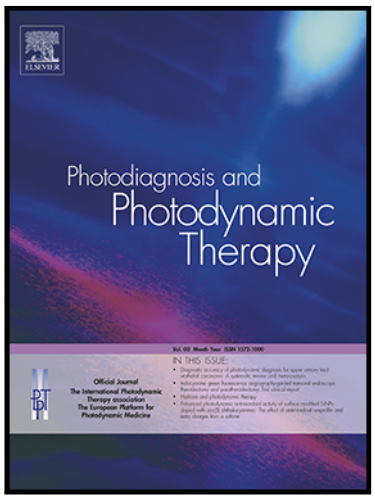
Journal Pre-proof



Noninvasive cholangitis and cholangiocarcinoma screening based on serum Raman spectroscopy and support vector machine

Na Su , Wubulitalifu Dawuti , Yan Hu , Hui Zhao

PII: S1572-1000(22)00442-2
DOI: https://doi.org/10.1016/j.pdpdt.2022.103156
Reference: PDPDT 103156

To appear in: *Photodiagnosis and Photodynamic Therapy*

Received date: 7 July 2022
Revised date: 17 September 2022
Accepted date: 11 October 2022

Please cite this article as: Na Su , Wubulitalifu Dawuti , Yan Hu , Hui Zhao , Noninvasive cholangitis and cholangiocarcinoma screening based on serum Raman spectroscopy and support vector machine, *Photodiagnosis and Photodynamic Therapy* (2022), doi: https://doi.org/10.1016/j.pdpdt.2022.103156

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier B.V.

Journal Pre-proof

Noninvasive cholangitis and cholangiocarcinoma screening based on serum Raman spectroscopy and support vector machine

Na Su [a,1], Wubulitalifu Dawuti[b,1], Yan Hu[c] *, Hui Zhao[a]*

[a] *Department of Clinical Laboratory, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, China*

[b] *School of Public Health, Xinjiang Medical University, Urumqi, Xinjiang, China*

[c] *Science and Technology Talent Development Center of Xinjiang Uygur Autonomous Region, Urumqi, China*

** Corresponding authors.*

E-mail addresses: 31853200@qq.com (Hui Zhao). 493611278@qq.com（*Yan* Hu）

[1] *Contribution equally.*

Journal Pre-proof

Highlights:

- The serum Raman spectra from cholangitis, cholangiocarcinoma, and healthy person, has significant difference in Raman peak intensities and peak wavelength, representing different metabolic component.

- The diagnostic sensitivity for cholangiocarcinoma of serum Raman spectroscopy combined with support vector machine algorithm is better than traditional serological methods.

- The results demonstrated great potentials for developing serum Raman spectroscopy methods into a novel clinical tool for rapid and non-invasive screening of cholangitis and cholangiocarcinoma.

Abstract: The feasibility of serum Raman spectroscopy for rapid screening of cholangitis and cholangiocarcinoma (CCA) was discussed. Raman spectra were collected from 49 patients with cholangitis, 38 patients with CCA, and 55 healthy volunteers. Normalized mean Raman spectra and spectral attributions reveal disease-specific biomolecular differences. Support vector machine (SVM) was used to establish the two-way (cholangitis vs healthy, CCA vs healthy etc.) and 3-way (cholangitis vs CCA vs healthy) classification model, and leave-one-out cross-validation (LOOCV) was used to verify these models' performance. Based on the support vector machine algorithm, serum Raman spectroscopy could identify cholangitis and CCA. Its diagnostic sensitivity, and specificity were 89.80%, 94.55%, and 89.50%, 98.18%, respectively. This study demonstrates that label-free serum Raman spectroscopy analysis technique combined with SVM diagnostic algorithm has great potential for noninvasive cholangitis and CCA screening.

Keywords: Raman Spectroscopy; Serum; Cholangitis; Cholangiocarcinoma (CCA); Screening; Support vector machine (SVM)

1. Introduction

Cholangiocarcinoma (CCA) is the second-largest primary liver malignancy after hepatocellular carcinoma, accounting for about 15% of all primary liver tumors and 3% of all gastrointestinal tumors [1]. The incidence and mortality of CCA has been increasing worldwide in the past decades. Globally, CCA has an incidence of 0.3-6/100,000 per year and a mortality rate of 1-6/100,000 per year. Specific regions such as South Korea, China and Thailand have higher rates, more than six in 100,000 per year [2]. Cholangitis incidence and prevalence vary depending on geographic location, with incidence ranging from 0-1.3/100000 and prevalence from 0-16.2/10000 [3]. Although often having an insidious onset and variable disease course, cholangitis generally leads to the development of cirrhosis and other disease complications [4]. However, cholangitis is positively correlated with the risk of hepatobiliary malignancy, especially CCA. As one of the most serious complications of cholangitis, CCA accounts for about 1/3 of the death causes of patients with

cholangitis [5]. In addition, most CCA is not detected until the late stage, when the available treatment options are very limited, resulting in a poor prognosis of CCA and increasing mortality [6], as many as 72% of patients with cholangitis who develop CCA die within 1 year [7]. Hence the gravity of this condition and need for early differentiating between benign and malignant bile duct disease.

