Original Article

LEARN algorithm: a novel option for predicting non-alcoholic steatohepatitis

Gang Li[1#], Tian-Lei Zheng[2#], Xiao-Ling Chi[3], Yong-Fen Zhu[4], Jin-Jun Chen[5,6], Liang Xu[7], Jun-Ping Shi[8], Xiao-Dong Wang[9], Wei-Guo Zhao[2], Christopher D. Byrne[10], Giovanni Targher[11], Rafael S. Rios[1], Ou-Yang Huang[1], Liang-Jie Tang[1], Shi-Jin Zhang[2], Shi Geng[2], Huan-Ming Xiao[3], Sui-Dan Chen[12], Rui Zhang[13], Ming-Hua Zheng[1,9,14]

[1]MAFLD Research Center, Department of Hepatology, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China; [2]Artificial Intelligence Unit, Department of Medical Equipment, The Affiliated Hospital of Xuzhou Medical University, Xuzhou, China; [3]Department of Hepatology, Guangdong Provincial Hospital of Chinese Medicine, the Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China; [4]Department of Hepatology and Infection, Sir Run Run Shaw Hospital, Affiliated with School of Medicine, Zhejiang University, Hangzhou, China; [5]Hepatology Unit, Department of Infectious Diseases, Nanfang Hospital, Southern Medical University, Guangzhou, China; [6]Hepatology Unit, Zengcheng Branch, Nanfang Hospital, Southern Medical University, Guangzhou, China; [7]Department of Hepatology, Tianjin Second People's Hospital, Tianjin, China; [8]Department of Liver Diseases, Hangzhou Normal University Affiliated Hospital, Hangzhou, China; [9]Key Laboratory of Diagnosis and Treatment for the Development of Chronic Liver Disease in Zhejiang Province, Wenzhou, China; [10]Southampton National Institute for Health and Care Research Biomedical Research Centre, University Hospital Southampton & University of Southampton, Southampton General Hospital, Southampton, UK; [11]Section of Endocrinology, Diabetes and Metabolism, Department of Medicine, University of Verona, Verona, Italy; [12]Department of Pathology, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China; [13]Department of Nutrition, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China; [14]Institute of Hepatology, Wenzhou Medical University, Wenzhou, China

Contributions: (I) Conception and design: G Li, MH Zheng; (II) Administrative support: MH Zheng; (III) Provision of Study Material: MH Zheng; (IV) Collection and assembly of data: G Li, XL Chi, YF Zhu, JJ Chen, L Xu, JP Shi, XD Wang, WG Zhao, OY Huang, LJ Tang, HM Xiao, R Zhang; (V) Data analysis and interpretation: G Li, TL Zheng, WG Zhao, SJ Zhang, S Geng; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work and should be considered as co-first authors.

Correspondence to: Ming-Hua Zheng, MD, PhD. MAFLD Research Center, Department of Hepatology, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China. Email: zhengmh@wmu.edu.cn.

Background: There is an unmet need for accurate non-invasive methods to diagnose non-alcoholic steatohepatitis (NASH). Since impedance-based measurements of body composition are simple, repeatable and have a strong association with non-alcoholic fatty liver disease (NAFLD) severity, we aimed to develop a novel and fully automatic machine learning algorithm, consisting of a deep neural network based on impedance-based measurements of body composition to identify NASH [the bioeLectrical impEdance Analysis foR Nash (LEARN) algorithm].

Methods: A total of 1,259 consecutive subjects with suspected NAFLD were screened from six medical centers across China, of which 766 patients with biopsy-proven NAFLD were included in final analysis. These patients were randomly subdivided into the training and validation groups, in a ratio of 4:1. The LEARN algorithm was developed in the training group to identify NASH, and subsequently, tested in the validation group.

