Hepatology International (2023) 17:339–349
https://doi.org/10.1007/s12072-022-10444-2

ORIGINAL ARTICLE

Fully connected neural network-based serum surface-enhanced Raman spectroscopy accurately identifies non-alcoholic steatohepatitis

Feng Gao[1] · De-Chan Lu[2] · Tian-Lei Zheng[3,4] · Shi Geng[3] · Jun-Cheng Sha[5] · Ou-Yang Huang[6] · Liang-Jie Tang[6] · Pei-Wu Zhu[7] · Yang-Yang Li[8] · Li-Li Chen[6] · Giovanni Targher[9] · Christopher D. Byrne[10] · Zu-Fang Huang[2] · Ming-Hua Zheng[6,11,12] on behalf of CHESS-MAFLD consortium

Received: 13 August 2022 / Accepted: 23 October 2022 / Published online: 11 November 2022

Abstract

Background/purpose of the study There is a need to find a standardized and low-risk diagnostic tool that can non-invasively detect non-alcoholic steatohepatitis (NASH). Surface enhanced Raman spectroscopy (SERS), which is a technique combining Raman spectroscopy (RS) with nanotechnology, has recently received considerable attention due to its potential for improving medical diagnostics. We aimed to investigate combining SERS and neural network approaches, using a liver biopsy dataset to develop and validate a new diagnostic model for non-invasively identifying NASH.

Methods Silver nanoparticles as the SERS-active nanostructures were mixed with blood serum to enhance the Raman scattering signals. The spectral data set was used to train the NASH classification model by a neural network primarily consisting of a fully connected residual module.

Results Data on 261 Chinese individuals with biopsy-proven NAFLD were included and a prediction model for NASH was built based on SERS spectra and neural network approaches. The model yielded an AUROC of 0.83 (95% confidence interval [CI] 0.70–0.92) in the validation set, which was better than AUROCs of both serum CK-18-M30 levels (AUROC 0.63, 95% CI 0.48–0.76, $p=0.044$) and the HAIR score (AUROC 0.65, 95% CI 0.51–0.77, $p=0.040$). Subgroup analyses showed that the model performed well in different patient subgroups.

Conclusions Fully connected neural network-based serum SERS analysis is a rapid and practical tool for the non-invasive identification of NASH. The online calculator website for the estimated risk of NASH is freely available to healthcare providers and researchers (http://www.pan-chess.cn/calculator/RAMAN_score).

Keywords NAFLD (non-alcoholic fatty liver disease) · NASH (non-alcoholic steatohepatitis) · Raman · SERS (surface-enhanced Raman spectroscopy)

Abbreviations

AgNPs	Silver nanoparticles
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
AUROC	Area under the receiver operating characteristic curve
BMI	Body mass index
CI	Confidence interval
CK-18	Cytokeratin-18 fragments
CT	Computed tomography
FCNN	Fully connected neural network
GGT	γ-Glutamyltranspeptidase
HOMA-IR	Homeostasis model assessment of insulin resistance
HDL-C	High-density lipoprotein cholesterol
LDL-C	Low-density lipoprotein cholesterol
MetS	Metabolic syndrome
MRE	Magnetic resonance elastography
NAFL	Non-alcoholic fatty liver

Feng Gao, De-Chan Lu and Tian-Lei Zheng contributed equally to this work.

Article guarantor**:** Ming-Hua Zheng and Zu-Fang Huang.

✉ Zu-Fang Huang
zfhuang@fjnu.edu.cn

✉ Ming-Hua Zheng
zhengmh@wmu.edu.cn

Extended author information available on the last page of the article

NAFLD	Non-alcoholic fatty liver disease
NASH	Non-alcoholic steatohepatitis
NAS	NAFLD activity score
TC	Total cholesterol
TG	Triglycerides
T2DM	Type 2 diabetes mellitus
SERS	Surface enhanced Raman spectroscopy

Introduction

Non-alcoholic fatty liver disease (NAFLD) is a public health threat that affects about one third of the global adult population [1]. Non-alcoholic steatohepatitis (NASH) is the progressive and inflammatory subtype of NAFLD, characterized by steatosis, lobular inflammation, and hepatocellular ballooning [2]. In recent years, NASH has become the most rapidly increasing indication for liver transplantation in the United States and Europe [3]. In addition, NASH is closely associated with an increased risk of developing important extra-hepatic diseases, such as cardiovascular disease, chronic kidney disease, and type 2 diabetes mellitus (T2DM) [4–7]. Early lifestyle intervention is associated with a favorable prognosis in patients with NASH, but most of these patients are diagnosed late with advanced fibrosis or even hepatocellular carcinoma, which has a poor survival rate [8].

To date, liver biopsy (LB) remains the 'gold standard' for the diagnosis of NASH, but this is an invasive diagnostic procedure with associated risks, such as post-procedural bleeding [9]. In recent years, some biomarkers and imaging techniques have been investigated to predict NASH. Serum levels of cytokeratin-18 fragments (CK-18 neoepitope M30), reflecting the degree of apoptosis, have been proposed as a differentiator between non-alcoholic fatty liver (NAFL) and NASH, though its performance to predict NASH is moderate [10, 11]. As the pathogenesis of NASH is complex and involves several biological pathways, it is unlikely that a single biomarker could accurately discriminate between NAFL and NASH [12]. Thus, there is currently a need to find a non-invasive and low-risk diagnostic tool that can accurately detect NASH at an early stage [13].

Surface enhanced Raman spectroscopy (SERS) is a technique combining Raman spectroscopy (RS) with nanotechnology that has recently received considerable scientific attention, due to its great potential for improving clinical diagnosis [14]. RS is an established technique that provides label-free fingerprinting type information on the vibrational and rotational modes of proteins, lipids, and nucleic acids [15]. SERS maintains the advantages of normal RS but enhances the signal intensity of Raman scattering for sensitive detection [16]. In recent years, a number of studies have tested the applicability and effectiveness of SERS technique in various disease diagnoses. For example, Feng et al. reported that SERS is a promising tool for the non-invasive nasopharyngeal cancer detection and screening [17]. Shin et al. reported that SERS could be also used to accurately identify early-stage lung cancer [18]. In addition to early detection of some cancers, SERS is also a promising tool for early diagnosis of non-cancer diseases, such as diabetic keratopathy and heparin-induced thrombocytopenia [19, 20]. Since each molecule produces a unique "spectral fingerprint", we hypothesized that different biomolecules with characteristic Raman signatures in NASH sera could also facilitate the non-invasive identification of NASH.