Currently, there is no standardized diagnostic method for early detection of CCA and cholangitis in clinical practice. Imaging modalities including abdominal ultrasound, computed tomography (CT), and/or MRI/magnetic resonance cholangiopancreatography (MRI/MRCP) are used to assess for and characterize dominant strictures [8]. However, all these imaging modalities lack a high positive predictive value (PPV) desired for diagnosis of CCA. For example, the use of MRI/ MRCP for diagnosing CCA, the PPV is only 23% [9]. In addition, many traditional imaging techniques carry radiological damage [10]. Due to the aforementioned shortcomings, many experts recommend the measurement of serum tumor marker Carbohydrate Antigen 19-9 (CA 19-9) in addition to imaging [11]. CA19-9 is the most used tumor marker for CCA, but this assay detects CCA with only 62% sensitivity and 63% specificity. Nonetheless, very high levels of CA 19-9 (≥A1000 U/mL) have been associated with metastatic CCA, so this assay might be used in disease staging rather than diagnosis [12]. Therefore, it is very important to find a convenient, non-invasive, rapid, and highly sensitive diagnostic technique for cholangitis and CCA.

Raman spectroscopy is a label-free, rapid, and non-invasive optical detection technique [13]. Currently, Raman spectroscopy is widely used in basic and applied research in food safety, environmental monitoring, materials science, and archaeological research [14, 15]. Raman spectroscopy has attracted much attention in biomedical research and is emerging as a potentially useful clinical adjunct in cancer diagnosis, identifying subtle changes associated with cancer progression in biological samples such as tissues, cells, serum, urine, and saliva [16]. Recent studies have reported the significance of the application of serum Raman spectroscopy in cancer detection for prostate cancer [17], breast cancer [18], and esophageal cancer [19]. These

Journal Pre-proof

exploratory works indicate that serum Raman spectroscopy is very promising for cancer detection. In addition, the study of serum Raman spectroscopy in cholangitis and CCA detection has not been reported.

This study evaluated the feasibility of using a serum Raman spectroscopy-based technique for label-free detection of serum samples from patients with cholangitis, CCA, and healthy volunteers. Support vector machine (SVM) algorithm was used to establish the diagnostic model, and the linear and Gaussian radial basis kernel functions were compared. This pilot study may lead to developing a rapid, convenient, and non-invasive method for cholangitis and CCA detection.

2. Materials and methods

2.1. Serum sample collection and preparation

In this study, 142 serum samples were collected from the First Affiliated Hospital of Xinjiang Medical University, including 49 from cholangitis patients, 38 from CCA patients, and 55 from healthy volunteers, all of which met the clinical diagnostic criteria. Table 1 lists simple information about these participants, including age and gender distribution. This study has been authorized by the Ethics Committee of the First Affiliated Hospital of Xinjiang Medical University (Ethics No. K202107-16). 3 ml of fresh blood from the patients in fasting state was taken and centrifuged in a high-speed centrifuge at 3000 r/min for 10 min without adding any anticoagulant, and then the serums were obtained by taking the upper clear liquid layer and placed in a centrifuge tube and stored in a refrigerator at -80℃.

Table 1. Age and gender information of patients with Cholangitis, CCA, and healthy volunteers.

	Cholangitis (n=49)	CCA(n=38)	Healthy volunteers(n=55)
Age			
Mean	52	57.3	48
Median	50.8	56.5	46.8
Gender			
Male	29	23	28
Female	20	15	27

Journal Pre-proof

Journal Pre-proof

2.2. Serum Raman spectrum measurements

Raman spectra of serum samples were recorded using Raman micro-spectrometer (ATR6500-785, Optosky, China) equipped with a 785-nm laser. Before testing, each sample was thawed at room temperature, and then 10 μl serum sample was placed on aluminum foil for Raman measurements without air drying. The spectrum acquisition was carried out in a dark environment, the spectral data were obtained by a 40X objective lens, and the wavelength calibration was carried out before each online acquisition, spectral range of 300-1800 cm^{-1}, laser power of 200 mW. The resolution of the spectrometer was about 4.5 cm^{-1}. The acquisition time of 5 seconds and accumulations of 3 times was used for each spectrum. Three tests were carried out on different positions of each sample, and the average spectrum was taken as the analysis spectrum [20]. A total of 142 Raman spectra were obtained, 49 of which were from cholangitis patient serum samples, 38 from CCA patient serum samples, and 55 from healthy people serum samples.

2.3. Spectral data pre-processing

The pre-processing of Raman data includes smoothing, baseline correction, and normalization. We used Savitzky-Golay (SG) algorithm to smooth filter the collected serum spectral data, which can not only improve the smoothness of the spectrum but also reduce the interference of noise. The baseline was removed using the adaptive iteratively reweighted penalized least squares (airPLS) algorithm [21]. Performed vector normalization processing on each spectrum [22].

2.4. Data analysis

After pre-processing the spectra, the spectral data were analyzed using Principal Component Analysis (PCA) methods. Analysis of Variance (ANOVA) was used to select principal components (PCs) with statistical differences and the highest contribution rate. PCA is an unsupervised learning algorithm. In general, the first few

PCs can contain the variance of most raw Raman spectral data，which may identify patterns or clustering in the data set [23]. Although PCA can effectively extract data features, its recognition ability is low. Therefore, the diagnostic model was established by the support vector machine (SVM) algorithm. For SVM classifications, the LIBSVM 3.22 [24] package was used, and the Linear kernel function and Gaussian radial basis kernel function (RBF) were selected as the comparison. The model was evaluated by leave-one-out cross-validation (LOOCV) methods. LOOCV means to use one sample in the data set as validation data and the remaining spectra as training data. Our samples of cholangitis, CCA and healthy people have all been verified by the LOOCV methods [23]. Matlab R2020a completed the whole data analysis process. Different classification diagnosis models' diagnostic accuracy, specificity, and sensitivity were analyzed and compared.