Results: The LEARN algorithm utilizing impedance-based measurements of body composition along with age, sex, pre-existing hypertension and diabetes, was able to predict the likelihood of having NASH. This algorithm showed good discriminatory ability for identifying NASH in both the training and validation groups [area under the receiver operating characteristics (AUROC): 0.81, 95% CI: 0.77–0.84 and AUROC: 0.80, 95% CI: 0.73–0.87, respectively]. This algorithm also performed better than serum cytokeratin-18 neoepitope M30 (CK-18 M30) level or other non-invasive NASH scores (including HAIR, ION, NICE) for identifying NASH (P value <0.001). Additionally, the LEARN algorithm performed well in identifying

NASH in different patient subgroups, as well as in subjects with partial missing body composition data.

Conclusions: The LEARN algorithm, utilizing simple easily obtained measures, provides a fully automated, simple, non-invasive method for identifying NASH.

Keywords: Non-alcoholic fatty liver disease (NAFLD); non-alcoholic steatohepatitis (NASH); bioeLectrical impEdance Analysis foR Nash (LEARN) algorithm; body composition

Submitted Dec 12, 2021. Accepted for publication May 09, 2022. Published online Mar 30, 2023.
doi: 10.21037/hbsn-21-523
View this article at: https://dx.doi.org/10.21037/hbsn-21-523

Introduction

Non-alcoholic steatohepatitis (NASH) is a major public health concern worldwide and, compared with hepatic steatosis alone, the annual incidence of hepatocellular carcinoma in patients with NASH-related cirrhosis is as high as 1–2% (1,2). NASH is more likely to lead to advanced liver fibrosis, cirrhosis and eventually liver-related illness and death (3-5). Therefore, due to its high prevalence and increased health risks, NASH is a significant economic and healthcare burden. The current definitive diagnosis of NASH is based not only on hepatocyte fat accumulation (steatosis), but also on histological evidence of hepatocyte ballooning and lobular inflammation (6). Given that the majority of patients with NASH are asymptomatic, the acceptability of liver biopsy (i.e., the gold standard) is relatively low and, because of liver biopsy-associated morbidity and even mortality, developing screening strategies to identify those individuals at risk of progressive NASH, remains an unmet need. Furthermore, non-invasive tests that may accurately predict disease progression (as part of the natural history of NASH), or identify regression (in response to treatment), are urgently needed to decrease the reliance on repeat liver biopsies (7-9).

Machine learning techniques require uploading a large amount of data to a computer program, and then selecting a model to "fit" these data for computer prediction, which creates new possibilities in medicine for diagnosing diseases (10-12). In previous studies, machine learning has facilitated success in cancer diagnosis and diagnosis of liver fibrosis (13,14). Recently, in Sagimet's NASH FASCINATE-2 Phase 2b Clinical Trial, stain-free artificial intelligence (AI)-based digital pathology was incorporated as secondary and exploratory efficacy endpoints. These advances would have been unimaginable without machine learning. To date, however, there is no a validated, non-invasive, simple, machine learning-based algorithm (MLA) for diagnosing NASH.

Bioelectrical impedance analysis (BIA) is a simple, commonly used, non-invasive and inexpensive method for assessing body composition (15). This method can provide >20 parameters on different dimensions of body composition, such as body fat content, muscle mass, bone mineral content and metabolic rate. Interestingly, there is evidence that body composition in non-alcoholic fatty liver disease (NAFLD) is different from that of non-steatotic control subjects (16-19). However, the abundant body composition outputs from BIA have not yet been fully evaluated and exploited in the diagnosis and treatment of NAFLD.

Therefore, the main aim of our multicenter cross-sectional study was to establish and validate a novel MLA, referred to as a deep neural network algorithm for non-invasively identifying NASH by impedance-based measures of body composition [named as the LEARN (bioeLectrical impEdance Analysis foR Nash) algorithm]. We present this article in accordance with the TRIPOD reporting checklist (available at https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-523/rc).

Methods

Ethical statement

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study has been approved by the Institutional Review Board of the First Affiliated Hospital of Wenzhou Medical University (2016-246). Informed consent was taken from all individual participants.