Fully connected neural network (FCNN) is a data driven application of artificial intelligence, in which systems automatically learn and improve without the need for explicit programming [21]. Accordingly, FCNN is able to extract useful information from high-dimensional and large data sets. FCNN is now widely used to develop prognostic and diagnostic prediction models in the field of medicine [22, 23].

In this exploratory study, we aimed for the first time to investigate combining SERS and FCNN approaches, using a liver-biopsy dataset to develop and validate an improved diagnostic model for non-invasively identifying NASH and then to compare its diagnostic performance with existing validated non-invasive models available in clinical practice (Fig. 1).

Patients and methods

Study population and design

This analysis was undertaken in a subset of our well-characterized Prospective Epidemic Research Specifically of the NASH (PERSONS) cohort. The entire cohort comprised 448 Han Chinese individuals diagnosed with suspected NAFLD (based on the presence of imaging-defined hepatic steatosis and/or persistently elevated serum transaminase levels in subjects with coexisting metabolic risk factors, such as overweight/obesity, T2DM or metabolic syndrome, and without significant alcohol consumption), who were enrolled between November 2016 and September 2019 at the First Affiliated Hospital of Wenzhou Medical University in Wenzhou (China). All these individuals underwent a diagnostic liver biopsy. Subjects with at least one of the following criteria were excluded from the analysis: (i) significant alcohol consumption ($\geq$ 140 g/week in male or $\geq$ 70 g/week in female); (ii) presence of viral hepatitis, autoimmune hepatitis, drug-induced liver injury, or other known chronic liver diseases; (iii) missing important clinical data; and (iv) hepatic fat content $<$ 5% on histology. As a consequence of these exclusion criteria, a total of 261 individuals with

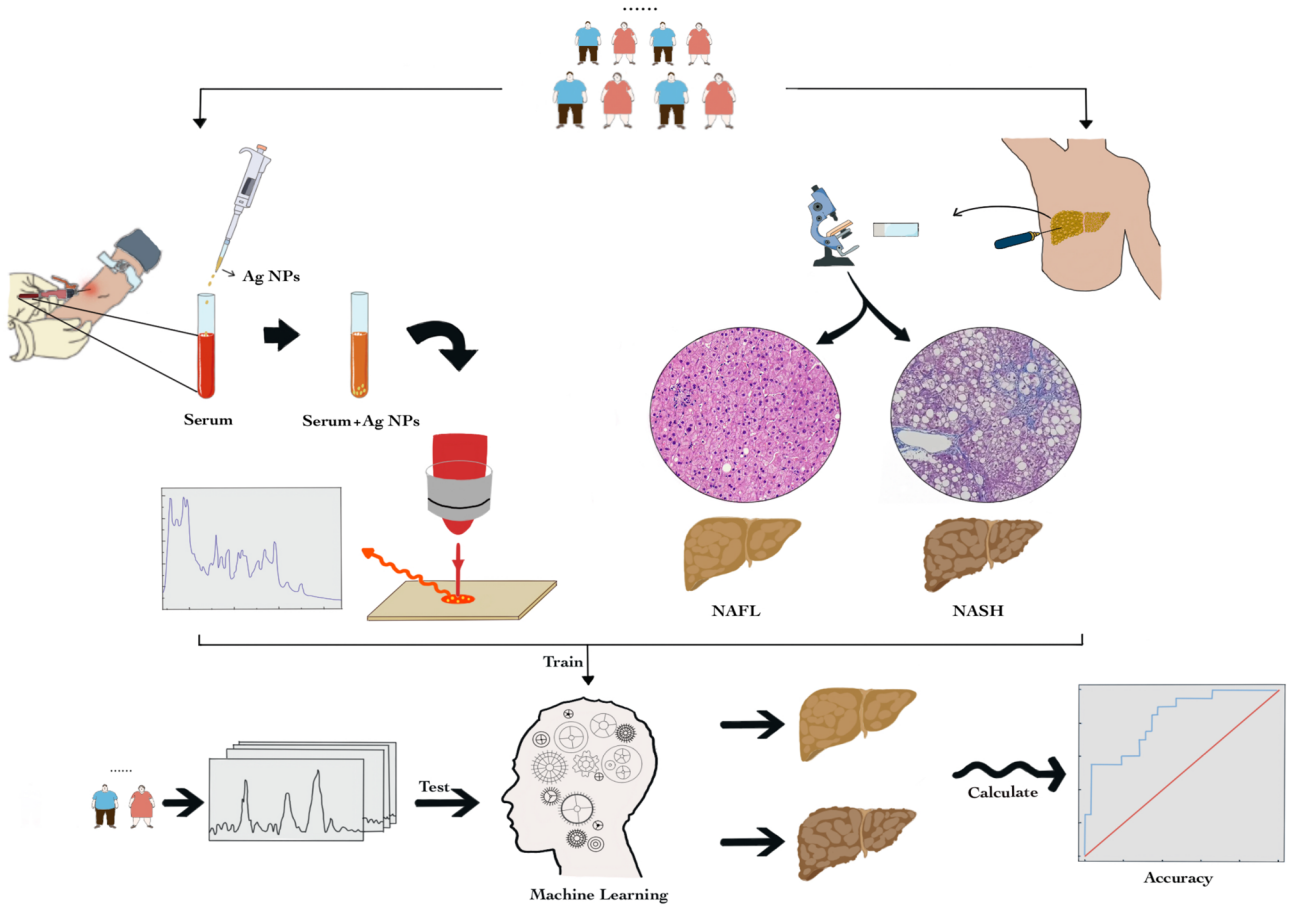

Fig. 1 Schematic illustration of combining silver nanoparticle-based SERS spectroscopy with fully connected neural network analysis to differentiate NASH from NAFL. **a** Collection of spectroscopic data of serum by SERS; **b** Collection of histology data by ultrasound-guided liver biopsy; **c** In the validation cohort, using machine learning to differentiate NAFL from NASH patients

biopsy-proven NAFLD were included in the final analysis (as detailed in Figure S1). These subjects were used as a training cohort ($n = 209$). Another sample of 52 subjects with NAFLD who underwent liver biopsy between April 2019 and September 2019 were used as an independent validation cohort.