3. Results and discussion

3.1. Spectral preprocessing results

Figure 1 shows the spectral preprocessing of a randomly selected sample of the healthy control group. During the spectral collection of serum samples, some noise will inevitably be introduced due to instrument and external environment noise. Therefore, before establishing the classification diagnosis model, it is necessary to carry out de-noising processing to improve the signal-to-noise ratio of spectral signals. In addition, for biological samples, the generation of Raman signal is accompanied by the generation of the self-fluorescence signal, which is reflected in the baseline drift in the spectrum. Hence, the baseline correction of the spectrum is required after noise removal. Finally, to eliminate the influence of different dimensions, the vector normalization operation is carried out after the baseline correction. The spectral preprocessing was carried out in the order of (a)→(b)→(c). (Note: the same pre-processing process was used for all spectra.)

Journal Pre-proof

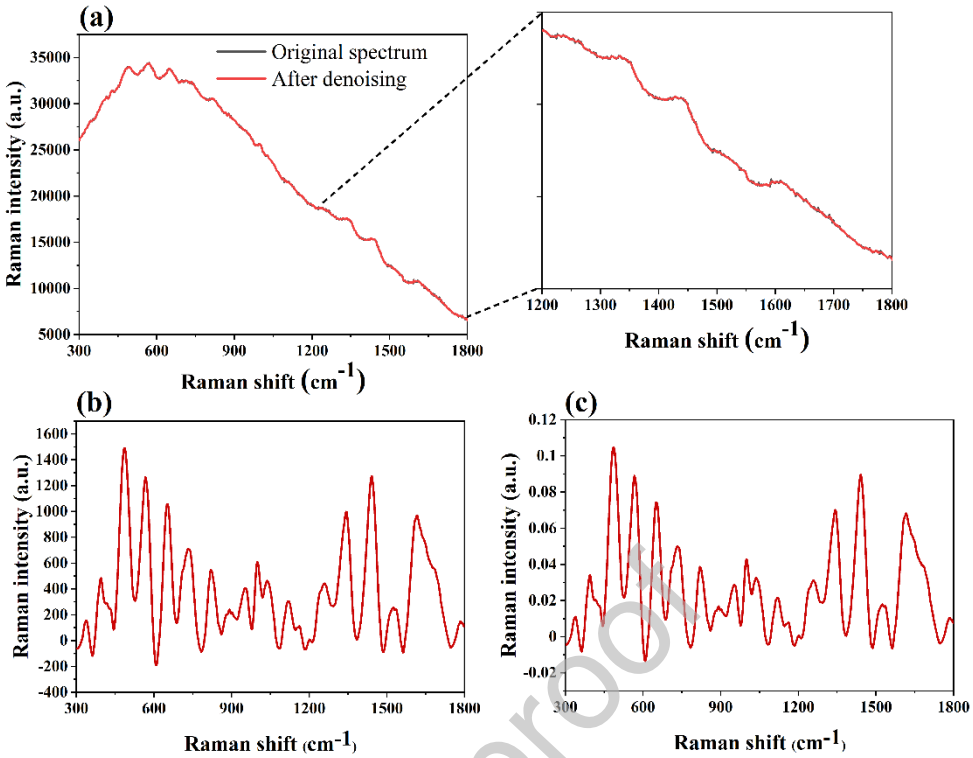


Figure1. Raman spectral pre-processing results of a randomly selected representative sample ;(a) Original spectrum and the original spectrum after denoising ;(b) Baseline corrected spectrum ;(c) Normalized spectrum.

3.2. Raman spectral analysis

In this study, the spectrum of the serum sample was recorded in the characteristic spectral region (300-1800 cm^{-1}) because it contains the most useful information for diagnosis. Figure 2(a) shows normalized mean Raman spectra of serum from patients with cholangitis, CCA, and healthy volunteers in the 300-1800 cm^{-1} wavelength range, in which the shaded areas represent the standard deviations of the means. The main characteristic peaks of serum Raman spectra of the three groups were 415, 486, 566, 650, 736, 886, 941, 1003, 1036, 1127, 1254, 1340, 1445, 1530 and 1654 cm^{-1}. The attribution of each spectral peak was shown in Table 2. The spectrum contains the most information about various amino acids, DNA/RNA bases, glucose, β-carotene, and other biomolecules in human serum [25]. As the spectral profiles of the three groups were relatively close, to observe the spectral differences between them more

Journal Pre-proof

clearly, the difference spectra of the average spectra of the three groups were given in Figure 2(b). As can be seen from the difference spectrogram, the Raman spectra of serum of CCA patients were significantly different from those of cholangitis patients and healthy controls. Serum samples from patients with cholangitis and CCA showed high spectral intensity at 486, 566, 650, 736, and 1036 cm^{-1}. However, it was lower at 886, 1003, 1445, and 1654 cm^{-1} than the healthy control group. These differences reflect the changes in serum biomolecular components in CCA patients and cholangitis patients during the progression of the disease.

The Raman spectrum peaks of Guanine (486 cm^{-1}) and Thymine (736 cm^{-1}) in the serum of the CCA group were significantly higher than those of normal healthy people, indicating that the nucleic acid content in the blood of CCA patients was higher than that of normal healthy people. There were many literature reports that the same trend has been found in other cancer research, such as colorectal cancer [26], nasopharyngeal carcinoma [27], and breast cancer [28]. Abnormal changes in the content of free DNA in serum may be detected when human cancer occurs. As for the increased free nucleic acid content observed in the blood of cancer patients, it has not been fully understood. One possible explanation is that cells release free DNA into the blood during apoptosis, necrosis, or breakage [29]. It is suggested that the detection of free DNA Raman signal in serum may be used to monitor the occurrence and development of cancer.

