective treatment and supposedly will increase PDAC patients' survival rates; Adapted with permission from[3]. Citation: Szymoński K, Milian-Ciesielska K, Lipiec E, Adamek D. Current Pathology Model of Pancreatic Cancer. *Cancers (Basel)* 2022; 14: 2321. Copyright© The Authors 2020. Published by MDPI. The image may be redistributed without special permissions–source: https://www.mdpi.com/openaccess. PDAC: Pancreatic ductal adenocarcinoma; LIF: Leukemia inhibitory factor; IL-6: Interleukin-6; cfDNA: Cell-free DNA; FTIR: Fourier transform infrared spectroscopy; Raman: Raman spectroscopy; SERS: Surface-enhanced Raman spectroscopy; NGS: Next-generation sequencing; ICB: Immune checkpoint blockers; NRF2: Nuclear factor-erythroid factor 2-related factor 2; GEM: Gemcitabine; PANR: PDAC-associated neural remodeling; CAFs: Cancer-associated fibroblasts; CSCs: Cancer stem cells; EGFR: Epithelial growth factor receptor.
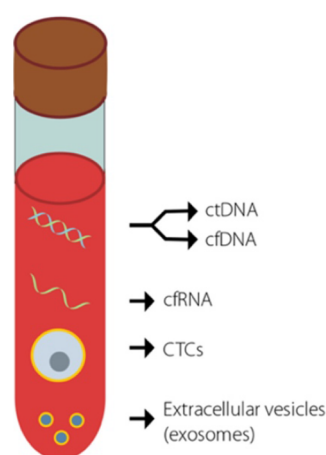

Figure 2 Potential cancer biomarkers that are detectable using liquid biopsies[16]. Key biomarkers that are currently used in an attempt to detect early-stage cancer. Adapted with permission from[16]. Citation: Jaworski JJ, Morgan RD, Sivakumar S. Circulating Cell-Free Tumour DNA for Early Detection of Pancreatic Cancer. *Cancers (Basel)* 2020; 12: 3704. Copyright© The Authors 2020. Published by MDPI. The image may be redistributed without special permissions–source: https://www.mdpi.com/openaccess. ctDNA: Circulating tumor DNA; cfDNA: Cell-free DNA; cfRNA: Cell-free RNA; CTCs: Circulating tumor cells.

scopic fingerprint specific for certain tumors, including PDAC. The molecular characteristics that were obtained using VS, when combined with the use of convolutional neural networks (CNN), can be specified to pinpoint valuable information from the general all-in-one data, thus creating a tool for successful diagnosis.

DNA methylation

DNA methylation plays an important role in the regulation of gene expression, and cellular differentiation. Aberrant methylation is associated with the development and progression of various cancers[16, 30,31]. Moreover, these specific aberrations may act as cancer biomarkers, enabling early diagnosis. Standard methylation status evaluation techniques include bisulfite conversion assay, the sequencing or melting curve analysis, restricted enzymes-based assay, and affinity capture using methylated DNA binding proteins[31]. Some of these methods suffer from a laborious workflow and false positives. Others require specific reagents such as enzymes and binding proteins, which are costly and time-consuming. Therefore, we will discuss the usefulness of label-free spectroscopic methods, such as surface-enhanced Raman spectroscopy (SERS), in DNA methylation status detection, and we will address potential issues that arise when using these techniques.

PDAC HETEROGENEITY AND CHEMORESISTANCE

PC is well known to be very heterogeneous in molecular and morphological phenotype. It is one of the reasons, aside from the lack of adequate early diagnosis methods, for patients' poor prognosis, because current treatment options do not consider tumor heterogeneity and thus give insufficient results[32,33]. When designing studies based on PC diagnosis, one must differentiate the results concerning the histomorphological type of the tumor. This distinction is crucial for patient care, due to different molecular pathways governing the development and evolution of these tumors, as well as the prognosis assessment and potential treatment options[3]. In our previous study, we demonstrated that Raman spectroscopy (RS) is capable of detecting ampullary cancer in pancreatic tumor tissue slides[34]. A subsequent step could then differentiate between three groups of pancreatic tumors, specifically conventional PDAC (cPDAC), PDAC derived from intraductal papillary mucinous neoplasm (IPMC), and ampulla of Vater adenocarcinoma (AVAC). cPDAC is the most common form of PC developing *via* pancreatic intraepithelial neoplasia. It arises in the ductal epithelium localized in "normal" pancreatic tissue, sometimes with signs of chronic pancreatitis, or is combined with so-called "acinar-to-ductal metaplasia" regions. The KRAS mutation is the initiating event in this pathway of carcinogenesis[3]. IPMCs are carcinomas that arise from intraductal papillary mucinous neoplasms (IPMNs), cystic tumors that develop in the main or peripheral branches of pancreatic ducts. Guanine nucleotide-binding protein, alpha stimulating activity polypeptide proto-oncogene mutation, which is not found besides pancreatic tumors, plays a significant role in IPMN development[3]. The third relevant group, AVAC, is a cancer of the duodenal ampulla of Vater. AVAC tumors usually grow with large diameters and deeply infiltrate the pancreatic tissue. The histomorphological similarities of AVAC and cPDAC tumors often make them hard to distinguish from each other. Moreover, these tumors are treated clinically and diagnostically in the same way, but the latest reports suggest that they differ regarding the incidence of various prognostic factors, such as tumor differentiation, perineural invasion, venous invasion, or lymph node involvement[35]. The early occurrence of bile duct obstruction symptoms in IPMC or AVAC enables earlier diagnosis and thus may lead to a better prognosis. However, the reports supporting this are ambiguous[3]. Further subgrouping of PC tumors into morphologically distinct entities, such as large duct (cystic papillary), foamy glands, clear cell, adenosquamous, vacuolated-cell, or colloid, which are described in detail elsewhere[3], may benefit patients and clinicians, because of the different prognosis of some of these groups.

Recently, Mukhopadhyay *et al*[36] showed that the nuclear factor-erythroid 2–related factor 2 (NRF2) was responsible for gemcitabine chemoresistance, and the NRF2 expression level in PDAC tissues correlated with poor patient outcome. Another study by Patzak *et al*[37] described cytosolic 5'-nucleotidase 1A (NT5C1A) as a mediator of this resistance by reducing intracellular gemcitabine metabolites and limiting its efficacy. Gemcitabine is a standard chemotherapeutic agent for PDAC. The tumor's resistance to the therapy is among the main reasons for the drastically low 5-year survival rates of PC. Thus, recognizing this chemoresistance is of great importance. The utilization of VS is a step closer in this direction. VS may prove to be beneficial for identifying chemoresistant pancreatic tumors. More studies are required to evaluate NRF2 or/and NT5C1A expression levels in PDAC tissues, and compare them with spectroscopic data to identify spectral markers that correlate with gemcitabine chemoresistance. These studies will help select patients who might benefit from gemcitabine therapy.

Molecular spectroscopy can obtain all the information about the studied sample with a single measurement. There is no need for special labeling or selecting areas of interest that other genetic or biochemical methods may require. Analyzing spectral data plays a pivotal role. All data is ready for interpretation and provides information about the tumor, such as its subtype, differentiation level, specific chemoresistance, and other tumor-specific poor prognostic factors (hepatocyte nuclear factor-1B expression, or cancer stem cells). Whether or not we can decipher this information will determine the efficacy of this methodology in diagnostics.

MOLECULAR SPECTROSCOPY

Molecular VS was confirmed to play an important role in the characterization of the chemical structure and composition of malignant tissues[38-41] and the analysis of human blood serum[42]. Due to high chemical selectivity, RS and infrared spectroscopy (IR) might evolve into efficient pancreatic malignancy screening tools[43]. These methodologies provide information about various biologically significant molecules and functional groups in a tumor including phospholipids and triglycerides, proteins, nucleic acids, phosphates, and carbohydrates. Electromagnetic radiation (photons) can be absorbed and/or scattered by a sample because the energy of phonon excitations and vibrations, as well as oscillations of functional groups in molecules correspond to the energy in the infrared region of the electromagnetic radiation spectrum. Therefore, the presence of functional groups in the analyte can be detected *via* interaction with light, providing information about the molecular structure and composition of the investigated sample. IR and RS take advantage of absorption and inelastic scattering, respectively. The results obtained using these methods indicate differences in metabolic pathways typical for various neoplasms. The main advantage of the molecular spectroscopic approach is collecting information about samples in a label-free and noninvasive manner. The potential of spectroscopic methods has not yet been fully explored in the investigation of PC. However, preliminary results are promising[34,44].