Clinical and biochemical data

Clinical and biochemical data were obtained from all participants within 48 h of the liver biopsy. Blood samples were taken after at least 8 h of overnight fasting. Body mass index (BMI) was calculated using the formula weight (kilograms) divided by height (meters) squared. T2DM was diagnosed as either self-reported history of disease, a fasting glucose level ≥ 7.0 mmol/L (≥ 126 mg/dL), hemoglobin A1c ≥ 48 mmol/mol (≥ 6.5%), or use of any antihyperglycemic drugs. Hypertension was defined as systolic blood pressure ≥ 140 mmHg, diastolic blood pressure ≥ 90 mmHg, or use of any anti-hypertensive drugs. Metabolic syndrome (MetS) was defined as having at least three of the following metabolic risk factors: central obesity (waist circumference ≥ 90 cm in male and ≥ 80 cm in female), increased blood pressure (systolic blood pressure ≥ 130 mmHg or diastolic blood pressure ≥ 85 mmHg or anti-hypertensive treatment), increased fasting glucose (≥ 5.6 mmol/L or use of any antihyperglycemic agents), high triglycerides (> 1.7 mmol/L or use of any lipid-lowering drugs), and low high-density lipoprotein cholesterol levels (< 1.03 mmol/L in male and < 1.29 mmol/L in female or use of any lipid-lowering drugs) [24]. Methodological details for measurement of serum cytokeratin-18 fragments (cytokeratin-18 [CK-18] neoepitope M30) concentrations have been reported previously [25]. The HAIR score (0–3) for each subject was calculated by adding hypertension = 1, serum alanine aminotransferase (ALT) level > 40 IU/L = 1, and HOMA-estimated insulin resistance score > 5.0 = 1 [26].

Liver histology

An ultrasound-guided liver biopsy was performed under sedation using a 16-gauge Hepafix needle. The assessment of liver histology was undertaken by a single experienced liver histopathologist (who was blinded to the clinical and laboratory data of participants), according to the NASH-Clinical Research Network (CRN) Scoring System [27]. Steatosis (0–3), ballooning (0–2), and lobular inflammation (0–3) grades as well as fibrosis stage (F0–F4) and NAFLD activity score (NAS) were scored using the NASH-Clinical Research Network (NASH-CRN) scoring system. NAS was calculated as the sum of three histological components, including steatosis, ballooning and lobular inflammation. The diagnosis of definite NASH was established by the presence of hepatic steatosis, lobular inflammation and ballooning with NAS ≥ 5 [28]. Significant fibrosis was defined as stage F ≥ 2, and advanced fibrosis was defined as stage F ≥ 3 [29].

Spectral acquisition

1.5 μL of silver nanoparticles (AgNPs) were added onto the pre-cleaned aluminum plate, immediately the serum with the same volume was then gently added and mixed with AgNPs, and dried at room temperature. The SERS spectra were acquired with Renishaw confocal Raman spectrometer (Renishaw Invia, UK) using a 785 nm laser excitation, and a 1200 lines/mm grating was selected to provide a spectral resolution of ~ 2 cm^{-1}. The excitation laser (~ 2.38 mW) was delivered via a 20× objective, and the spectra were recorded with a 10 s exposure time over the fingerprint region from 400 to 1800 cm^{-1}. The detection of the SERS signal was carried out with a Peltier cooled charge-coupled device (CCD) camera. The software package WIRE 2.0 (Renishaw) was employed for spectral acquisition and analysis. After normalization, the whole SERS spectrum data set was fed into the python for FCNN analysis. For each serum sample, representative Raman spectra was a result of an average of three spectra obtained from random detection sites. More detailed information about the RS and SERS is presented in Supplementary Appendix.

AgNPs were synthesized referring to the Leopold's hydroxylamine hydrochloride reduction method. Briefly, 5 mL of the 60 mM hydroxylamine hydrochloride solution was added to 4.5 mL of the 0.1 M sodium hydroxide solution. The mixture was then added to 90 mL of 10^{-3} M sliver nitrate solution and stirred for 15 min until the color turned gray. Finally, 1 mL of AgNPs was centrifuged (10,000 rpm for 10 min), and the supernatant was removed to obtain 100 times the original concentration of AgNPs. Silver nitrate and hydroxylamine hydrochloride were purchased from Shanghai Aladdin Co., Ltd. Sodium hydroxide was purchased from Shanghai Guoyao Group Chemical Reagent Co., Ltd.

Statistical analysis

Continuous variables were expressed as means ± SD or medians with interquartile ranges (IQRs) and compared using either the unpaired Student's *t* test or the Mann–Whitney U test, as appropriate. Categorical variables were expressed as numbers (or percentages) and compared using the chi-squared test or the Fisher's exact test, as appropriate.

The spectral data set from the training cohort ($n = 206$) is used to train the predicting model for binary classification by FCNN. Long et al. proposed that a FCNN performs pixel-level classification to efficiently solve the problem of semantic-level image segmentation [30]. Unlike the classic CNN that uses the fully connected layer to obtain fixed-length feature vectors for classification after the convolutional layer, FCNN can accept input images of any size, and uses the deconvolutional layer to sample the feature map of the last convolutional layer. It restores the output to the same size as the input image, so that a prediction can be generated for each pixel, while also retaining the spatial information in the original input image, and finally performs pixel-by-pixel classification on the upsampled feature map.

This neural network (FCNN) has three layers: an input layer, a hidden layer, and an output layer. And the input layer has 1000 nodes, the hidden layer has 1568 nodes, and the output layer has 2 nodes. The input layer consists of a full connected layer, which mainly extracts features of input spectral data to generate a shallow feature. The shallow feature will further enter the hidden layer, which is composed of three fully connected residual modules, to generate a deep feature. Each fully connected residual module consists of three fully connected modules and residual structure. The first two fully connected modules include the fully connected layer, the batch normalization layer and the ReLU activation function, and the last module removes the activation function compared with the previous two modules. Batch normalization adds normalization "layer" between each fully connected layer and makes the model more stable by protecting against outlier weights and reducing overfitting. ReLU is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. ReLU activation function helps the model account for interaction and non-linear effects. Each fully connected residual module can be expressed as:

$$x_{\mathrm{m}} = F(x, \{W_1, W_2, W_3\})$$

$$y = \mathrm{Dropout}(ReLU(x_{\mathrm{m}} + x))$$

where W means the weight of the fully connected module and F indicates the combination of three fully connected

modules. First, the fully connected residual module uses three fully connected modules to extract an input feature x to generate an intermediate feature x_m. Then, x is added as a residual to x_m, and using the dropout function to generate y after ReLU activation. Dropout layer works by randomly setting the outgoing edges of hidden units (neurons that make up hidden layers) to 0 at each update of the training phase. Dropout regularization reduces overfitting and improves the generalization of the model. The value for dropout in the fully connected residual module is 0.5. Three fully connected residual modules can increase the depth of the model while suppressing the disappearance of the gradient, thus improving the performance of the model. Finally, the deep feature is sent to the output layer composed of a fully connected layer and a Softmax layer to generate the NASH classification result.