The 566cm^{-1} peak and 650 cm^{-1} peak in the serum Raman spectrum were classified as Tryptophan and Glucose, respectively. They were significantly up-regulated in the sera of patients with CCA and cholangitis and more significantly in the sera of patients with CCA. Glucose was a factor promoting inflammation, and tryptophan was a crucial regulator of inflammation and immunity. Relevant studies have shown that tryptophan metabolism may be the body's defense response to reduce the inflammatory response. Some key enzymes and products in the process of tryptophan metabolism can participate in immune regulation and then affect the tumor immune microenvironment, which usually occurs in cancer development [30].

Journal Pre-proof

At 886 cm^{-1}(Glutathione) and 941 cm^{-1}(Valine, Proline) peaks, the serum Raman spectral signal of CCA patients was significantly lower than that of normal healthy persons, indicating that the percentage of amino acids in the serum of CCA patients decreased. Similar changes have been seen in other cancer areas [31]. The Raman bands of 415 cm^{-1}(phospholipid), 1445 cm^{-1}(Phospholipid, Collagen), and 1654 cm^{-1}(Collagen) in serum of CCA patients show lower percentage signals than those of Healthy, suggesting that cancer patients may be caused by tumors that metabolize aggressively and consume large amounts of fat, a similar phenomenon found in other cancer studies [17, 32]. At the peak of 1003 cm^{-1} (Collagen), the Raman signal in the CCA group was lower than in the healthy group. The Collagen was lower in the level of serum proteins in the CCA group than in the normal group with associated the depletion of cytoplasmic mucin, resulting in the concentration of metalloproteinase was increased in the process of cancer cell dysplasia, and thereby the rupture of Collagen in the stromal layer [33]. At 1036cm^{-1} (Phenylalanine), the serum Raman signal in the CCA group was higher than that in the healthy group. It can be seen that the content of Phenylalanine in the serum of CCA patients was significantly higher than that of normal people. Similar situations were also found in cervical cancer and other related cancer fields [34].

At 1127 cm^{-1} (Amide II) and 1254 cm^{-1} (Amide III) peaks, the serum signal intensity of CCA group was higher than that of normal healthy group, indicating that the concentration of Amide was increased in the total Raman-active components of CCA samples. This change could be attributed to vibrational modes of the amino acid bonds of the secondary structure of proteins [35]. Moreover, Depciuch et al., also showed changes in the Raman spectra in the range corresponding to amide II and amide I vibrations, they noticed increase of the protein vibrations in CCA [36]. At 1530 cm^{-1} (β-carotene) peaks, the serum Raman signal intensity of the CCA and cholangitis groups was lower than that of the healthy group. The decreased carotene of the diseased serum samples was consistent with previous research, including some cancers [37,38]. In a word, relevant studies have shown that due to the specific anatomical position of the bile duct, when the bile duct is damaged, the metabolic

Journal Pre-proof

function of the liver will be directly affected due to the obstruction of bile flow, which will lead to the disorder of bilirubin metabolism, lipid metabolism and amino acid metabolism [39]. The above differences revealed that the changes in serum between patients with CCA, and cholangitis could be detected by Raman spectra, suggesting that serum Raman spectra technology can provide a potential clinical evaluation methodology for CCA and cholangitis screening.