Raman hyperspectral mapping

RS is a VS method that can deliver a detailed molecular fingerprint of a studied sample almost real-time, without the need for special labeling. A technique called Raman hyperspectral mapping (RHM) provides high-resolution imaging at a relatively low cost compared to other well-established medical imaging techniques, such as magnetic resonance imaging (MRI)[43]. RHM relies on multiple measurements of adjacent "pixels" of tissue and integrating the resulting spectra into a single map image. Such imaging allows the selection of points of interest in the studied sample and precise distinction between particular tissue elements, such as cancer cells (with nuclei and cytoplasm) or the stroma compartment (Figure 3).

The RHM methodology of tissue samples usually involves slicing 2.5 μm thick tissue sections with a microtome from a standard formalin-fixed paraffin-embedded tissue block. Because glass gives substantial interference in the Raman readings, for the slide mounting, special calcium fluoride (CaF$_2$) windows are used for the slide mounting instead. Another important aspect of RHM is the selection of the region of interest which is usually a part of the cancerous glands or stroma compartment. This should be completed by an experienced pathologist. Subsequently, a complete paraffin removal ought to be conducted involving a 12-h xylene bath and graded ethanol rehydration. On such preprocessed tissue slides, the Raman measurements can be performed. This methodology was already described in our previous paper on ampullary cancer detection with RHM[34].

RS is accurate, and the information obtained is characterized by good resolution. Nevertheless, RS requires a sophisticated methodology and equipment because the Raman effect is very weak in nature [43,45]. Another drawback of RS is the substantial sample pre-processing, manual selection of points of interest, and further data analysis.

SERS

To augment the Raman signal strength, another method called SERS is used[43,45]. SERS is a label-free, ultrasensitive tool, capable of DNA methylation analysis. SERS utilizes the same physical phenomenon as RS (Raman effect), but the effect is significantly enhanced using specially synthesized plasmonic nanoparticles[46] into which the molecules (such as DNA) are absorbed [31]. Furthermore, SERS is efficient for liquid biopsy measurements, and it does not require special labeling. However, the production of nanoparticles requires an experienced team and a proper methodology.

The unique ability for such sensitivity of SERS is achieved *via* the use of plasmonic nanostructures (SERS substrates), between which so-called "hot spots" are localized. The traditional methodology of SERS measurements involved mixing the sample with gold (Au) nanoparticles (20-50 nm diameter). One method of SERS substrate production prepares Au nanoparticles *via* chemical reduction of tetrachloroauric (III) acid using trisodium citrate under specific reaction conditions according to the procedure described by Frens[46]. However, this synthesis method is characterized by random and nonuniform hot spot distribution, which leads to the poor reproducibility of SERS[31,47,48]. Another issue, especially regarding DNA methylation studies, is the difficulty distinguishing between DNA methylation signals and the adjacent nucleotide signals due to their similarity[49]. Additionally, surfactants and/or capping agents can affect signal purity[50]. Overcoming these issues is crucial to utilizing SERS for diagnostics largely based on DNA methylation analysis. To achieve this, a methodology of proper SERS substrate development is required. Such a substrate is characterized by a large Raman enhancement, regularly arranged hot spots, and an open and easily accessible surface topology[31]. Luo *et al*[31] proposed the use of a plasmonic gold nanohole array (PGNA) as a SERS substrate. Originally, the authors described the use of electron beam lithography (EBL) for the PGNA substrates production, but a focused ion beam (FIB) might be even better, due to its higher resolution [51]. Both EBL and FIB can be used to obtain a periodic matrix of holes (plasmonic nano-holes array) in a
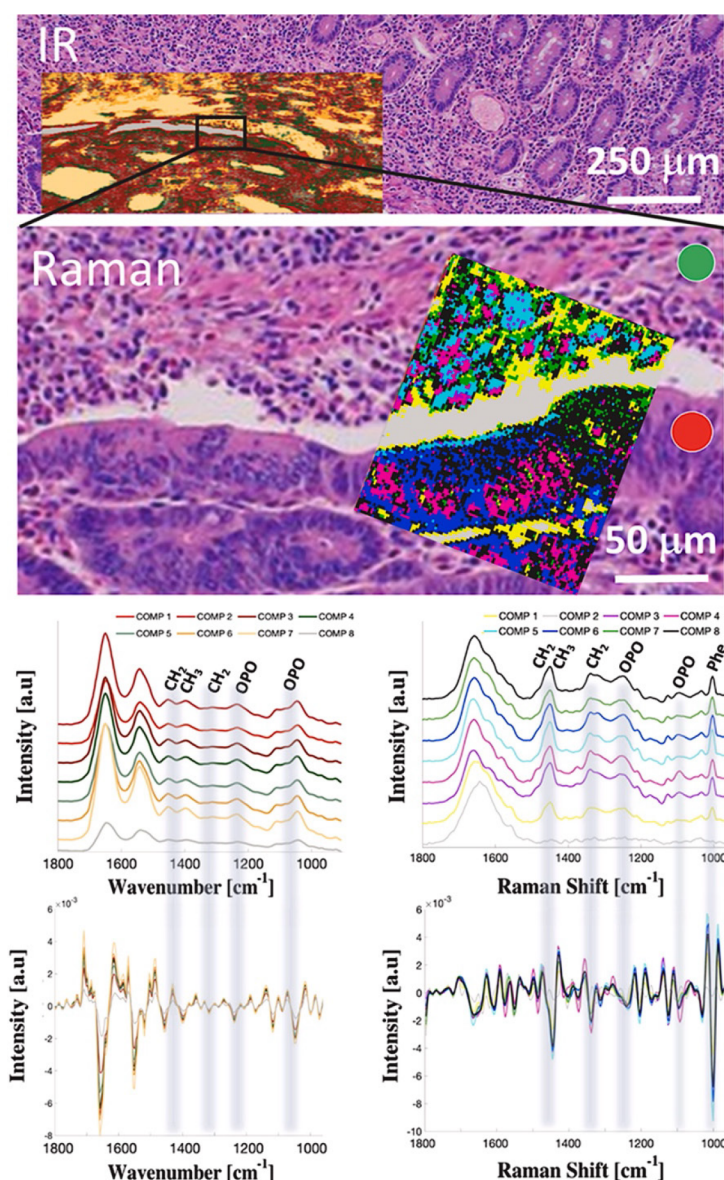
Figure 3 Spectroscopic mapping of ampullary adenocarcinoma[34]. A hematoxylin-eosin slide image of ampullary cancer tissue with superimposed Fourier transform infrared spectroscopy and Raman hyperspectral maps treated with hierarchical cluster analysis and averaged spectra with corresponding second derivatives from each cluster. Spectroscopic maps cover both cancerous (red circle) and noncancerous–stroma (green circle) tissue fragments. Adapted with permission from[34]. Citation: Szymoński K, Lipiec E, Sofińska K, Skirlińska-Nosek K, Milian-Ciesielska K, Szpor J, Czaja M, Seweryn S, Wilkosz N, Birarda G, Piccirilli F, Vaccari L, Szymoński M. Spectroscopic screening of pancreatic cancer. *Clin Spect* 2021; 3: 100016. Copyright© The Authors 2020. Published by ELSEVIER.

Au layer evaporated onto an atomically flat non-plasmonic substrate[31] (Figure 4).

SERS is characterized by a large amount of work required for substrate production, but the sample pre-processing is minimal. Raw blood serum samples can be analyzed, without requiring preselection by a specialist. This is a valuable asset in the search for a diagnostic option.