Study subjects and design

A total of 1,259 consecutive subjects with suspected

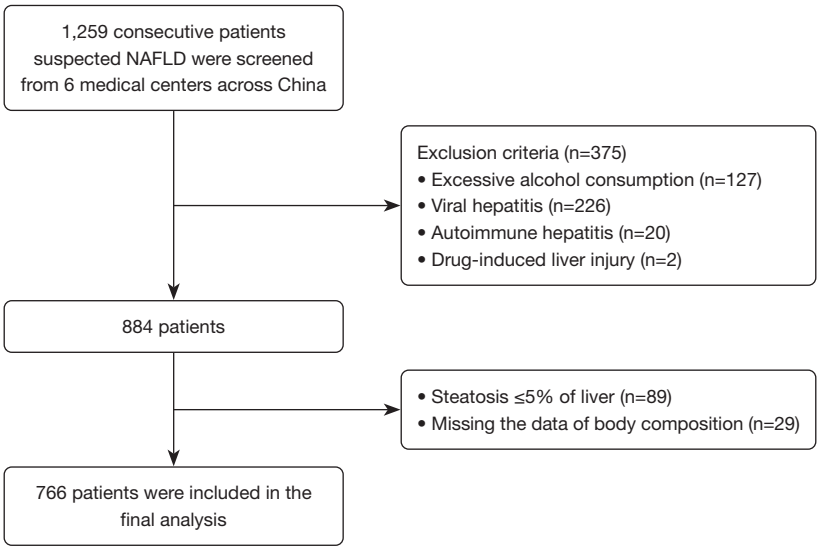

Figure 1 Flowchart of the study. NAFLD, non-alcoholic fatty liver disease.

NAFLD were initially screened from six medical centers across China from September 2016 to April 2021. Inclusion criteria were as follows: (I) elevated serum aminotransferase concentrations and/or evidence of hepatic steatosis on imaging methods (irrespective of serum aminotransferase levels); (II) agreement to undergo a liver biopsy; (III) agreement to undergo BIA method within 1 month of liver biopsy; and (IV) age range from 18 to 75 years. A total of 493 subjects were excluded due to the following criteria: (I) excessive alcohol consumption (more than 20 and 10 grams per day for men and women, respectively); (II) other coexisting chronic liver diseases, such as viral hepatitis, autoimmune hepatitis, or drug-induced liver injury; (III) absence of hepatic steatosis on histology (steatotic hepatocytes ≤5%); and (IV) no BIA measurement. As a consequence of these exclusion criteria, 766 Chinese adults with biopsy-confirmed NAFLD were included in the final analysis (*Figure 1*).

Clinical and laboratory data

For every patient, demographic data, anthropometry, clinical biochemical parameters and concomitant diseases were measured and collected during liver biopsy, at each center, within 48 hours from liver biopsy. Hypertension was defined as blood pressure ≥130/85 mmHg or current use any of anti-hypertensive drugs. Presence of diabetes was defined as self-reported physician diagnosis of diabetes, use of anti-hyperglycemic drugs, fasting glucose levels ≥7 mmol/L or hemoglobin A1c (HbA1c) ≥6.5% (≥48 mmol/moL). Homeostasis model assessment (HOMA-IR) was used to estimate insulin resistance, and body mass index (BMI) ≥25 kg/m^2 was diagnosed as overweight/obese. The specific methods for assessing HOMA-IR and BMI have been described in our previous study (20).

Measurement of cytokeratin-18 neoepitope M30 (CK-18 M30)

Serum CK-18 M30 level was measured only in the Wenzhou cohort. Serum CK-18 M30 level was determined by a commercially ELISA kit (Herui Biomed Company Limited, Suzhou, China), according to the manufacturer's recommendation. The specific detection details have been described in our previous study (21).