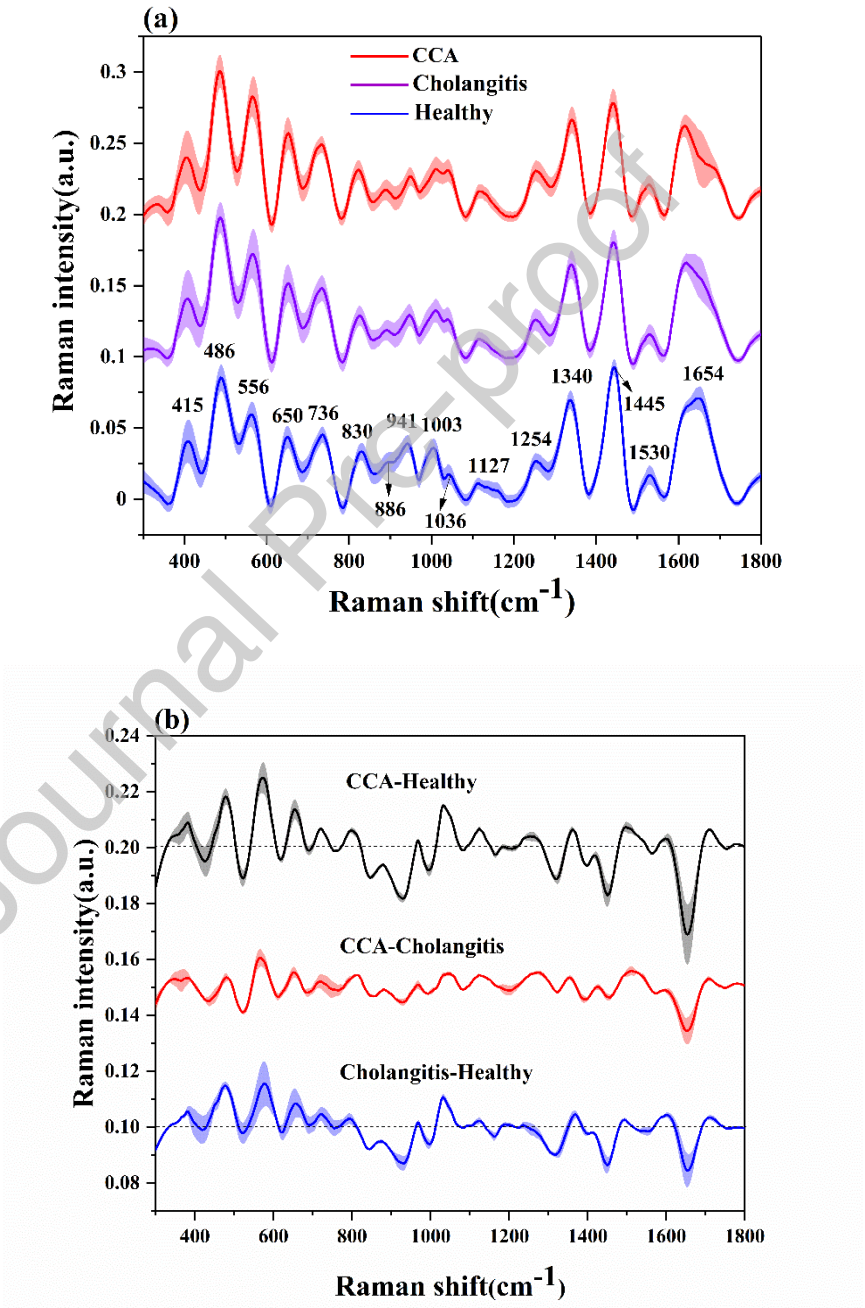


Figure 2. (a)Comparison of the normalized average Raman spectra from Cholangitis, CCA and healthy volunteers' serum samples in the different bands;(b) The difference spectra calculated from the mean

spectra among the three groups. Shade area represents the standard deviations.

Table 2. The tentative assignments for serum Raman spectra

Raman shift (cm^{-1})	Tentative assignment
415	phospholipid
486	DNA, guanine
556	Tryptophan
650	D-glucose, lactose
736	Thymine
886	Glutathione，
941	Valine, proline
1003	Phenylalanine (of collagen)
1036	Phenylalanine，
1127	Amide II, Protein
1254	Amide III Proteins
1340	Nucleic acid mode
1445	Collagen, phospholipid
1530	β-carotene
1654	Collagen

"Assignments taken from References [17, 40-47]"

3.3. Classification results based on PCA and SVM methods

To reduce the Raman spectra data's complexity and multidimensionality, PCA was conducted on the cholangitis, CCA, and healthy volunteers' serum samples Raman spectrum data, in a range of 300-1800 cm^{-1}. After computing the PCs, Analysis of Variance (ANOVA) was used to select PCs with statistical differences and the highest contribution rate (PC1, 50.2%; PC2, 18.7%; [$p < 0.05$]). Figure. 3 showed the two-dimensional scatter plots of principal component scores of serum Raman spectra. According to PC1, spectra of the cholangitis, CCA, and healthy volunteers were differentiated demonstrating that Raman spectroscopy of serum samples has the capability to discriminate the cholangitis, CCA, and healthy volunteers. Further, for each pre-processed spectrum, the SVM algorithm was used to establish a classification diagnosis model, and the retention cross-validation (LOOCV) method

Journal Pre-proof

was used to verify the model's performance. A total of 142 cases were collected, including 49 cases in the Cholangitis group, 38 cases in the CCA group, and 55 in the healthy volunteer group. The SVM algorithm of the two functions has the same overall accuracy for the classification of the three groups, 83.80%. The detailed results were shown in Table 3. Specifically, for 55 spectra of healthy people, 52 were correctly classified, 1 was classified as CCA, and 2 was classified as Cholangitis. Of the 38 CCA spectra, 29 spectra were successfully classified, 2 were classified as Healthy and 7 were classified as Cholangitis. For 49 Cholangitis spectra, 38 spectra were successfully classified, 4 were classified as Healthy, and 7 were classified as CCA. Thus, the diagnostic sensitivities of 94.55%, 76.32%, and 77.55%, and specificities of 77.01%, 86.54%, and 87.10%, respectively, were yielded for differentiation three types of serum Raman spectra. Cholangitis, as a kind of risk factor for CCA, has become a chronic stimulus leading to the occurrence of malignant tumors. Therefore, it is of positive significance to carry out the differential diagnosis of cholangitis and other bile duct diseases as soon as possible [48]. Compared with the current conventional diagnostic tools for bile duct diseases, the results of this study suggest that CCA, cholangitis and healthy people can be differentiated at one time.

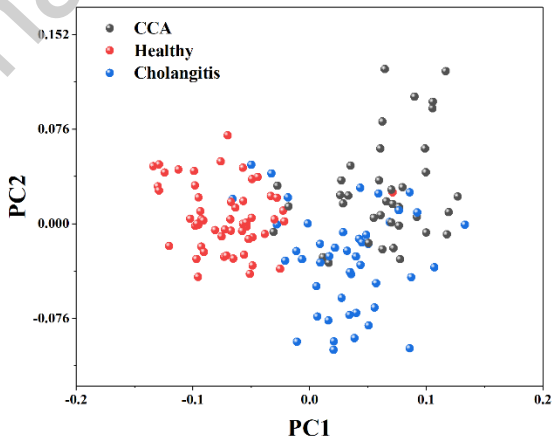


Figure 3. Two-dimensional scatter plots of principal component scores of serum Raman spectra

Table3. Cross validation results of SVM based on Gaussian radial and linear basis function

	Actual group	LOOCV		
		Healthy	CCA	Cholangitis
TRUE	Healthy	52	1	2

CCA	2	**29**	7
Cholangitis	4	7	**38**
Sensitivity	94.55%	76.32%	77.55%
Specificity	77.01%	86.54%	87.10%
Accuracy		83.80%	

Besides, we also explore the effect of dichotomies based on serum Raman spectroscopy. Table 4 shows the results of SVM leave-one cross-validation binary classification. The healthy group was compared with CCA and cholangitis group. The sensitivity and specificity of classification diagnosis were 89.50%, 98.18%, and 89.80%, 94.55%, respectively. The sensitivity and specificity of CCA group and cholangitis group were 78.30% and 83.70%, respectively. Finally, we combined the cholangitis group and CCA into the case group, and the corresponding classification sensitivity, specificity, and accuracy were 92.72%, 94.55%, and 94.36%, respectively. Serum levels of CA 19-9, a tumor biomarker, can aid in diagnosing CCA, but this assay detects CCA with only 62% sensitivity and 63% specificity [49]. Moreover, increased levels of CA 19-9 were also observed in patients with benign diseases such as bacterial cholangitis or choledocholithiasis [50].