The accuracy of the aforementioned model was evaluated in the validation cohort ($n = 52$) by assessing its discriminatory power and calibration capability. The diagnostic discriminatory capability was assessed by calculating the area under the receiver operating characteristic curve (AUROC), and compared using the DeLong test. The calibration capability was assessed by the calibration curve. Statistical analyses were two-sided and significance was set at $p < 0.05$. Origin 2017 software was used to plot the average spectrum and difference spectrum of the two sets of data (OriginLab Inc., Northampton, USA), and the R (Version 3.3.1 The R Foundation) was used for the heatmap, ROC, and calibration curve plot.

Results

Baseline characteristics of participants

A total of 261 Chinese individuals with biopsy-proven NAFLD were analyzed. The mean age of participants was 42.6 ± 12.3 years and 70.5% of the subjects were male. Their mean value of BMI was 27.0 ± 3.7 kg/m^2. Ninety-two subjects (35.2%) had established T2DM, 135 (51.7%) had hypertension and 206 (78.9%) had MetS, respectively. The median value of the histological NAS score of participants was 4.0 (IQR 3–5). Presence of definite NASH (as defined in the Methods section above) was histologically diagnosed in 117 subjects (44.8%). By study design, 209 subjects and 52 subjects with biopsy-confirmed NAFLD were assigned to the training and validation cohorts according to the time of admission, respectively. Baseline clinical and biochemical characteristics, as well as the NAFLD histology features of participants are summarized in Table 1.

SERS for NASH prediction

The SERS spectra obtained from the patient sera exhibited vibrational features reflecting the molecular composition of serum samples. All measured SERS spectra were normalized to the integrated area under the curve in the 400–1800 cm^{-1} wavenumber range after the removal of fluorescence background from the original SERS data. This reduces the spectral intensity variations between different spectra and facilitates more accurate spectral shape analyses. Figure 2a compares the normalized mean SERS spectra obtained from serum samples of both NAFL and NASH patients. It can be seen that while significant SERS spectral differences exist between NAFL and NASH serum samples, primary SERS peaks at 497, 533, 593, 638, 726, 744, 766, 813, 887, 959, 1005, 1071, 1095, 1133, 1204, 1270, 1330, 1366, 1395, 1443, 1500, 1580, and 1653 cm^{-1} can all be observed in both NAFL and NASH serum samples. However, the normalized intensities of SERS peaks at 726, 744, 959, 1095, 1270, 1330, 1395, 1443, 1580, and 1653 cm^{-1} are greater in NAFL serum samples than in NASH serum samples, while SERS bands at 497, 533, 593, 638, 766, 813, 887, 1005, 1071, 1133, 1204, 1366, and 1500 cm^{-1} are greater in NASH serum samples.

The association between the presence of NASH on liver histology and these selected SERS bands is presented in the heatmap (Fig. 2b). To better understand the molecular basis for the observed SERS of patient sera, supplementary Table 1 lists tentative assignments for the observed SERS bands, according to literature data.

A prediction model for NASH was built by the FCNN. This model yielded an AUROC of 0.83 (95% confidence interval [CI] 0.70–0.92) in the validation cohort, which was significantly better than those of both serum CK-18-M30 levels (AUROC 0.63, 95% CI 0.48–0.76, $p = 0.044$) and the HAIR score (AUROC 0.65, 95% CI 0.51–0.77, $p = 0.040$) for identifying patients with NASH (Fig. 2c). The calibration curve of the model for the probability of having NASH showed good agreement between prediction and observation in the validation cohort (Fig. 2d), which suggests that there was a little departure from a perfect fit.

Demographics (age and sex), metabolic risk factors (BMI, T2DM, hypertension, and MetS) and laboratory parameters (HOMA-IR, serum ALT, glucose, triglycerides, total cholesterol, and CK18-M30 levels) of the patients were included in the FCNN model containing the spectra data. However, the performance of the model in identifying NASH was not significantly improved by adding these additional data (AUROC: 0.834 vs. 0.828; $p > 0.05$). These results could be explained by the fact that each molecule produces a unique "spectral fingerprint" in SERS, and the different biomolecules with characteristic Raman signatures in sera already contain the information

Table 1 Baseline characteristics of participants with biopsy-proven NAFLD

	Overall population $N=261$	Training set $N=209$	Validation set $N=52$	p value
Demographics				
Age, years	42.6 ± 12.3	41.5 ± 12.2	46.9 ± 12.0	0.004
Male, *n* (%)	184 (70.5%)	151 (72.2%)	33 (63.5%)	0.214
Metabolic risk factors				
BMI, kg/m^2	27.0 ± 3.7	27.1 ± 3.7	26.8 ± 3.3	0.635
Type 2 diabetes, *n* (%)	92 (35.2%)	74 (35.4%)	18 (34.6%)	0.915
Hypertension, *n* (%)	135 (51.7%)	106 (50.7%)	29 (55.8%)	0.514
Metabolic syndrome, *n* (%)	206 (78.9%)	162 (77.5%)	44 (84.6%)	0.261
Laboratory parameters				
ALT, IU/L	47 (27–82)	51 (30–888)	34 (23–68)	0.004
Glucose, mmol/L	5.4 (4.9–6.3)	5.3 (4.9–6.2)	5.5 (5.0–6.5)	0.152
HOMA-IR score	3.7 (2.3–5.4)	3.7 (2.3–5.6)	3.7 (2.2–5.0)	0.546
TG, mmol/L	1.8 (1.3–2.6)	1.8 (1.3–2.6)	1.9 (1.4–2.6)	0.571
TC, mmol/L	5.3 ± 1.1	5.3 ± 1.1	5.3 ± 1.2	0.970
HDL-C, mmol/L	1.0 ± 0.2	1.0 ± 0.2	1.0 ± 0.2	0.306
LDL-C, mmol/L	3.1 ± 0.9	3.1 ± 0.9	2.9 ± 0.9	0.142
CK18-M30, IU/L	135 (70–282)	150 (77–317)	87 (59–177)	0.002
Liver histology				
Liver biopsy length (cm)	1.5 (1.5–2.0)	1.5 (1.5–2.0)	1.5 (1.5–2.0)	0.712
Significant fibrosis[a], *n* (%)	61/261 (23.4%)	48/209 (23.0%)	13/52 (25.0%)	0.757
Steatosis grade, *n* (%)				0.185
S1	123 (47.1%)	93 (44.5%)	30 (57.7%)	
S2	85 (32.6%)	73 (34.9%)	12 (23.1%)	
S3	53 (20.3%)	43 (20.6%)	10 (19.2%)	
Ballooning grade, *n* (%)				0.422
B0	32 (12.3%)	28 (13.4%)	4 (7.7%)	
B1	130 (49.8%)	106 (50.7%)	24 (46.2%)	
B2	98 (37.5%)	74 (35.4%)	24 (46.2%)	
Lobular inflammation grade, *n* (%)				0.472
L0	25 (9.6%)	21 (10.0%)	4 (7.7%)	
L1	138 (52.9%)	106 (50.7%)	32 (61.5%)	
L2	95 (36.4%)	80 (38.3%)	15 (28.8%)	
L3	3 (1.1%)	2 (1.0%)	1 (1.9%)	
Histologic NAS score	4 (3–5)	4 (3–6)	4 (3–5)	0.709
NASH[b]	117 (44.8%)	97 (46.4%)	20 (38.5%)	0.302