Body composition measurement

Each patient was examined for body composition (within 48 hours of the liver biopsy) by professionally trained personnel at each center in accordance with uniform operating instructions. Specifically, BIA (InBody 720; Biospace, Seoul, Korea) was employed to measure body composition. According to operating instructions, the subjects took off their shoes and removed their belongings and coats and stood on the designated electrodes. The thumb of both hands was placed on the thumb electrode button, and the other four fingers were all placed on the electrode, under

the handle, with arms straightened. The impedance of left arm, right arm, trunk, right leg and left leg were measured at six frequencies (1, 5, 50, 250, 500, and 1,000 kHz). Based on the aforementioned impedances, the system automatically produced information of body composition, which included 20 parameters, such as intracellular water, extracellular water, total body water, soft lean mass, fat free mass, weight, skeletal muscle mass, body fat mass, percent body fat, waist-hip ratio, right arm, left arm, trunk, as well as right leg, left leg, visceral fat area, body cell mass, bone mineral content, basal metabolic rate, arm circumference and arm muscle circumference.

Liver biopsy

Liver biopsies were performed using a 16-gauge needle under ultrasound guidance as previous described (21). All liver biopsy specimens were interpreted by an experienced pathologist from each center. Diagnostic criteria for NAFLD were the evidence of steatotic hepatocytes >5% on histology. NASH was diagnosed only when the NAFLD activity score (NAS) was ≥4 and each component of its three histological features (i.e., steatosis, ballooning and lobular inflammation) was ≥1 (22). Liver fibrosis stage was graded from 0 to 4, according to the Brunt's histologic criteria.

Development of the LEARN algorithm

An experienced AI team used neural network algorithms to build a prediction model that provided clinicians with an individual's probability of having NASH. As shown in the *Figure 2*, the data from our 766 patients with biopsy-proven NAFLD were first subdivided randomly into a training set (613 patients) and a validation set (153 patients), in a 4:1 ratio. The data included in the LEARN algorithm included age, sex, diabetes and hypertension status, as well as 20 body composition parameters obtained by BIA. The MLA process is currently the subject of a patent application. In particular, we normalized processing, and inputted these data to the input layer composed of the full connection network. In this layer, we further analyzed the 20 body composition parameters. Each parameter was automatically assigned to a different weight in the neural network model, and the best choices of the first six body composition information parameters (i.e., arm circumference, body fat percentage, bone mineral content, basal metabolic rate, body cell mass and visceral fat area) for prediction of NASH were selected after the method of exhaustion (as shown in *Figure 2*). The six body composition parameters along with age, sex, and prior history of diabetes or hypertension were re-sent into the input layer, to extract the feature matrix. The feature matrix was then inputted into the residual network layer, composed of four residual modules, which can also be called the hidden layer. 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 Tanh activation function, and the last module removes the activation function compared with the previous two modules. A single fully connected residual module may be expressed as:

$$x_m = F\left(x, \{W_1, W_2, W_3\}\right) \quad [1]$$

$$y = Dropout\left(Tanh\left(x_m + x\right)\right) \quad [2]$$

where x is the input feature, y is the output feature, W indicates the weight of the fully connected module and F indicates the combination of three fully connected modules. For the feature x of the input fully connected residual module, the module firstly uses three fully connected modules to extract feature successively to generate intermediate feature x_m. Then add x as residuals to x_m, and use the Dropout function to generate y after Tanh activation. Four fully connected residual modules can increase the depth of the model while suppressing the disappearance of gradient, thus improving the performance of the model. Finally, the extracted features are inputted into the output layer, composed of the fully connected network and softmax activation function, to calculate the probability of having NASH.

Other widely used non-invasive NASH scores

As liver biopsy can be fraught with major acute complications, there are some widely used non-invasive scores for diagnosing NASH, such as ION, HAIR, NICE and model, which are based on combinations of laboratory indicators and metabolic factors (23-25). In particular, these three non-invasive NASH scores can be calculated as follows:

The index of NASH (*ION*) = 1.33 waist-to hip ratio + 0.03 × triglycerides (mg/dL) + 0.18 × alanine aminotransferase (ALT) (U/L) + 8.53 × HOMA-IR – 13.93 in men; 0.02 × triglycerides (mg/dL) + 0.24 × ALT (U/L) + 9.61 × HOMA-IR – 13.99 in women.