The results of this study showed that the sensitivity (89.50%) and specificity (98.18%) of CCA diagnosis were higher than the above commonly used CCA diagnosis methods. It suggests the potential clinical significance of serum Raman spectroscopy combined with machine learning algorithm in differential diagnosis of bile duct diseases. This paper is a preliminary study on the detection of CCA and cholangitis by serum Raman spectroscopy. In the next step, we will conduct further in-depth studies on a broader range of research objects.

Table4. Cross validation results of SVM based on Gaussian radial basis function

Comparison groups	**SVM**		
	Sensitivity	**Specificity**	**Accuracy**
Healthy vs CCA	89.50%	98.18%	94.62%
Healthy vs Cholangitis	89.80%	94.55%	93.27%
CCA vs Cholangitis	78.30%	83.70%	80.46%
Healthy vs Cases*	92.72%	94.55%	94.36%

*Cases, CCA and Cholangitis group.

Journal Pre-proof

4. Conclusion

This study is based on the potential label-free serum detection of serum Raman spectrum and combined with machine learning methods to realize the rapid diagnosis of cholangitis and CCA. We found significant spectral differences between the average Raman spectra of serum from patients with cholangitis, CCA, and healthy volunteers. This indicates that some biomolecular components in serum were changed in humans suffering from bile duct diseases. They have potential value as diagnostic markers of CCA or cholangitis. With the portable Raman spectrometer, the Raman spectrum data of patients with cholangitis, CCA, and healthy volunteers can be obtained with only a small amount of serum. The difference between the spectra can reflect the changes of biochemical components in patients' serum compared with healthy people. Combined with SVM algorithm, Raman spectra of different serum types were classified, and good diagnostic accuracy was obtained. Compared with traditional clinical diagnostic techniques, serum Raman spectroscopy has the advantages of rapid and non-invasive. This preliminary study is expected to provide a new method for the clinical detection of CCA and cholangitis.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest related to this article.

Acknowledgments

This work was supported by the Natural Science Foundation of Xinjiang Uygur Autonomous Region(2018D01C195).

References

[1] Sarcognato S, Sacchi D, Fassan M, et al. Cholangiocarcinoma[J]. Pathologica, 2021,113(3):158-169. https://doi.org/10.32074/1591-951X-252.

[2] Blechacz B. Cholangiocarcinoma: Current Knowledge and New Developments[J]. Gut and liver, 2017,11(1):13-26. https://doi.org/ 10.5009/gnl15568.

[3] Boonstra K, Beuers U, Ponsioen C Y. Epidemiology of primary sclerosing cholangitis and primary

Journal Pre-proof

biliary cirrhosis: A systematic review[J]. Journal of Hepatology, 2012,56(5):1181-1188. https://doi.org/ 10.1016/j.jhep.2011.10.025.

[4] Takakura W R, Tabibian J H, Bowlus C L. The evolution of natural history of primary s clerosing cholangitis[J]. Current opinion in gastroenterology, 2017,33(2):71-77. https://doi.org/10.1097/MOG.0000000000000333.

[5] Fung B M, Tabibian J H. Cholangiocarcinoma in patients with primary sclerosing cholang itis[J]. Current Opinion in Gastroenterology,2020,36(2):77-84. https://doi.org/10.1097/MOG.0000000000000616.

[6] Rizvi S, Khan S A, Hallemeier C L, et al. Cholangiocarcinoma - evolving concepts and t herapeutic strategies[J]. Nature reviews. Clinical oncology, 2018,15(2):95-111. https://doi.org/ 10.1038/nrclinonc.2017.157.

[7] Gulamhusein A F, Eaton J E, Tabibian J H, et al. Duration of Inflammatory Bowel Disea se Is Associated with Increased Risk of Cholangiocarcinoma in Patients with Primary Scle rosing Cholangitis and IBD[J]. Am J Gastroenterol, 2016,111(5):705-711. https://doi.org/ 10.1038/ajg.2016.55.

[8] Lindor K D, Kowdley K V, Harrison E M. ACG Clinical Guideline: Primary Sclerosing Cholangitis[J]. American Journal of Gastroenterology, 2015,110(5):646-659. https://doi.org/ 10.1038/ajg.2015.112.

[9] Aabakken L, Karlsen T, Albert J, et al. Role of endoscopy in primary sclerosing cholangitis: European Society of Gastrointestinal Endoscopy (ESGE) and European Association for the Study of the Liver (EASL) Clinical Guideline[J]. Endoscopy, 2017,49(06):588-608. https://doi.org/ 10.1055/s-0043-107029.

[10] Goeppert B, Toth R, Singer S, et al. Integrative Analysis Defines Distinct Prognostic Subgroups of Intrahepatic Cholangiocarcinoma[J]. Hepatology (Baltimore, Md.), 2019,69(5):2091-2106. https://doi.org/ 10.1002/hep.30493.