ALT alanine aminotransferase, *BMI* body mass index, *CK-18* cytokeratin-18 fragments, *HOMA-IR* homeostasis model assessment-insulin resistance, *HDL-C* high-density lipoprotein cholesterol, *LDL-C* low-density lipoprotein cholesterol, *NASH* nonalcoholic steatohepatitis, *NAS* NAFLD activity score, *TG* triglycerides, *TC* total cholesterol

[a]Significant fibrosis was defined as F2–F4 stages

[b]NASH was defined as the presence of hepatic steatosis, lobular inflammation and ballooning with histologic NAS ≥ 5

of demographics, metabolic risk factors and laboratory parameters.

Since the histologic definition of NASH can vary between the published studies, we also tested the diagnostic performance of our model in identifying either NASH (defined as a NAS score ≥ 3 with at least each of grade 1 steatosis, ballooning, and lobular inflammation) or active NASH (defined as a NAS score ≥ 4 with at least each of grade 1 steatosis, ballooning, and lobular inflammation). As shown in Supplementary Table 2, the discriminatory capability of the model performed well for diagnosing both NASH (NAS ≥ 3), and active NASH (NAS ≥ 4). Furthermore, in the validation set, the AUROCs of the model were 0.784 and 0.785 for diagnosing NASH (NAS ≥ 3) and active NASH (NAS ≥ 4), respectively.

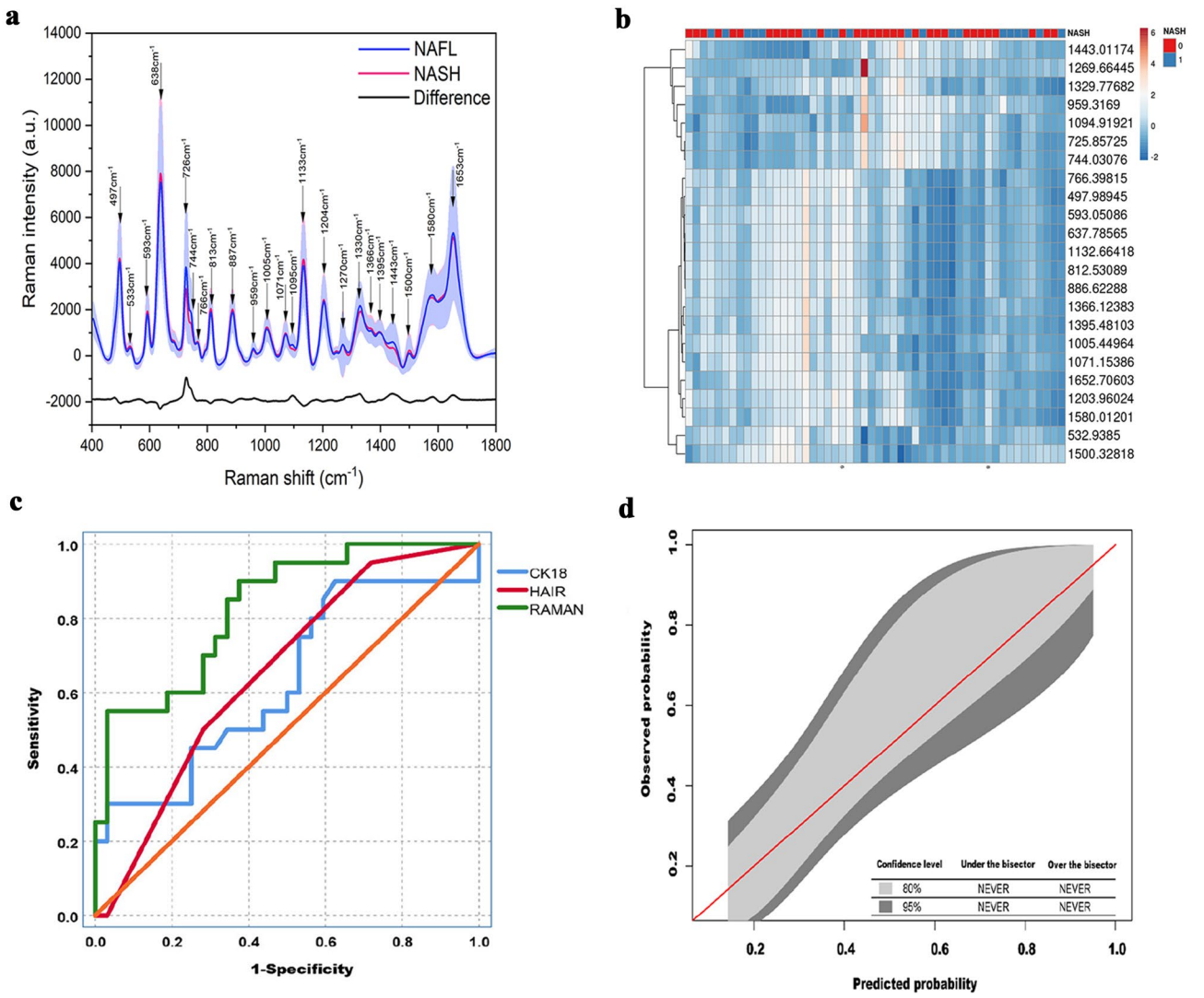

Fig. 2 Diagnostic performance of the SERS analysis for the non-invasive diagnosis of NASH. **a** comparison of the mean spectrum for NAFL blood serum samples (black line) versus NASH blood serum samples (red line). **b** heat map of the association between NASH and selected SERS bands. **c** AUROC of the validation cohort, and **d** calibration curve of the validation cohort

We also examined the diagnostic performance of our model according to different stages of liver fibrosis (Supplementary Table 3). In the validation cohort, the model identified significant fibrosis (stage F ≥ 2) and advanced fibrosis (stage F ≥ 3) with AUROCs of 0.72 and 0.96, respectively. Since only three of our patients had previously unknown cirrhosis (stage F4) in our cohort, we were unable to test the diagnostic performance of our model in identifying cirrhosis.