Attenuated total reflection Fourier transform infrared spectroscopy

Attenuated total reflection Fourier transform infrared spectroscopy (ATR-FTIR) is a complementary method to SERS and RS, which measures absorption as an infrared spectrum of a biochemical fingerprint. It is widely accessible, inexpensive, and easy to use and implement, while still providing substantial information about the sample being analyzed. ATR-FTIR is characterized by high sensitivity in biomedical diagnostics[52] and can detect spectral markers of many pathologies in physiological liquids such as saliva or blood serum. ATR-FTIR was used to detect lung[53], bile duct[54], ovary[55, 56], breast[57], and brain[58,59] tumors in the blood serum samples. In a study by Butler *et al*[59], the team described the methodology of using ATR-FTIR in an early screening of brain tumor patients with sensitivity and specificity of 93.2% and 92.8%, respectively. The high sensitivity of ATR-FTIR can be achieved due to the design of the instrumental set-up. In the ATR-FTIR, a sample droplet is placed on an internal reflection element (IRE), also called the ATR crystal. IREs with high refractive indexes, such
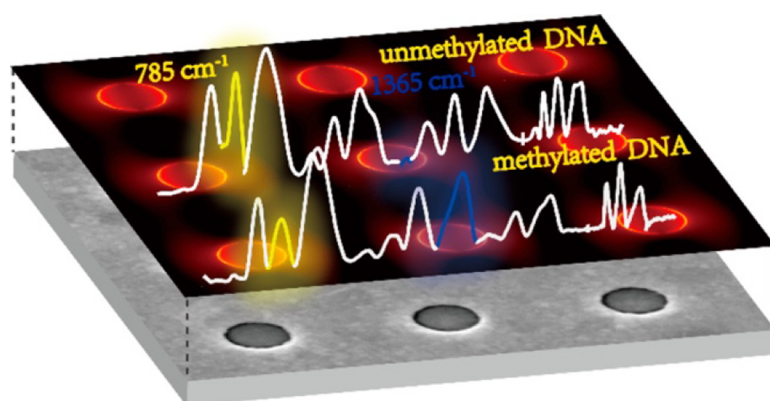
Figure 4 A schematic of a plasmonic gold nanohole array as a surface-enhanced Raman spectroscopy-active substrate used in the detection of DNA methylation[31]. Citation: Luo X, Xing Y, Galvan DD, Zheng E, Wu P, Cai C, Yu Q. Plasmonic Gold Nanohole Array for Surface-Enhanced Raman Scattering Detection of DNA Methylation. *ACS Sens* 2019; 4: 1534–1542. Copyright© The Authors 2020. Published by American Chemical Society.

as diamond, zinc selenide, germanium, or silicon are used. The incident IR beam is directed through the IRE and the resultant evanescent wave extends beyond the ATR crystal surface and penetrates the sample. As with SERS, there is no need for sample pre-processing or special labeling. No additional substrate production is required as well. However, achieving sensitive information about the samples, and their interpretation remains challenging. Employing CNNs to analyze ATR-FTIR data is a viable solution that has already shown success[59].

Molecular spectroscopy in diagnostics

Multiple studies have shown the usefulness of VS in the detection of DNA methylation status of cells, as well as other diagnostically important factors. For example, DNA methylation aberrations were successfully detected using SERS[31,49,60-62], and FTIR[63-69] methods. Recently, in an interesting study by Ho *et al*[70], the authors successfully utilized RS and deep neural networking that enabled culture-free serum bacteria identification and antibiotic susceptibility testing with about 97% accuracy. In other studies[71,72], RS and SERS were used to detect previously untraceable concentrations of biomarkers (matrix metalloproteinase 7, and mucin 4) in the serum of PDAC patients.

To date, no report has combined detecting aberrant DNA methylation markers of PDAC, obtained by analyzing patients' serum with the use of VS methods, specifically the ATR-FTIR combined with CNN classification. Although partial results are available and show promise, additional investigations are needed to support the combination of VS and CNN for PDAC detection.

Residual disease monitoring

Currently, there are no efficient methods of patient monitoring for minimal residual disease (MRD) in PC. Generally, postoperative surveillance methods, including monitoring of clinical symptoms, blood tumor markers, and computed tomography or MRI are used, but these methods lack sensitivity and specificity for MRD[73]. Recently, ctDNA detection using genetic methods in a liquid biopsy was highly advocated for MRD monitoring[73-76]. VS could be a better option for MRD monitoring compared to genetic mutation detection, because of VS's ability to identify DNA methylation aberrations[20]. However, there are currently no data supporting this. Notably, as highlighted by Henriksen *et al*[77], surgical trauma elevates the serum ctDNA levels up to 4 wk after the surgery. This should be taken into consideration when designing a MRD monitoring study using liquid biopsy analysis.

MULTIVARIATE DATA ANALYSIS

In spectroscopic data evaluation, it is very important to draw proper conclusions. Therefore, various methods of multivariate analysis are used to help with data interpretation[78]. Hyperspectral mapping with K-means clustering and principal component analysis (PCA) is commonly performed to explore spectral variation[79-81]. Pre-processing of the spectroscopic data involves cosmic rays removal, baseline correction, and smoothing (adaptive multi-round smoothing based on the Savitzky-Golay filter). Minimal necessary operations are performed to explore marker bands of significant aspects, such as DNA methylation while preventing the loss of important spectral information. Multivariate analysis is carried out to reduce data dimensionality and extract the most important parameters from the acquired information. Some of these methods are briefly described below.

Unsupervised hierarchical cluster analysis is a clustering algorithm designed to group the obtained spectra or to produce false-colored maps based on spectral similarity and variability.

PCA is based on a linear transformation of the data to a new space described by orthogonal axes, the so-called principal components. The most significant results are the "score" values, which represent the data in multidimensional space corresponding to the principal components, and the loading values, which identify the variables causing the data separation according to their influence on the scores. Additionally, the results of the PCA are complemented by the explained variance. The explained variance gives the percentage of the total variance of the original data set, which is explained by a certain principal component.

Partial least squares regression (PLSR) involves a linear transition of numerous original descriptors to a new variable space based on a small number of orthogonal factors (latent variables). In other words, PLSR allows the construction of predictive models when the factors are highly collinear. This analysis estimates unobservable variables as exact linear combinations of their empirical indicators. The estimated proxies are treated as substitutes for the latent variables. The selected case values capture most of the independent variables' variance. This variance is used for predicting the dependent variable.

Non-negative matrix factorization (NMF) is a useful tool for the analysis of high-dimensional data. Besides detecting a compressed representation, NMF provides insights into the structure and features of the given data by extracting easily interpretable factors. With the use of NMF, basic spectral components for proteins, lipids, phospholipids, or nucleic acids can be compared.

All of these methods depend on human decision-making, although minimal. As a result, these methods might disturb and lose seemingly irrelevant data. To overcome this and enhance the sensitivity and specificity of spectral data interpretation, deep neural networks are used.

CNNs

RS is characterized by a very low strength of the measured effect (Raman effect) and is thus very sensitive to distortion factors, such as fluorescence, thermal noise, the quality of the measuring equipment, and research team experience. Removing noise artifacts requires various pre-processing methods on acquired spectra (*i.e.*, cosmic rays removal, baseline correction, and smoothing). All of these depend on human input and definition. On one hand, this prevents automation, but on the other, some seemingly irrelevant data might be lost during the pre-processing. The successful use of CNNs in spectroscopic data evaluation and classification was shown in multiple studies[82,83] including those using RS[84], SERS[70,85], and ATR-FTIR[86]. As a source, the CNN is fed with raw spectral data, without human interaction-related pre-processing. This approach gives better results in classification than conventional methods[82] and makes the methodology more universal. CNN objectives should be clear and simple. Similarly, the proper selection of training data is of great importance. For each objective, different training sets should be created. All spectral results from each group are split (2:1) to form the training (two-thirds of data) and testing (one-third of data) datasets. Training datasets should include positive and negative cases, preferably with many variations. For example, when designing a training dataset for the CNN that will decide whether the results are from a malignant pancreatic tumor or not from a malignant pancreatic tumor, one might include cases from PC, but as a negative control also include malignancies of other sites (*i.e.*, colorectal or breast carcinomas), and benign pancreatic entities (*i.e.*, IPMN, mucinous cystic neoplasm, or groove pancreatitis).

There are some issues related to utilizing neural networks (NNs) with multiple layers (deep NNs), specifically very deep NNs. One obstacle is the vanishing/exploding gradient problem. A training process of a NN in simplest words usually involves updating the "weights" of the algorithm to better cope with the problem, that the NN is exposed to. A great benefit of a CNN is that it extracts features of the task on its own. Feature extraction is done using an optimization algorithm, such as "gradient descent", which simply finds values of a function's parameters to minimize the cost function. If gradients that update the weights shrink, the weights are no longer updated, and the learning stops. This is called a vanishing gradient problem. Similarly, if gradients grow, weights do not update reasonably, and the learning becomes unstable, resulting in the exploding gradient problem. The so-called "skip connection" technique is utilized to overcome the vanishing/exploding gradient problem [83,87]. This is the basis for a Residual Network CNN architecture[87]. Another issue of CNNs is data overfitting. When a model used to train the CNN is very complex, and there is a limited amount of learning data available, the CNN learns to know the training dataset well, but performs poorly against any new data (*i.e.*, validating/testing dataset or the implementation data). To improve this shortcoming, proper data augmentation techniques might be needed. Usually, in spectral data analysis, augmentation requires making additional artificial spectra by small spectral shifting, expanding spectral range, adding Gaussian noise, or superimposing the spectra for each real result[83].