[11] Folseraas T, Boberg K M. Cancer Risk and Surveillance in Primary Sclerosing Cholangitis[J]. Clinics in Liver Disease, 2016,20(1):79-98. https://doi.org/ 10.1016/j.cld.2015.08.014.

[12] Baj J, Bryliński A, Woliński F, et al. Biomarkers and Genetic Markers of Hepatocellular Carcinoma and Cholangiocarcinoma-What Do We Already Know[J]. Cancers, 2022,14(6):1493. https://doi.org/ 10.3390/cancers14061493.

[13] Lü G, Zheng X, Lü X, et al. Label-free detection of echinococcosis and liver cirrhosis based on serum Raman spectroscopy combined with multivariate analysis[J]. Photodiagnosis and photodynamic therapy, 2021,33:102164. https://doi.org/ 10.1016/j.pdpdt.2020.102164.

[14] Li Z, Jiang S, Huo Y, et al. 3D silver nanoparticles with multilayer graphene oxide as a spacer for surface enhanced Raman spectroscopy analysis[J]. Nanoscale, 2018,10(13):5897-5905. https://doi.org/10.1039/C7NR09276H.

[15] Li C, Liu A, Zhang C, et al. Ag gyrus-nanostructure supported on graphene/Au film with nanometer gap for ideal surface enhanced Raman scattering[J]. Optics express, 2017,25(17):20631-20641. https://doi.org/ 10.1364/OE.25.020631.

Journal Pre-proof

[16] Žukovskaja O, Ryabchykov O, Straßburger M, et al. Towards Raman spectroscopy of urine as screening tool[J]. Journal of Biophotonics, 2020,13(1). https://doi.org/ 10.1002/jbio.201900143.

[17] Li S, Zhang Y, Xu J, et al. Noninvasive prostate cancer screening based on serum surface-enhanced Raman spectroscopy and support vector machine[J]. Applied Physics Letters, 2014,105(9):91104. https://doi.org/ 10.1063/1.4892667.

[18] Lin X, Wu Q, Lin J, et al. Label-free liquid biopsy based on urine analysis using surface-enhanced Raman spectroscopy for noninvasive gastric and breast cancer detection[J]. Journal of Raman spectroscopy, 2020,51(11):2245-2254. https://doi.org/ 10.1002/jrs.5978.

[19] Li S X, Zeng Q Y, Li L F, et al. Study of support vector machine and serum surface-enhanced Raman spectroscopy for noninvasive esophageal cancer detection[J]. Journal of biomedical optics, 2013,18(2):27008. https://doi.org/ 10.1117/1.JBO.18.2.027008.

[20] Tong D, Chen C, Zhang J, et al. Application of Raman spectroscopy in the detection of hepatitis B virus infection[J]. Photodiagnosis and Photodynamic Therapy, 2019,28:248-252. https://doi.org/10.1016/j.pdpdt.2019.08.006.

[21] Zhang Z, Chen S, Liang Y. Baseline correction using adaptive iteratively reweighted penalized least squares[J]. Analyst (London), 2010,135(5):1138-1146. https://doi.org/ 10.1039/b922045c.

[22] Gautam R, Vanga S, Ariese F. Review of multidimensional data processing approaches for Raman and infrared spectroscopy[J]. EPJ techniques and instrumentation, 2015,2(1):1. https://doi.org/ 10.1140/epjti/s40485-015-0018-6.

[23] Lussier F, Thibault V, Charron B, et al. Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering[J]. TrAC, Trends in analytical chemistry (Regular ed.), 2020,124:115796. https://doi.org/10.1016/j.trac.2019.115796.

[24] Dawuti W, Zheng X, Liu H, et al. Urine surface-enhanced Raman spectroscopy combined with SVM algorithm for rapid diagnosis of liver cirrhosis and hepatocellular carcinoma[J]. Photodiagnosis and photodynamic therapy, 2022,38:102811. https://doi.org/ 10.1016/j.pdpdt.2022.102811.

[25] Ralbovsky N M, Lednev I K. Towards development of a novel universal medical diagnostic method: Raman spectroscopy and machine learning[J]. Chemical Society reviews, 2020,49(20):7428-7453. https://doi.org/ 10.1039/d0cs01019g.

[26] Lin J, Huang Z, Lin X, et al. Rapid and label-free urine test based on surface-enhanced Raman spectroscopy for the non-invasive detection of colorectal cancer at different stages[J]. Biomedical Optics Express, 2020,11(12):7109. https://doi.org/ 10.1364/BOE.406097.

[27] Feng S, Chen R, Lin J, et al. Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis[J]. Biosensors and Bioelectronics, 2010,25(11):2414-2419. https://doi.org/ 10.1016/j.bios.2010.03.033.

[28] Lin Y, Gao J, Tang S, et al. Label-free diagnosis of breast cancer based on serum protein purification assisted surface-enhanced Raman spectroscopy[J]. Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy, 2021,263:120234. https://doi.org/ 10.1016/j.saa.2021.120234.

Journal Pre-proof

[29] Gormally E, Caboux E, Vineis P, et al. Circulating free DNA in plasma or serum as biomarker of carcinogenesis: Practical aspects and biological significance[J]. Mutation research, 2007,635(2-3):105-117. https://doi.org/ 10.1016/j.mrrev.2006.11.002.

[30] Satriano L, Lewinska M, Rodrigues P M, et al. Metabolic rearrangements in primary liver cancers: cause and consequences[J]. Nat Rev Gastroenterol Hepatol, 2019,16(12):748-766. https://doi.org/ 10.1038/s41575-019-0217-8.

[31] Feng S, Chen R, Lin J, et al. Gastric cancer detection based on blood plasma surface-enhanced Raman spectroscopy excited by polarized laser light[J]. Biosens Bioelectron, 2011,26(7):3167-3174. https://doi.org/ 10.1016/j.bios.2010.12.020.