Six traditional machine learning algorithms, convolutional neural network (CNN), LightGBM, XGBoost, random forest, logistic regression, and gradient boosting were employed to train models to predict NASH based on SERS spectra. These methods classified NASH with AUROCs of 0.75 (CNN), 0.74 (LightGBM), 0.73 (XGBoost), 0.72 (random forest), 0.75 (logistic regression), and 0.74 (gradient boosting), respectively (Supplementary Table 4). The results suggest our machine learning model (FCNN) more accurately identified NASH than conventional approaches investigating spectroscopic data.

A user-friendly online calculator based on FCNN and SERS approaches is available freely, allowing health care providers and researches to estimate the individual patient's risk of NASH at http://www.pan-chess.cn/calculator/RAMAN_score.

Subgroup analysis

As shown in Table 2, we tested the diagnostic performance of the model in the validation cohort in various patient

Table 2 Diagnostic performance of the model for the non-invasive identification of NASH in different patient subgroups

	Percentage	AUROC (95% CI)	*p* value
Sex			
Male	63.5% (33/52)	0.88 (0.76–0.99)	0.326
Female	36.5% (19/52)	0.75 (0.53–0.98)	
Age			
< 50 years	55.8% (29/52)	0.89 (0.78–1.00)	0.216
≥ 50 years	44.2% (23/52)	0.74 (0.53–0.95)	
T2DM			
Yes	65.4% (34/52)	0.84 (0.70–0.98)	0.682
No	34.6% (18/52)	0.88 (0.72–1.00)	
Hypertension			
Yes	55.8% (29/52)	0.76 (0.58–0.94)	0.331
No	44.2% (23/52)	0.87 (0.72–1.00)	
BMI			
< 28 kg/m^2	69.2% (36/52)	0.84 (0.71–0.97)	0.781
≥ 28 kg/m^2	30.8% (16/52)	0.80 (0.57–1.00)	
ALT levels			
< 35 IU/L	51.9% (27/52)	0.70 (0.50–0.90)	0.261
≥ 35 IU/L	48.1% (25/52)	0.85 (0.68–1.00)	

ALT alanine aminotransferase, *AUROC* area under the receiver operating characteristics, *BMI* body mass index, *CI* confidence intervals, *MetS* metabolic syndrome, *NASH* nonalcoholic steatohepatitis, *T2DM* type 2 diabetes mellitus

subgroups. The validation cohort was stratified by age, sex, presence of T2DM, hypertension, obesity or elevated serum ALT levels. Notably, there were no significant differences in each of these patient subgroups (with AUROCs for the model ranging from nearly 0.70–0.88).

Discussion

In our exploratory study, a surface-enhanced Raman spectroscopy (SERS) method was developed for serum biochemical analysis with the specific aim of developing a simple blood test for the non-invasive identification of NASH. In our study, AgNPs as the SERS-active nanostructures were mixed with blood sera to enhance the Raman scattering signals, while suppressing fluorescence emissions at the same time. In addition, the Renishaw micro-Raman system can obtain high quality SERS spectrum from blood serum–Ag NP mixture within 10 s.[17] Notably, the results of our exploratory study suggest great potential for SERS in the non-invasive diagnosis of NASH.

The changing levels of a variety of serum biomarkers may reflect NAFLD severity. These biomarkers range from simple biochemical (serum aminotransferases, bilirubin and ferritin levels), metabolic (hemoglobin A1c, fasting insulin and HOMA-insulin resistance score) and lipid parameters (triglycerides and total cholesterol) to complex biomarkers reflecting specific molecular mechanisms underlying the pathogenesis and progression of NAFLD, including inflammation, oxidative stress, apoptosis, and glucose and lipid metabolism [31]. As the pathogenesis of NASH is complex and likely involves multiple biological abnormalities, it is unlikely that a single biomarker could accurately discriminate between NAFL and NASH. SERS typically features a number of peaks, and each peak corresponds to a specific molecular bond vibration, including individual bonds such as C–C, C=C, N–O, C–H etc. SERS can provide label-free fingerprinting type information on vibrational and rotational modes of different chemical structures, including metabolites, glucose, proteins, lipids, and nucleic acids [15]. Therefore, it is reasonable to assume that the distinctive SERS spectral features and intensity differences between NASH and NAFL patient groups can reflect molecular and cellular changes associated with the progression from NAFL to NASH.

Biochemical biomarkers and scoring systems derived from them generally have shown a relatively low sensitivity for the diagnosis and monitoring of NASH [32]. In addition, current serum biomarkers are usually measured by laboratory enzyme-linked immunosorbent assay (ELISA) methods, whereas the SERS-based analysis is faster than ELISA measurements. Moreover, the SERS-based analysis also has a simpler analysis protocol, and requires only a drop of serum [33]. Imaging techniques, such as ultrasonography, computed tomography, vibration-controlled transient elastography or magnetic resonance elastography, are a significant breakthrough for the non-invasive diagnosis and staging of liver steatosis and fibrosis, but they are still of limited value for the diagnosis of NASH. In addition, the relatively high costs, limited availability or exposure to ionizing radiations can limit the use of these imaging techniques for the screening and monitoring of patients with NASH. Finally, patients with NAFLD and severe obesity always have a long skin-to-liver capsule distance, which consistently limits the performance of liver ultrasonography [34].

Recently, a number of studies have tested the applicability and effectiveness of the SERS technique for the diagnosis of different diseases. For example, Yang et al. reported a label-free diagnostic platform that combines SERS and machine learning for the rapid and accurate detection of thirteen respiratory virus species, including SARS-CoV-2, common human coronaviruses, influenza viruses, and others [35]. Xie et al. reported an artificial intelligent SERS strategy for the diagnosis of breast cancer and assessment of its surgical outcomes [36]. In the specific field of NAFLD, Minamikawa et al. have accurately visualized the distribution of intrahepatic lipid droplets in a NASH mouse model, by applying Raman microscopy [37]. In such study, the Raman imaging analysis was

able to characterize NAFLD in terms of both molecular species and structures. Recently, Gurian et al. recruited 67 severely obese female, and found that there were significant differences in the band intensity of SERS spectra mainly related to hypoxanthine and uric acid between patients with NAFL and those with NASH [33]. Urasaki et al. found that hyperspectral stimulated Raman scattering microscopy could quantitatively measure liver composition of protein, DNA and lipid without labeling, and sensitively detect early-stage hepatic steatosis in a few minutes [38]. Szafraniec et al. characterized lipid droplets in liver sinusoidal endothelial cells, where their chemical composition was analyzed along the progression of NAFLD at the level of single lipid droplet using Raman imaging [39]. Yan et al. combined Raman micro-spectroscopy and machine learning techniques to develop a classification model based on a well-established NASH mouse model [40]. The results of such study showed that the classification model was capable of accurately detecting NASH in these mice (AUROC: 0.85). Collectively, our study further validated these results in humans. To our knowledge, this is the first study that has applied SERS analysis to diagnose NASH in a cohort of both male and female with biopsy-proven NAFLD. Our subgroup analyses showed that there were no significant differences in the diagnostic accuracy of SERS analysis between sexes.