Table 1 Drawbacks of spectroscopic-related methodologies, and proposed solutions

Methodology	Issue	Solution
Raman spectroscopy (random measurement points selection)	(1) Not differentiating the results with cancer subtype leads to mixed results; and (2) Mixing measurements of cancer cells or tumor stroma with pancreatitis, inflammation, necrosis, colloid, *etc*	(1) Results should be analyzed with regard to cancer subtype; (2) Raman hyperspectral mapping allows the selection of specific points of interest, whether cancer cells or tumor stroma areas; and (3) The selection should be done by an experienced pancreatic pathologist
SERS	(1) Poor reproducibility, due to random and nonuniform hot spots distribution in nanostructures production[31,47,48]; (2) The distinction between methylation signals and those from adjacent nucleotides is difficult, because of their similarity[49]; and (3) Signal purity is affected by the use of surfactants and/or capping agents[50]	(1) PGNA as a SERS substrate; and (2) The use of a FIB to obtain a periodic matrix of holes (plasmonic nano-holes array) in a gold layer evaporated on an atomically flat non-plasmonic substrate[31]
ATR-FTIR	(1) Low sensitivity and specificity; (2) Human-dependent pre-processing and analysis of spectral data; and (3) Not yet confirmed in pancreatic cancer diagnosis	To increase the sensitivity and specificity, and to limit the human intervention, convolutional neural networks are used
Multivariate data analysis	(1) Results might be disturbed by the human-dependant actions and seemingly irrelevant data will be lost; and (2) Losing data lowers the sensitivity and specificity of the method	Using of convolutional neural networks, which are fed with raw spectral data
CNN	(1) Very deep neural networks are characterized by a vanishing/exploding gradient problem; and (2) Overfitting of the CNN trained on a limited amount of data	(1) The use of ResNet CNN architecture, with the use of so-called "skip connections"; and (2) Proper data augmentation to prevent overfitting and sensitize the CNN to deal with various "scenarios"

PGNA: Plasmonic gold nanohole array; FIB: Focused ion beam; SERS: Surface-enhanced Raman spectroscopy; ResNet: Residual Network; CNN: Convolutional neural networks; ATR-FTIR: Attenuated total reflection Fourier transform infrared spectroscopy.

CONCLUSION

When designing a study investigating VS methods in PC diagnosis, it is important to bear in mind a couple of aspects. First, the study is required to address all the drawbacks of RS, SERS, and ATR-FTIR methodology, described briefly above. Another crucial issue is the proper cooperation of multidiscip-linary teams, including medical specialists, such as pathologists and clinicians, and spectroscopic specialists, such as physicists and chemists. For example, comparing PC cases with pancreatic neuroen-docrine neoplasms is not reasonable because the malignancies represent different entities with very distinct pathways governing their initiation and progression, as much as patient survival. When dealing with PC, detailed knowledge of the tumor's molecular and pathological nature is required. In particular, one should interpret the results of measurements in the context of PC subtypes and some other prognostic factors[3]. Proper analysis of measured spectroscopic data is impartant. For example, when conducting a RS on complex tissue samples, such as PC sections, the random spots of measurements will yield ambiguous results. The spectral data might be disturbed by numerous interfering phenomena, such as inflammation, tumor necrosis, or fibrosis. Additionally, the cytoplasmic and nuclear regions of cancer cells significantly differ as well. Interpretation of the PC stroma compartment, including the complicated cancer-stroma "cross-talk", is another aspect to address[3]. Careful selection of regions of interest is important too. This should be done by a specialized pancreatic pathologist familiar with the spectroscopic methodology. Finally, in the search for a PC diagnostic tool, we look for universality and automation. Thus, the interpretation of spectral data obtained from liquid biopsy, which relies on human-dependent pre-processing is not a good path to follow. CNNs are invaluable here, but proper design and training is the key to success. With these aims, we designed a study that will compre-hensively evaluate the potential of VS methods used in diagnosing PC, by systemically evaluating liquid biopsy samples (Figure 5). In conclusion, VS seems to be leading the way in the race, with most of the methodology drawbacks resolved, at least partially (Table 1).
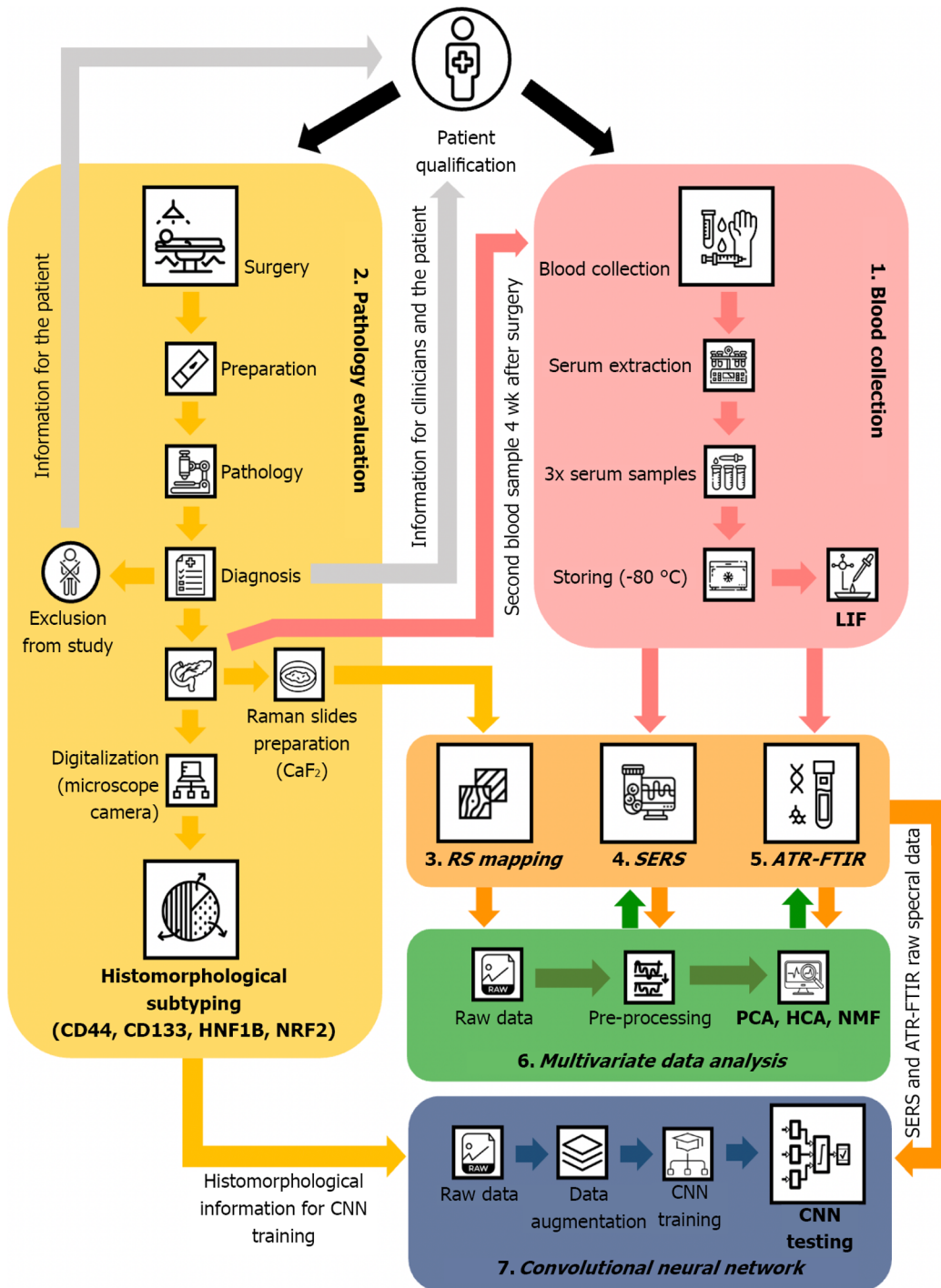
Figure 5 A Designed study regarding the diagnostic potential of vibrational spectroscopic methods used for diagnosing pancreatic cancer by systemically evaluating liquid biopsy samples. LIF: Leukemia inhibitory factor; RS: Raman spectroscopy; SERS: Surface-enhanced Raman spectroscopy; ATR-FTIR: Attenuated total reflection Fourier transform infrared spectroscopy; CNN: Convolutional neural network; NRF2: Nuclear factor-erythroid 2–related factor 2; HNF1B: Hepatocyte nuclear factor-1B; PCA: Principal component analysis; HCA: Hierarchical clustering analysis; NMF: Non-negative matrix factorization.