[32] Maziak D E M, Do M T M, Shamji F M M, et al. Fourier-transform infrared spectroscopic study of characteristic molecular structure in cancer cells of esophagus: An exploratory study[J]. Cancer detection and prevention, 2007,31(3):244-253. https://doi.org/ 10.1016/j.cdp.2007.03.003

[33] Wei Zheng K Y H M. Near-infrared Raman spectroscopy for gastric precancer diagnosis[J]. Journal of Raman spectroscopy, 2009,48(8):908-914. https://doi.org/ 10.1002/jrs.2197.

[34] Feng S, Lin D, Lin J, et al. Blood plasma surface-enhanced Raman spectroscopy for non-invasive optical detection of cervical cancer[J]. Analyst, 2013,138(14):3967-3974. https://doi.org/ 10.1039/c3an36890d.

[35] Talari A C S, Movasaghi Z, Rehman S, et al. Raman Spectroscopy of Biological Tissues[J]. Applied Spectroscopy Reviews, 2014,50(1):46-111. https://doi.org/ 10.1080/05704928.2014.923902.

[36] Depciuch J, Parlinska-Wojtan M, Rahmi S K, et al. Differential of cholangiocarcinoma disease using Raman spectroscopy combined with multivariate analysis[J]. Spectrochim Acta A Mol Biomol Spectrosc, 2022,272:121006. https://doi.org/ 10.1016/j.saa.2022.121006.

[37] Gonzalez-Solis J L, Martinez-Espinosa J C, Torres-Gonzalez L A, et al. Cervical cancer detection based on serum sample Raman spectroscopy[J]. Lasers in medical science, 2014,29(3):979-985. https://doi.org/ 10.1007/s10103-013-1447-6.

[38] Sahu A, Sawant S, Mamgain H, et al. Raman spectroscopy of serum: an exploratory stud y for detection of oral cancers[J]. Analyst, 2013,138(14):4161-4174. https://doi.org/ 10.1039 /c3an00308f.

[39] Krasinskas A M. Cholangiocarcinoma[J]. Surgical Pathology Clinics, 2018,11(2):403-429. https://doi.org/ 10.1016/j.path.2018.02.005.

[40] Gao N, Wang Q, Tang J, et al. Non-invasive SERS serum detection technology combined with multivariate statistical algorithm for simultaneous screening of cervical cancer and breast cancer[J]. Analytical and Bioanalytical Chemistry, 2021,413(19):4775-4784. https://doi.org/ 10.1007/s00216-021-03431-3.

[41] Lilo T, Morais C L M, Shenton C, et al. Revising Fourier-transform infrared (FT-IR) and Raman spectroscopy towards brain cancer detection[J]. Photodiagnosis and photodynamic therapy, 2022,38:102785. https://doi.org/ 10.1016/j.pdpdt.2022.102785.

[42] Xiao R, Zhang X, Rong Z, et al. Non-invasive detection of hepatocellular carcinoma seru

Journal Pre-proof

m metabolic profile through surface-enhanced Raman spectroscopy[J]. Nanomedicine: Nanot echnology, Biology and Medicine, 2016,12(8):2475-2484. https://doi.org/ 10.1016/j.nano.201 6.07.014.

[43] Paraskevaidi M, Morais C L M, Halliwell D E, et al. Raman Spectroscopy to Diagnose Alzheimer's Disease and Dementia with Lewy Bodies in Blood[J]. ACS chemical neuroscience, 2018,9(11):2786-2794. https://doi.org/ 10.1021/acschemneuro.8b00198.

[44] Feng S, Pan J, Wu Y, et al. Study on gastric cancer blood plasma based on surface-enha nced Raman spectroscopy combined with multivariate analysis[J]. Science China. Life scie nces, 2011,54(9):828-834. https://doi.org/ 10.1007/s11427-011-4212-8.

[45] Patel S K, Rajora N, Kumar S, et al. Rapid Discrimination of Malaria- and Dengue-Infected Patients Sera Using Raman Spectroscopy[J]. Analytical chemistry (Washington), 2019,91(11):7054-7062. https://doi.org/ 10.1021/acs.analchem.8b05907.

[46] Zheng X, Lv G, Zhang Y, et al. Rapid and non-invasive screening of high renin hyperten sion using Raman spectroscopy and different classification algorithms[J]. Spectrochimica A cta Part A: Molecular and Biomolecular Spectroscopy, 2019,215:244-248. https://doi.org/ 1 0.1016/j.saa.2019.02.063.

[47] Ullah R, Khan S, Farman F, et al. Demonstrating the application of Raman spectroscopy together with chemometric technique for screening of asthma disease[J]. Biomedical optics express, 2019,10(2):600-609. https://doi.org/ 10.1364/BOE.10.000600.

[48] Khan S A, Tavolari S, Brandi G. Cholangiocarcinoma: Epidemiology and risk factors[J]. Liver International, 2019,39(S1):19-31. https://doi.org/10.1111/liv.14095.

[49] Blechacz B, Gores G J. Cholangiocarcinoma: Advances in pathogenesis, diagnosis, and tre atment[J]. Hepatology (Baltimore, Md.), 2008,48(1):308-321. https://doi.org/ 10.1002/hep.223 10.

[50] Moazzami B, Majidzadeh-A K, Dooghaie-Moghadam A, et al. Cholangiocarcinoma: State of the Art[J]. Journal of Gastrointestinal Cancer, 2020,51(3):774-781. https://doi.org/ 10.1007/s12029-020-00390-3.

Journal Pre-proof