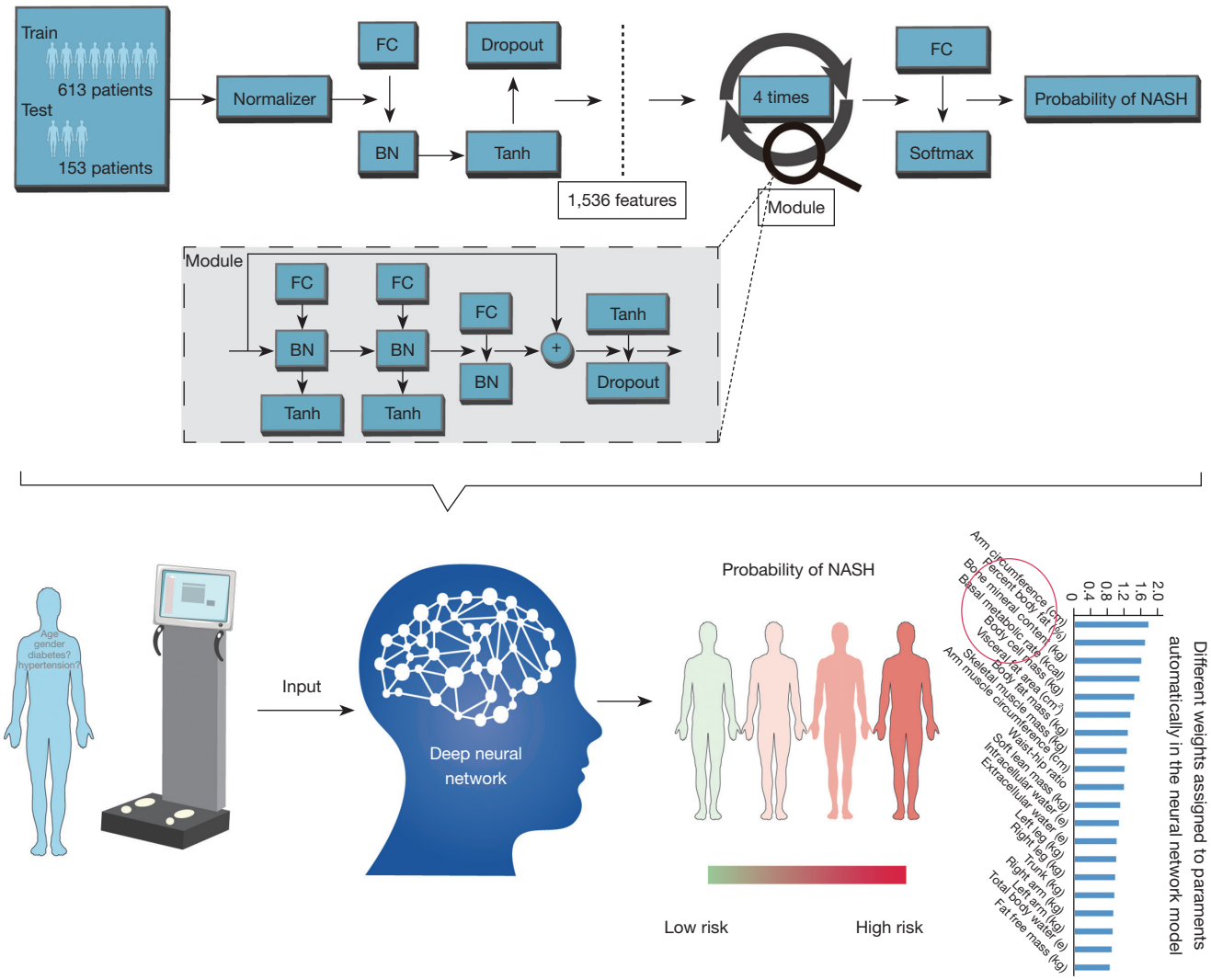

Figure 2 Flowchart of the deep neural network algorithm for prediction of NASH (namely the LEARN algorithm). Input layer: input the normalized data into this layer consisting of four modules. First, a FC layer, to synthesize the features extracted from the previous section. Second, a BN layer, to simplify the calculation and make the data retain its original expression ability as far as possible after the normalization processing. Third, Tanh function, a nonlinear function to help machines learn complex mappings. Last, the Dropout layer, reducing overfitting, extract a matrix containing 1,536 features. Hidden layer: the matrix containing 1,536 features is input into this layer, and the data needs to be looped four times through the residual module. Output layer: the output layer is consisted of a FC layer and a Softmax function. Through the Softmax function, we can map the output values to the interval (0, 1) for the final classification. FC, full connected; BN, batch normalization; NASH, non-alcoholic steatohepatitis; LEARN, bioeLectrical impEdance Analysis foR Nash.

The NICE model = −5.654 + 3.780E-02 × ALT (IU/L) + 2.215E-03 × CK18 fragment (IU/L) + 1.825 × (presence of metabolic syndrome =1).

The HAIR score was calculated by adding hypertension =1, ALT >40 U/L =1, and HOMA-IR index >5.0 =1 for each patient (0–3).

Statistical analysis

Continuous and categorical data were expressed as means ± standard deviations, and medians (1st quartile, 3rd quartile), or proportions, respectively. For the purpose of determining statistical differences between the training

and the validation groups, the unpaired Student's *t*-test (for normally distributed continuous data), the Mann-Whitney U-test (for non-normally distributed continuous data) and the chi-square test (for categorical variables) were used. PASS15 was used to estimate the sample size. The area under the receiver operator characteristic (AUROC) curve was 0.80, α =0.05 (bilateral), β =0.1 (test efficiency was 0.9), and the ratio between groups was 3:2. It was found that a minimum of 42 subjects, including 25 patients and 17 controls, needed to be enrolled. A sample size of 776 is completely sufficient. The process of establishing the deep neural network algorithm (the LEARN algorithm) was