FOOTNOTES

ORCID number: Krzysztof Szymoński 0000-0002-6127-1918; Łukasz Chmura 0000-0003-2768-0643; Ewelina Lipiec 0000-0003-2946-7387; Dariusz Adamek 0000-0002-5570-2259.

Corresponding Author's Membership in Professional Societies: European Society of Pathology, No. 0000030740.

REFERENCES

1   **Rawla P**, Sunkara T, Gaduputi V. Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. *World J Oncol* 2019; **10**: 10-27 [PMID: 30834048 DOI: 10.14740/wjon1166]

2   **Sung H**, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021; **71**: 209-249 [PMID: 33538338 DOI: 10.3322/caac.21660]

3   **Szymoński K**, Milian-Ciesielska K, Lipiec E, Adamek D. Current Pathology Model of Pancreatic Cancer. *Cancers (Basel)* 2022; **14** [PMID: 35565450 DOI: 10.3390/cancers14092321]

4   **Zimmermann C**, Wolk S, Aust DE, Meier F, Saeger HD, Ehehalt F, Weitz J, Welsch T, Distler M. The pathohistological subtype strongly predicts survival in patients with ampullary carcinoma. *Sci Rep* 2019; **9**: 12676 [PMID: 31481741 DOI: 10.1038/s41598-019-49179-w]

5   **Takaori K**, Bassi C, Biankin A, Brunner TB, Cataldo I, Campbell F, Cunningham D, Falconi M, Frampton AE, Furuse J, Giovannini M, Jackson R, Nakamura A, Nealon W, Neoptolemos JP, Real FX, Scarpa A, Sclafani F, Windsor JA, Yamaguchi K, Wolfgang C, Johnson CD; IAP/EPC study group on the clinical managements of pancreatic cancer. International Association of Pancreatology (IAP)/European Pancreatic Club (EPC) consensus review of guidelines for the treatment of pancreatic cancer. *Pancreatology* 2016; **16**: 14-27 [PMID: 26699808 DOI: 10.1016/j.pan.2015.10.013]

6   **Gui JC**, Yan WL, Liu XD. CA19-9 and CA242 as tumor markers for the diagnosis of pancreatic cancer: a meta-analysis. *Clin Exp Med* 2014; **14**: 225-233 [DOI: 10.1007/s10238-013-0234-9]

7   **Błogowski W**, Deskur A, Budkowska M, Sałata D, Madej-Michniewicz A, Dąbkowski K, Dołęgowska B, Starzyńska T. Selected cytokines in patients with pancreatic cancer: a preliminary report. *PLoS One* 2014; **9**: e97613 [PMID: 24849506 DOI: 10.1371/journal.pone.0097613]

8   **Talar-Wojnarowska R**, Gasiorowska A, Smolarz B, Romanowicz-Makowska H, Kulig A, Malecka-Panas E. Clinical significance of interleukin-6 (IL-6) gene polymorphism and IL-6 serum level in pancreatic adenocarcinoma and chronic pancreatitis. *Dig Dis Sci* 2009; **54**: 683-689 [PMID: 18661238 DOI: 10.1007/s10620-008-0390-z]

9   **Mroczko B**, Groblewska M, Gryko M, Kedra B, Szmitkowski M. Diagnostic usefulness of serum interleukin 6 (IL-6) and C-reactive protein (CRP) in the differentiation between pancreatic cancer and chronic pancreatitis. *J Clin Lab Anal* 2010; **24**: 256-261 [PMID: 20626020 DOI: 10.1002/jcla.20395]

10 **Bressy C**, Lac S, Nigri J, Leca J, Roques J, Lavaut MN, Secq V, Guillaumond F, Bui TT, Pietrasz D, Granjeaud S, Bachet JB, Ouaissi M, Iovanna J, Vasseur S, Tomasini R. LIF Drives Neural Remodeling in Pancreatic Cancer and Offers a New Candidate Biomarker. *Cancer Res* 2018; **78**: 909-921 [PMID: 29269518 DOI: 10.1158/0008-5472.CAN-15-2790]

11 **Shi Y**, Gao W, Lytle NK, Huang P, Yuan X, Dann AM, Ridinger-Saison M, DelGiorno KE, Antal CE, Liang G, Atkins AR, Erikson G, Sun H, Meisenhelder J, Terenziani E, Woo G, Fang L, Santisakultarm TP, Manor U, Xu R, Becerra CR, Borazanci E, Von Hoff DD, Grandgenett PM, Hollingsworth MA, Leblanc M, Umetsu SE, Collisson EA, Scadeng M, Lowy AM, Donahue TR, Reya T, Downes M, Evans RM, Wahl GM, Pawson T, Tian R, Hunter T. Targeting LIF-mediated paracrine interaction for pancreatic cancer therapy and monitoring. *Nature* 2019; **569**: 131-135 [PMID: 30996350 DOI: 10.1038/s41586-019-1130-6]

12 **Wrona E**, Potemski P, Sclafani F, Borowiec M. Leukemia Inhibitory Factor: A Potential Biomarker and Therapeutic Target in Pancreatic Cancer. *Arch Immunol Ther Exp (Warsz)* 2021; **69**: 2 [PMID: 33630157 DOI: 10.1007/s00005-021-00605-w]

13 **Vaddepally RK**, Kharel P, Pandey R, Garje R, Chandra AB. Review of Indications of FDA-Approved Immune Checkpoint Inhibitors per NCCN Guidelines with the Level of Evidence. *Cancers (Basel)* 2020; **12** [PMID: 32245016 DOI: 10.3390/cancers12030738]

14 **Loriot Y**, Marabelle A, Guégan JP, Danlos FX, Besse B, Chaput N, Massard C, Planchard D, Robert C, Even C, Khettab M, Tselikas L, Friboulet L, André F, Nafia I, Le Loarer F, Soria JC, Bessede A, Italiano A. Plasma proteomics identifies leukemia inhibitory factor (LIF) as a novel predictive biomarker of immune-checkpoint blockade resistance. *Ann Oncol* 2021; **32**: 1381-1390 [PMID: 34416362 DOI: 10.1016/j.annonc.2021.08.1748]

15 **Jiang W**, Bai W, Li J, Liu J, Zhao K, Ren L. Leukemia inhibitory factor is a novel biomarker to predict lymph node and distant metastasis in pancreatic cancer. *Int J Cancer* 2021; **148**: 1006-1013 [PMID: 32914874 DOI: 10.1002/ijc.33291]

16   **Jaworski JJ**, Morgan RD, Sivakumar S. Circulating Cell-Free Tumour DNA for Early Detection of Pancreatic Cancer. *Cancers (Basel)* 2020; **12** [PMID: 33317202 DOI: 10.3390/cancers12123704]

17   **Shinjo K**, Hara K, Nagae G, Umeda T, Katsushima K, Suzuki M, Murofushi Y, Umezu Y, Takeuchi I, Takahashi S, Okuno Y, Matsuo K, Ito H, Tajima S, Aburatani H, Yamao K, Kondo Y. A novel sensitive detection method for DNA methylation in circulating free DNA of pancreatic cancer. *PLoS One* 2020; **15**: e0233782 [PMID: 32520974 DOI: 10.1371/journal.pone.0233782]

18   **Gall TMH**, Belete S, Khanderia E, Frampton AE, Jiao LR. Circulating Tumor Cells and Cell-Free DNA in Pancreatic Ductal Adenocarcinoma. *Am J Pathol* 2019; **189**: 71-81 [PMID: 30558725 DOI: 10.1016/j.ajpath.2018.03.020]