The neural network is an important algorithm in machine learning. It connects multiple neurons into a network structure and mimics information analysis of biological nerve cells [41], and our methodology combined silver nanoparticle-based SERS spectroscopy with FCNN analysis to differentiate the blood serum of NASH patients from that of NAFL patients with high diagnostic accuracy. That said, we believe that the advent of the SERS analysis allows a non-invasive, standardized and rapid identification of patients with NASH at regular intervals. Moreover, our subgroup analyses showed that the diagnostic performance of the model performed well in different patient subgroups.

There are some limitations to our study that should be mentioned. Firstly, our patients with biopsy-confirmed NAFLD are from one country (China), and of a single ethnic group (Han people), and therefore these results are not necessarily applicable to other ethnic groups. Secondly, our sample size was relatively small. However, to our knowledge, this is the largest clinical study that has validated the use of the SERS analysis for the non-invasive diagnosis of NASH.

In conclusion, the results of our exploratory study suggest that the SERS blood serum analysis provides a non-invasive, rapid and practical way for accurately identifying the presence of NASH. Future studies in different ethnic cohorts of individuals with NAFLD are needed to further validate the use of the SERS analysis in clinical practice, and to test whether this methodology can be also used to non-invasively monitor therapeutic responses to emerging pharmacotherapies for NASH.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12072-022-10444-2.

Acknowledgements We thank Herui Biomed Company Limited (Suzhou, China) for providing CK-18 M30 ELISA kits.

Author contributions Study concept and design: FG, Z-FH, and M-HZ. Acquisition of data: D-CL, O-YH, L-JT, P-WZ, L-LC and Y-YL. Drafting of the manuscript: FG. Critical revision: GT and CDB. Statistical analysis: T-LZ, SG, and J-CS. Study supervision: M-HZ and Z-FH. All authors contributed to the manuscript for important intellectual content and approved the final version of the manuscript.

Funding This work was supported by grants from the National Natural Science Foundation of China (82070588), High Level Creative Talents from Department of Public Health in Zhejiang Province (S2032102600032) and Project of New Century 551 Talent Nurturing in Wenzhou. GT is supported in part by grants from the School of Medicine, University of Verona, Verona, Italy. CDB is supported in part by the Southampton NIHR Biomedical Research Centre (IS-BRC-20004).

Availability of data and material All data included in this study are available upon request by contact with the corresponding author.

Declarations

Conflict of interest Feng Gao, De-Chan Lu, Tian-Lei Zheng, Shi Geng, Jun-Cheng Sha, Ou-Yang Huang, Liang-Jie Tang, Pei-Wu Zhu, Yang-Yang Li, Li-Li Chen, Giovanni Targher, Christopher D. Byrne, Zu-Fang Huang, Ming-Hua Zheng: nothing to declare.

Ethical approval Ethical approval for the study was obtained from the ethics committee of the First Affiliated Hospital of Wenzhou Medical University. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments.

Informed consent Written informed consent was obtained from all individual participants included in the study.

References

1. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology (Baltimore, MD). 2016;64(1):73–84
2. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol. 1999;94(9):2467–2474
3. Younossi ZM, Stepanova M, Ong J, et al. Nonalcoholic steatohepatitis is the most rapidly increasing indication for liver transplantation in the United States. Clin Gastroenterol Hepatol. 2021;19(3):580–589.e5
4. Chen Z, Liu J, Zhou F, et al. Nonalcoholic fatty liver disease: an emerging driver of cardiac arrhythmia. Circ Res. 2021;128(11):1747–1765