summarized in *Figure 2*. The AUROC curve was calculated to evaluate the discrimination of the machine learning intelligent diagnostic model. Cut-off values for the diagnosis of NASH were identified in the training group, corresponding to 90% sensitivity and 90% specificity, respectively. At the same time, the specificity, sensitivity, negative predictive value (NPV), positive predictive value (PPV) and the gray zone were also calculated corresponding to each cut-off value in the training and validation groups. Statistical significance was two sided, set at P value less than 0.05. All statistical tests were performed by R (http://www.r-project.org) and SPSS version 22.0 (SPSS Inc.).

Results

Patient characteristics

A total of 766 Chinese adult patients with biopsy-confirmed NAFLD from six hospitals were randomly assigned to the training (n=613) and validation groups (n=153) in a ratio of 4:1. As shown in *Table 1*, there were no statistical differences in demographic and biochemical measurements, body composition data, as well as other widely used non-invasive NASH scores (ION, HAIR, NICE model) and individual features of liver histology between the two groups. It is worth noting that a complete body composition examination report included 20 parameters, involving also skeletal muscle mass and abdominal fat area. However, due to the loss of data during the collection process, some patient's body composition examination reports were incomplete and reported between 12 and 19 measurements, which we referred to as subjects with partial missing data on body composition (PMBC). The baseline clinical, biochemical and BIA characteristics of patients stratified by those without missing BIA data (WMBC) (n=690) and those with PMBC (n=76) are summarized in Table S1. Table S2 shows the baseline characteristics of patients with biopsy-proven NAFLD, stratified by NASH or non-NASH, both in the training and validation groups. Compared with those with non-NASH, patients with NASH differed in terms of age, BMI, percent body fat, visceral fat area, fasting glucose, plasma lipid profile and serum transaminases. Notably, LEARN algorithm, ION, HAIR, and NICE model, as well as histological stages of fibrosis also significantly differed between NASH or non-NASH patients, both in the training or validation groups.