19   **Imaoka H**, Sasaki M, Hashimoto Y, Watanabe K, Ikeda M. New Era of Endoscopic Ultrasound-Guided Tissue Acquisition: Next-Generation Sequencing by Endoscopic Ultrasound-Guided Sampling for Pancreatic Cancer. *J Clin Med* 2019; **8** [PMID: 31387310 DOI: 10.3390/jcm8081173]

20   **Huerta M**, Roselló S, Sabater L, Ferrer A, Tarazona N, Roda D, Gambardella V, Alfaro-Cervelló C, Garcés-Albir M, Cervantes A, Ibarrola-Villava M. Circulating Tumor DNA Detection by Digital-Droplet PCR in Pancreatic Ductal Adenocarcinoma: A Systematic Review. *Cancers (Basel)* 2021; **13** [PMID: 33673558 DOI: 10.3390/cancers13050994]

21   **Raposo G**, Stoorvogel W. Extracellular vesicles: exosomes, microvesicles, and friends. *J Cell Biol* 2013; **200**: 373-383 [PMID: 23420871 DOI: 10.1083/jcb.201211138]

22   **Colombo M**, Raposo G, Théry C. Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. *Annu Rev Cell Dev Biol* 2014; **30**: 255-289 [PMID: 25288114 DOI: 10.1146/annurev-cellbio-101512-122326]

23   **Becker A**, Thakur BK, Weiss JM, Kim HS, Peinado H, Lyden D. Extracellular Vesicles in Cancer: Cell-to-Cell Mediators of Metastasis. *Cancer Cell* 2016; **30**: 836-848 [PMID: 27960084 DOI: 10.1016/j.ccell.2016.10.009]

24   **Martins VR**, Dias MS, Hainaut P. Tumor-cell-derived microvesicles as carriers of molecular information in cancer. *Curr Opin Oncol* 2013; **25**: 66-75 [PMID: 23165142 DOI: 10.1097/CCO.0b013e32835b7c81]

25   **Amintas S**, Vendrely V, Dupin C, Buscail L, Laurent C, Bournet B, Merlio JP, Bedel A, Moreau-Gaudry F, Boutin J, Dabernat S, Buscail E. Next-Generation Cancer Biomarkers: Extracellular Vesicle DNA as a Circulating Surrogate of Tumor DNA. *Front Cell Dev Biol* 2020; **8**: 622048 [PMID: 33604335 DOI: 10.3389/fcell.2020.622048]

26   **Logozzi M**, De Milito A, Lugini L, Borghi M, Calabrò L, Spada M, Perdicchio M, Marino ML, Federici C, Iessi E, Brambilla D, Venturi G, Lozupone F, Santinami M, Huber V, Maio M, Rivoltini L, Fais S. High levels of exosomes expressing CD63 and caveolin-1 in plasma of melanoma patients. *PLoS One* 2009; **4**: e5219 [PMID: 19381331 DOI: 10.1371/journal.pone.0005219]

27   **Hoshino A**, Costa-Silva B, Shen TL, Rodrigues G, Hashimoto A, Tesic Mark M, Molina H, Kohsaka S, Di Giannatale A, Ceder S, Singh S, Williams C, Soplop N, Uryu K, Pharmer L, King T, Bojmar L, Davies AE, Ararso Y, Zhang T, Zhang H, Hernandez J, Weiss JM, Dumont-Cole VD, Kramer K, Wexler LH, Narendran A, Schwartz GK, Healey JH, Sandstrom P, Labori KJ, Kure EH, Grandgenett PM, Hollingsworth MA, de Sousa M, Kaur S, Jain M, Mallya K, Batra SK, Jarnagin WR, Brady MS, Fodstad O, Muller V, Pantel K, Minn AJ, Bissell MJ, Garcia BA, Kang Y, Rajasekhar VK, Ghajar CM, Matei I, Peinado H, Bromberg J, Lyden D. Tumour exosome integrins determine organotropic metastasis. *Nature* 2015; **527**: 329-335 [PMID: 26524530 DOI: 10.1038/nature15756]

28   **Valadi H**, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat Cell Biol* 2007; **9**: 654-659 [PMID: 17486113 DOI: 10.1038/ncb1596]

29   **Skog J**, Würdinger T, van Rijn S, Meijer DH, Gainche L, Sena-Esteves M, Curry WT Jr, Carter BS, Krichevsky AM, Breakefield XO. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat Cell Biol* 2008; **10**: 1470-1476 [PMID: 19011622 DOI: 10.1038/ncb1800]

30   **Locke WJ**, Guanzon D, Ma C, Liew YJ, Duesing KR, Fung KYC, Ross JP. DNA Methylation Cancer Biomarkers: Translation to the Clinic. *Front Genet* 2019; **10**: 1150 [PMID: 31803237 DOI: 10.3389/fgene.2019.01150]

31   **Luo X**, Xing Y, Galvan DD, Zheng E, Wu P, Cai C, Yu Q. Plasmonic Gold Nanohole Array for Surface-Enhanced Raman Scattering Detection of DNA Methylation. *ACS Sens* 2019; **4**: 1534-1542 [PMID: 31074265 DOI: 10.1021/acssensors.9b00008]

32   **Sántha P**, Lenggenhager D, Finstadsveen A, Dorg L, Tøndel K, Amrutkar M, Gladhaug IP, Verbeke C. Morphological Heterogeneity in Pancreatic Cancer Reflects Structural and Functional Divergence. *Cancers (Basel)* 2021; **13** [PMID: 33672734 DOI: 10.3390/cancers13040895]

33   **Verbeke C**. Morphological heterogeneity in ductal adenocarcinoma of the pancreas - Does it matter? *Pancreatology* 2016; **16**: 295-301 [PMID: 26924665 DOI: 10.1016/j.pan.2016.02.004]

34   **Szymoński K**, Lipiec E, Sofińska K, Skirlińska-Nosek K, Milian-Ciesielska K, Szpor J, Czaja M, Seweryn S, Wilkosz N, Birarda G, Piccirilli F, Vaccari L, Szymoński M. Spectroscopic screening of pancreatic cancer. *Clin Spect* 2021; **3**: 100016 [DOI: 10.1016/j.clispe.2021.100016]

35   **Ferchichi M**, Jouini R, Koubaa W, Khanchel F, Helal I, Hadad D, Bibani N, Chadli-Debbiche A, BenBrahim E. Ampullary and pancreatic adenocarcinoma-a comparative study. *J Gastrointest Oncol* 2019; **10**: 270-275 [PMID: 31032094 DOI: 10.21037/jgo.2018.09.09]

36   **Mukhopadhyay S**, Goswami D, Adiseshaiah PP, Burgan W, Yi M, Guerin TM, Kozlov SV, Nissley DV, McCormick F. Undermining Glutaminolysis Bolsters Chemotherapy While NRF2 Promotes Chemoresistance in KRAS-Driven Pancreatic Cancers. *Cancer Res* 2020; **80**: 1630-1643 [PMID: 31911550 DOI: 10.1158/0008-5472.CAN-19-1363]

37   **Patzak MS**, Kari V, Patil S, Hamdan FH, Goetze RG, Brunner M, Gaedcke J, Kitz J, Jodrell DI, Richards FM, Pilarsky C, Gruetzmann R, Rümmele P, Knösel T, Hessmann E, Ellenrieder V, Johnsen SA, Neesse A. Cytosolic 5'-nucleotidase 1A is overexpressed in pancreatic cancer and mediates gemcitabine resistance by reducing intracellular gemcitabine metabolites. *EBioMedicine* 2019; **40**: 394-405 [PMID: 30709769 DOI: 10.1016/j.ebiom.2019.01.037]

38   **Krafft C**, Popp J. Micro-Raman spectroscopy in medicine. *Phy Sci Rev* 2019; **4** [DOI: 10.1515/psr-2017-0047]

39   **Krafft C**, Sergo V. Biomedical applications of Raman and infrared spectroscopy to diagnose tissues. *Spectroscopy* 2006; **20**: 195-218 [DOI: 10.1155/2006/738186]

40   **Diem M**, Miljković M, Bird B, Chernenko T, Schubert J, Marcsisin E, Mazur A, Kingston E, Zuser E, Papamarkakis K,

Laver N. Applications of infrared and raman microspectroscopy of cells and tissue in medical diagnostics: Present status and future promises. *Spectroscopy (New York)* 2012; **27**: 463-496 [DOI: 10.1155/2012/848360]