5. Byrne CD, Targher G. NAFLD: a multisystem disease. J Hepatol. 2015;62(1 Suppl):S47-64
6. Sun DQ, Jin Y, Wang TY, et al. MAFLD and risk of CKD. Metabolism. 2021;115:154433
7. Wang TY, Wang RF, Bu ZY, et al. Association of metabolic dysfunction-associated fatty liver disease with kidney disease. Nat Rev Nephrol. 2022;18(4):259–268
8. Sheka AC, Adeyi O, Thompson J, Hameed B, Crawford PA, Ikramuddin S. Nonalcoholic steatohepatitis: a review. JAMA. 2020;323(12):1175–1183
9. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD. Liver biopsy. Hepatology (Baltimore, MD). 2009;49(3):1017–1044
10. Kwok R, Tse YK, Wong GL, et al. Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease—the role of transient elastography and plasma cytokeratin-18 fragments. Aliment Pharmacol Ther. 2014;39(3):254–269
11. Cusi K, Chang Z, Harrison S, et al. Limited value of plasma cytokeratin-18 as a biomarker for NASH and fibrosis in patients with non-alcoholic fatty liver disease. J Hepatol. 2014;60(1):167–174
12. Dorairaj V, Sulaiman SA, Abu N, Abdul Murad NA. Nonalcoholic fatty liver disease (NAFLD): pathogenesis and noninvasive diagnosis. Biomedicines. 2021;10(1):15
13. Zhou YJ, Wong VW, Zheng MH. Consensus scoring systems for nonalcoholic fatty liver disease: an unmet clinical need. Hepatobiliary Surg Nutr. 2021;10(3):388–390
14. Zong C, Xu M, Xu LJ, et al. Surface-enhanced Raman spectroscopy for bioanalysis: reliability and challenges. Chem Rev. 2018;118(10):4946–4980
15. Devitt G, Howard K, Mudher A, Mahajan S. Raman spectroscopy: an emerging tool in neurodegenerative disease research and diagnosis. ACS Chem Neurosci. 2018;9(3):404–420
16. Fan M, Andrade GF, Brolo AG. A review on the fabrication of substrates for surface enhanced Raman spectroscopy and their applications in analytical chemistry. Anal Chim Acta. 2011;693(1–2):7–25
17. Feng S, Chen R, Lin J, et al. Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis. Biosens Bioelectron. 2010;25(11):2414–2419
18. Shin H, Oh S, Hong S, et al. Early-stage lung cancer diagnosis by deep learning-based spectroscopic analysis of circulating exosomes. ACS Nano. 2020;14(5):5435–5444
19. Huang Z, Siddhanta S, Zheng G, Kickler T, Barman I. Rapid, label-free optical spectroscopy platform for diagnosis of heparin-induced thrombocytopenia. Angew Chem (Int Ed Engl). 2020;59(15):5972–5978
20. Guan H, Huang C, Lu D, et al. Label-free Raman spectroscopy: a potential tool for early diagnosis of diabetic keratopathy. Spectrochim Acta Part A Mol Biomol Spectrosc. 2021;256:119731
21. Rashidi HH, Tran NK, Betts EV, Howell LP, Green R. Artificial intelligence and machine learning in pathology: the present landscape of supervised methods. Acad Pathol. 2019;6:2374289519873088
22. Zhao D, Chen M, Shi K, Ma M, Huang Y, Shen J. A long short-term memory-fully connected (LSTM-FC) neural network for predicting the incidence of bronchopneumonia in children. Environ Sci Pollut Res Int. 2021;28(40):56892–56905
23. Rehman KU, Li J, Pei Y, Yasin A, Ali S, Mahmood T. Computer vision-based microcalcification detection in digital mammograms using fully connected depthwise separable convolutional neural network. Sensors (Basel, Switzerland). 2021;21(14):4854
24. Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation. 2009;120(16):1640–1645
25. Gao F, Huang JF, Zheng KI, et al. Development and validation of a novel non-invasive test for diagnosing fibrotic non-alcoholic steatohepatitis in patients with biopsy-proven non-alcoholic fatty liver disease. J Gastroenterol Hepatol. 2020;35(10):1804–1812
26. Dixon JB, Bhathal PS, O'Brien PE. Nonalcoholic fatty liver disease: predictors of nonalcoholic steatohepatitis and liver fibrosis in the severely obese. Gastroenterology. 2001;121(1):91–100
27. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology (Baltimore, MD). 2005;41(6):1313–1321
28. Brunt EM, Kleiner DE, Wilson LA, Belt P, Neuschwander-Tetri BA. Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. Hepatology (Baltimore, MD). 2011;53(3):810–820
29. Koo BK, Kim D, Joo SK, et al. Sarcopenia is an independent risk factor for non-alcoholic steatohepatitis and significant fibrosis. J Hepatol. 2017;66(1):123–131
30. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Boston, USA; 2015. pp. 3431–3430
31. Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: Clinical prediction rules and blood-based biomarkers. J Hepatol. 2018;68(2):305–315
32. Pirmoazen AM, Khurana A, El Kaffas A, Kamaya A. Quantitative ultrasound approaches for diagnosis and monitoring hepatic steatosis in nonalcoholic fatty liver disease. Theranostics. 2020;10(9):4277–4289
33. Gurian E, Giraudi P, Rosso N, et al. Differentiation between stages of non-alcoholic fatty liver diseases using surface-enhanced Raman spectroscopy. Anal Chim Acta. 2020;1110:190–198
34. Gao F, He Q, Li G, et al. A novel quantitative ultrasound technique for identifying non-alcoholic steatohepatitis. Liver Int. 2022;42(1):80–91
35. Yang Y, Xu B, Murray J, et al. Rapid and quantitative detection of respiratory viruses using surface-enhanced Raman spectroscopy and machine learning. Biosens Bioelectron. 2022;217: 114721
36. Xie Y, Su X, Wen Y, Zheng C, Li M. Artificial intelligent label-free SERS profiling of serum exosomes for breast cancer diagnosis and postoperative assessment. Nano Lett. 2022;22:7910–7918
37. Minamikawa T, Ichimura-Shimizu M, Takanari H, et al. Molecular imaging analysis of microvesicular and macrovesicular lipid droplets in non-alcoholic fatty liver disease by Raman microscopy. Sci Rep. 2020;10(1):18548
38. Urasaki Y, Zhang C, Cheng JX, Le TT. Quantitative Assessment of Liver Steatosis and Affected Pathways with Molecular Imaging and Proteomic Profiling. Sci Rep. 2018;8(1):3606
39. Szafraniec E, Kus E, Wislocka A, et al. Raman spectroscopy-based insight into lipid droplets presence and contents in liver sinusoidal endothelial cells and hepatocytes. J Biophotonics. 2019;12(4): e201800290
40. Yan J, Yu Y, Kang JW, et al. Development of a classification model for non-alcoholic steatohepatitis (NASH) using confocal Raman micro-spectroscopy. J Biophotonics. 2017;10(12):1703–1713
41. Kriegeskorte N, Golan T. Neural network models and deep learning. Curr Biol CB. 2019;29(7):R231-r236

Authors and Affiliations

Feng Gao[1] · De-Chan Lu[2] · Tian-Lei Zheng[3,4] · Shi Geng[3] · Jun-Cheng Sha[5] · Ou-Yang Huang[6] · Liang-Jie Tang[6] · Pei-Wu Zhu[7] · Yang-Yang Li[8] · Li-Li Chen[6] · Giovanni Targher[9] · Christopher D. Byrne[10] · Zu-Fang Huang[2] · Ming-Hua Zheng[6,11,12] on behalf of CHESS-MAFLD consortium

1 Department of Gastroenterology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

2 Key Laboratory of Opto-Electronic Science and Technology for Medicine of Ministry of Education, Fujian Provincial Key Laboratory for Photonics Technology, Fujian Normal University, Fuzhou 350000, China

3 Artificial Intelligence Unit, Department of Medical Equipment Management, Affiliated Hospital of Xuzhou Medical University, Xuzhou, China

4 School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China

5 Interventional Radiology, Affiliated Hospital of Xuzhou Medical University, Xuzhou, China

6 NAFLD Research Center, Department of Hepatology, The First Affiliated Hospital of Wenzhou Medical University, No. 2 Fuxue Lane, Wenzhou 325000, China

7 Department of Laboratory Medicine, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

8 Department of Pathology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

9 Section of Endocrinology, Diabetes and Metabolism, Department of Medicine, University and Azienda Ospedaliera Universitaria Integrata of Verona, Verona, Italy

10 Southampton National Institute for Health Research Biomedical Research Centre, University Hospital Southampton, Southampton General Hospital, Southampton, UK

11 Key Laboratory of Diagnosis and Treatment for the Development of Chronic Liver Disease in Zhejiang Province, Wenzhou, China

12 Institute of Hepatology, Wenzhou Medical University, Wenzhou, China