41 **Grzelak MM**, Wróbel PM, Lankosz M, Stęgowski Z, Chmura Ł, Adamek D, Hesse B, Castillo-Michel H. Diagnosis of ovarian tumour tissues by SR-FTIR spectroscopy: A pilot study. *Spectrochim Acta A Mol Biomol Spectrosc* 2018; **203**: 48-55 [PMID: 29859492 DOI: 10.1016/j.saa.2018.05.070]

42 **Acquarelli J**, van Laarhoven T, Gerretzen J, Tran TN, Buydens LMC, Marchiori E. Convolutional neural networks for vibrational spectroscopic data analysis. *Anal Chim Acta* 2017; **954**: 22-31 [PMID: 28081811 DOI: 10.1016/j.aca.2016.12.010]

43 **Kong K**, Kendall C, Stone N, Notingher I. Raman spectroscopy for medical diagnostics--From in-vitro biofluid assays to in-vivo cancer detection. *Adv Drug Deliv Rev* 2015; **89**: 121-134 [PMID: 25809988 DOI: 10.1016/j.addr.2015.03.009]

44 **Paraskevaidi M**, Matthew BJ, Holly BJ, Hugh BJ, Thulya CPV, Loren C, StJohn C, Peter G, Callum G, Sergei KG, Kamila K, Maria K, Kássio LMG, Pierre MHL, Evangelos P, Savithri P, John AA, Alexandra S, Marfran S, Josep SS, Gunjan T, Michael W, Bayden W. Clinical applications of infrared and Raman spectroscopy in the fields of cancer and infectious diseases. *App Spec Rev* 2021; **56**: 804-868 [DOI: 10.1080/05704928.2021.1946076]

45 **Bonifacio A**, Cervo S, Sergo V. Label-free surface-enhanced Raman spectroscopy of biofluids: fundamental aspects and diagnostic applications. *Anal Bioanal Chem* 2015; **407**: 8265-8277 [PMID: 25935674 DOI: 10.1007/s00216-015-8697-z]

46 **Frens G**. Controlled Nucleation for the Regulation of the Particle Size in Monodisperse Gold Suspensions. *Nature Phy Sci* 1973; **241**: 20-22 [DOI: 10.1038/physci241020a0]

47 **Klingsporn JM**, Jiang N, Pozzi EA, Sonntag MD, Chulhai D, Seideman T, Jensen L, Hersam MC, Van Duyne RP. Intramolecular insight into adsorbate-substrate interactions *via* low-temperature, ultrahigh-vacuum tip-enhanced Raman spectroscopy. *J Am Chem Soc* 2014; **136**: 3881-3887 [PMID: 24548252 DOI: 10.1021/ja411899k]

48 **Papadopoulou E**, Bell SE. DNA reorientation on Au nanoparticles: label-free detection of hybridization by surface enhanced Raman spectroscopy. *Chem Commun (Camb)* 2011; **47**: 10966-10968 [PMID: 21909525 DOI: 10.1039/c1cc13705k]

49 **Barhoumi A**, Halas NJ. Detecting Chemically Modified DNA Bases Using Surface Enhanced Raman Spectroscopy. *J Phys Chem Lett* 2011; **2**: 3118-3123 [PMID: 24427449 DOI: 10.1021/jz201423b]

50 **Xu LJ**, Lei ZC, Li J, Zong C, Yang CJ, Ren B. Label-free surface-enhanced Raman spectroscopy detection of DNA with single-base sensitivity. *J Am Chem Soc* 2015; **137**: 5149-5154 [PMID: 25835155 DOI: 10.1021/jacs.5b01426]

51 **Morimoto H**, Sasaki Y, Saitoh K, Watakabe Y, Kato T. Focused ion beam lithography and its application to submicron devices. *Microelectron Eng* 1986; **4**: 163-179 [DOI: 10.1016/0167-9317(86)90009-2]

52 **Baker MJ**, Trevisan J, Bassan P, Bhargava R, Butler HJ, Dorling KM, Fielden PR, Fogarty SW, Fullwood NJ, Heys KA, Hughes C, Lasch P, Martin-Hirsch PL, Obinaju B, Sockalingum GD, Sulé-Suso J, Strong RJ, Walsh MJ, Wood BR, Gardner P, Martin FL. Using Fourier transform IR spectroscopy to analyze biological materials. *Nat Protoc* 2014; **9**: 1771-1791 [PMID: 24992094 DOI: 10.1038/nprot.2014.110]

53 **Yang X**, Ou Q, Qian K, Yang J, Bai Z, Yang W, Shi Y, Liu G. Diagnosis of Lung Cancer by ATR-FTIR Spectroscopy and Chemometrics. *Front Oncol* 2021; **11**: 753791 [PMID: 34660320 DOI: 10.3389/fonc.2021.753791]

54 **Chatchawal P**, Wongwattanakul M, Tippayawat P, Kochan K, Jearanaikoon N, Wood BR, Jearanaikoon P. Detection of Human Cholangiocarcinoma Markers in Serum Using Infrared Spectroscopy. *Cancers (Basel)* 2021; **13** [PMID: 34680259 DOI: 10.3390/cancers13205109]

55 **Gajjar K**, Trevisan J, Owens G, Keating PJ, Wood NJ, Stringfellow HF, Martin-Hirsch PL, Martin FL. Fourier-transform infrared spectroscopy coupled with a classification machine for the analysis of blood plasma or serum: a novel diagnostic approach for ovarian cancer. *Analyst* 2013; **138**: 3917-3926 [PMID: 23325355 DOI: 10.1039/c3an36654e]

56 **Owens GL**, Gajjar K, Trevisan J, Fogarty SW, Taylor SE, Da Gama-Rose B, Martin-Hirsch PL, Martin FL. Vibrational biospectroscopy coupled with multivariate analysis extracts potentially diagnostic features in blood plasma/serum of ovarian cancer patients. *J Biophotonics* 2014; **7**: 200-209 [PMID: 24259229 DOI: 10.1002/jbio.201300157]

57 **Sitnikova VE**, Kotkova MA, Nosenko TN, Kotkova TN, Martynova DM, Uspenskaya MV. Breast cancer detection by ATR-FTIR spectroscopy of blood serum and multivariate data-analysis. *Talanta* 2020; **214**: 120857 [PMID: 32278436 DOI: 10.1016/j.talanta.2020.120857]

58 **Hands JR**, Clemens G, Stables R, Ashton K, Brodbelt A, Davis C, Dawson TP, Jenkinson MD, Lea RW, Walker C, Baker MJ. Brain tumour differentiation: rapid stratified serum diagnostics *via* attenuated total reflection Fourier-transform infrared spectroscopy. *J Neurooncol* 2016; **127**: 463-472 [PMID: 26874961 DOI: 10.1007/s11060-016-2060-x]

59 **Butler HJ**, Brennan PM, Cameron JM, Finlayson D, Hegarty MG, Jenkinson MD, Palmer DS, Smith BR, Baker MJ. Development of high-throughput ATR-FTIR technology for rapid triage of brain cancer. *Nat Commun* 2019; **10**: 4501 [PMID: 31594931 DOI: 10.1038/s41467-019-12527-5]

60 **Li L**, Lim SF, Puretzky A, Riehn R, Hallen HD. DNA Methylation Detection Using Resonance and Nanobowtie-Antenna-Enhanced Raman Spectroscopy. *Biophys J* 2018; **114**: 2498-2506 [PMID: 29874601 DOI: 10.1016/j.bpj.2018.04.021]

61 **Morla-Folch J**, Xie HN, Gisbert-Quilis P, Gómez-de Pedro S, Pazos-Perez N, Alvarez-Puebla RA, Guerrini L. Ultrasensitive Direct Quantification of Nucleobase Modifications in DNA by Surface-Enhanced Raman Scattering: The Case of Cytosine. *Angew Chem Int Ed Engl* 2015; **54**: 13650-13654 [PMID: 26447808 DOI: 10.1002/anie.201507682]

62 **Guerrini L**, Krpetić Ž, van Lierop D, Alvarez-Puebla RA, Graham D. Direct surface-enhanced Raman scattering analysis of DNA duplexes. *Angew Chem Int Ed Engl* 2015; **54**: 1144-1148 [PMID: 25414148 DOI: 10.1002/anie.201408558]

63 **Vidal Bde C**, Ghiraldini FG, Mello ML. Changes in liver cell DNA methylation status in diabetic mice affect its FT-IR characteristics. *PLoS One* 2014; **9**: e102295 [PMID: 25019512 DOI: 10.1371/journal.pone.0102295]

64 **Mello ML**, Vidal BC. Analysis of the DNA Fourier transform-infrared microspectroscopic signature using an all-reflecting objective. *Micron* 2014; **61**: 49-52 [PMID: 24792446 DOI: 10.1016/j.micron.2014.02.003]

65 **Mello ML**, Vidal BC. Changes in the infrared microspectroscopic characteristics of DNA caused by cationic elements, different base richness and single-stranded form. *PLoS One* 2012; **7**: e43169 [PMID: 22937023 DOI: 10.1371/journal.pone.0043169]

66   **Kelly JG**, Najand GM, Martin FL. Characterisation of DNA methylation status using spectroscopy (mid-IR versus Raman) with multivariate analysis. *J Biophotonics* 2011; **4**: 345-354 [PMID: 21520428 DOI: 10.1002/jbio.201000085]

67   **Kelly JG**, Martin-Hirsch PL, Martin FL. Discrimination of base differences in oligonucleotides using mid-infrared spectroscopy and multivariate analysis. *Anal Chem* 2009; **81**: 5314-5319 [PMID: 19499925 DOI: 10.1021/ac900546m]

68   **Banyay M**, Gräslund A. Structural effects of cytosine methylation on DNA sugar pucker studied by FTIR. *J Mol Biol* 2002; **324**: 667-676 [PMID: 12460569 DOI: 10.1016/s0022-2836(02)01104-x]

69   **Veronezi GMB**, Felisbino MB, Gatti MS v, Mello MLS, Vidal B de C. DNA Methylation Changes in Valproic Acid-Treated HeLa Cells as Assessed by Image Analysis, Immunofluorescence and Vibrational Microspectroscopy. *PLoS One* 2017; **12**: e0170740 [DOI: 10.1371/journal.pone.0170740]

70   **Ho CS**, Jean N, Hogan CA, Blackmon L, Jeffrey SS, Holodniy M, Banaei N, Saleh AAE, Ermon S, Dionne J. Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning. *Nat Commun* 2019; **10**: 4927 [PMID: 31666527 DOI: 10.1038/s41467-019-12898-9]

71   **Banaei N**, Foley A, Houghton JM, Sun Y, Kim B. Multiplex detection of pancreatic cancer biomarkers using a SERS-based immunoassay. *Nanotechnology* 2017; **28**: 455101 [PMID: 28937361 DOI: 10.1088/1361-6528/aa8e8c]

72   **Wang G**, Lipert RJ, Jain M, Kaur S, Chakraboty S, Torres MP, Batra SK, Brand RE, Porter MD. Detection of the potential pancreatic cancer marker MUC4 in serum using surface-enhanced Raman scattering. *Anal Chem* 2011; **83**: 2554-2561 [PMID: 21391573 DOI: 10.1021/ac102829b]

73   **Jiang J**, Ye S, Xu Y, Chang L, Hu X, Ru G, Guo Y, Yi X, Yang L, Huang D. Circulating Tumor DNA as a Potential Marker to Detect Minimal Residual Disease and Predict Recurrence in Pancreatic Cancer. *Front Oncol* 2020; **10**: 1220 [PMID: 32850360 DOI: 10.3389/fonc.2020.01220]

74   **Lee B**, Lipton L, Cohen J, Tie J, Javed AA, Li L, Goldstein D, Burge M, Cooray P, Nagrial A, Tebbutt NC, Thomson B, Nikfarjam M, Harris M, Haydon A, Lawrence B, Tai DWM, Simons K, Lennon AM, Wolfgang CL, Tomasetti C, Papadopoulos N, Kinzler KW, Vogelstein B, Gibbs P. Circulating tumor DNA as a potential marker of adjuvant chemotherapy benefit following surgery for localized pancreatic cancer. *Ann Oncol* 2019; **30**: 1472-1478 [PMID: 31250894 DOI: 10.1093/annonc/mdz200]

75   **Perets R**, Greenberg O, Shentzer T, Semenisty V, Epelbaum R, Bick T, Sarji S, Ben-Izhak O, Sabo E, Hershkovitz D. Mutant *KRAS* Circulating Tumor DNA Is an Accurate Tool for Pancreatic Cancer Monitoring. *Oncologist* 2018; **23**: 566-572 [PMID: 29371474 DOI: 10.1634/theoncologist.2017-0467]

76   **Montagut C**, Vidal J. ctDNA to detect minimal residual disease in pancreatic cancer: moving into clinical trials. *Ann Oncol* 2019; **30**: 1410-1413 [PMID: 31418008 DOI: 10.1093/annonc/mdz236]

77   **Henriksen TV**, Reinert T, Christensen E, Sethi H, Birkenkamp-Demtröder K, Gögenur M, Gögenur I, Zimmermann BG; IMPROVE Study Group, Dyrskjøt L, Andersen CL. The effect of surgical trauma on circulating free DNA levels in cancer patients-implications for studies of circulating tumor DNA. *Mol Oncol* 2020; **14**: 1670-1679 [PMID: 32471011 DOI: 10.1002/1878-0261.12729]

78   **Notarstefano V**, Sabbatini S, Conti C, Pisani M, Astolfi P, Pro C, Rubini C, Vaccari L, Giorgini E. Investigation of human pancreatic cancer tissues by Fourier Transform Infrared Hyperspectral Imaging. *J Biophotonics* 2020; **13**: e201960071 [PMID: 31648419 DOI: 10.1002/jbio.201960071]

79   **Gassner C**, Adegoke JA, Patel SK, Sharma VJ, Kochan K, Burrell LM, Raman J, Wood BR. Improved tissue preparation for multimodal vibrational imaging of biological tissues. *Clin Spect* 2022; **4**: 100021 [DOI: 10.1016/j.clispe.2022.100021]

80   **Meksiarun P**, Ishigaki M, Huck-Pezzei VA, Huck CW, Wongravee K, Sato H, Ozaki Y. Comparison of multivariate analysis methods for extracting the paraffin component from the paraffin-embedded cancer tissue spectra for Raman imaging. *Sci Rep* 2017; **7**: 44890 [PMID: 28327648 DOI: 10.1038/srep44890]

81   **Ali SM**, Bonnier F, Tfayli A, Lambkin H, Flynn K, McDonagh V, Healy C, Clive Lee T, Lyng FM, Byrne HJ. Raman spectroscopic analysis of human skin tissue sections ex-vivo: evaluation of the effects of tissue processing and dewaxing. *J Biomed Opt* 2013; **18**: 061202 [PMID: 23124057 DOI: 10.1117/1.JBO.18.6.061202]

82   **Liu J**, Osadchy M, Ashton L, Foster M, Solomon CJ, Gibson SJ. Deep convolutional neural networks for Raman spectrum recognition: a unified solution. *Analyst* 2017; **142**: 4067-4074 [PMID: 28993828 DOI: 10.1039/c7an01371j]

83   **Kazemzadeh M**, Hisey CL, Zargar-Shoshtari K, Xu W, Broderick NGR. Deep convolutional neural networks as a unified solution for Raman spectroscopy-based classification in biomedical applications. *Opt Commun* 2022; **510**: 127977 [DOI: 10.1016/j.optcom.2022.127977]

84   **Ma D**, Shang L, Tang J, Bao Y, Fu J, Yin J. Classifying breast cancer tissue by Raman spectroscopy with one-dimensional convolutional neural network. *Spectrochim Acta A Mol Biomol Spectrosc* 2021; **256**: 119732 [PMID: 33819758 DOI: 10.1016/j.saa.2021.119732]

85   **Weng S**, Yuan H, Zhang X, Li P, Zheng L, Zhao J, Huang L. Deep learning networks for the recognition and quantitation of surface-enhanced Raman spectroscopy. *Analyst* 2020; **145**: 4827-4835 [PMID: 32515435 DOI: 10.1039/d0an00492h]

86   **Berisha S**, Lotfollahi M, Jahanipour J, Gurcan I, Walsh M, Bhargava R, Van Nguyen H, Mayerich D. Deep learning for FTIR histology: leveraging spatial and spectral features with convolutional neural networks. *Analyst* 2019; **144**: 1642-1653 [PMID: 30644947 DOI: 10.1039/c8an01495g]

87   **He K**, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2015. [cited 10 July 2022]. Available from: https://blog.csdn.net/wspba/article/details/57074389