LEARN algorithm development

As shown in *Figure 2*, we have developed a novel and automatic machine learning model, called the LEARN algorithm, which included age, sex, prior hypertension, prior diabetes, as well as six body composition parameters (namely arm circumference, percent body fat, bone mineral content, basal metabolic rate, body cell mass and visceral fat area). Biochemical parameters were not included in the LEARN algorithm. The individual probability of having NASH was calculated via the LEARN algorithm. For example, for a patient who has undergone BIA-based measurements of body composition and who provides data on age, sex, and prior history of hypertension or diabetes, it is possible to calculate her/his probability of having NASH.

Diagnostic performance of LEARN algorithm in the training and validation groups (Figure 3)

The AUROC for LEARN algorithm in the training and validation groups were 0.81 (95% CI: 0.77–0.84) (*Figure 3A*), and 0.80 (95% CI: 0.73–0.87) (*Figure 3C*), respectively. In both patient groups, the LEARN algorithm performed well for diagnosing NASH. To more accurately identify NASH, we chose 0.492 (sensitivity =0.90) and 0.531 (specificity =0.91), as dual cut-off values in the training group (*Figure 3B*). As shown in *Table 2*, when we chose these two cut-off values obtained by the LEARN algorithm, there was a NPV of 0.70 to rule out NASH and a PPV of 0.93 to rule in NASH in the training group, respectively. Similarly, in *Figure 3D* and *Table 2*, when we used the same dual cut-off values in the validation group, the cut-off values of 0.492 (sensitivity =0.91) and 0.531 (specificity =0.89) gave a NPV of 0.71 to rule out NASH and a PPV of 0.90 to rule in NASH in the training group, respectively. The diagnostic efficiency in the validation group also showed the same level of discrimination as that in the training group. Also, *Figure 4* shows the boxplots of the LEARN algorithm *vs.*

Table 1 Baseline characteristics of patients with biopsy-proven NAFLD

Characteristics	Training group (n=613)	Validation group (n=153)	P value
Demographics			
Age (years)	41.4±12.6	39.9±12.1	0.213
Male sex, n (%)	440 (71.8)	108 (70.6)	0.770
BMI (kg/m^2)	26.9 (24.6, 29.3)	26.8 (24.2, 29.1)	0.618
Type 2 diabetes, n (%)	147 (24.0)	43 (28.1)	0.291
Hypertension, n (%)	137 (22.3)	37 (24.2)	0.628
Biochemical measurements			
Albumin (g/L)	46.1±3.9	46.1±4.0	0.999
Platelet count (×10^9)	243.3±62.0	246.3±62.8	0.595
AST (U/L)	34.0 (25.0, 52.0)	35.0 (24.0, 56.0)	0.614
ALT (U/L)	51.0 (32.0, 88.0)	49.0 (30.0, 100.0)	0.844
ALP (U/L)	82.0 (68.0, 99.0)	79.0 (63.0, 97.0)	0.222
Glucose (mmol/L)	5.3 (4.9, 6.2)	5.3 (4.9, 6.5)	0.485
Total cholesterol (mmol/L)	5.1 (4.3, 5.9)	5.2 (4.5, 5.9)	0.461
Triglycerides (mmol/L)	1.9 (1.4, 2.9)	1.7 (1.3, 2.7)	0.249
HDL-C (mmol/L)	1.0±0.3	1.0±0.2	0.610
LDL-C (mmol/L)	3.0±0.9	3.1±1.0	0.406
CK-18 M30 (U/L)	155.5 (72.8, 337.2) [428]	162.0 (80.0, 414.7) [109]	0.362
Body composition			
Body composition analysis			
Intracellular water (e)	24.6±4.7 [611]	24.5±4.4	0.815
Extracellular water (e)	15.0±2.6 [611]	14.8±2.5	0.423
Total body water (e)	39.7±7.3 [611]	39.3±6.8	0.590
Soft lean mass (g)	50.9±9.8 [611]	50.5±9.0	0.595
Fat free mass (kg)	54.2±10.3 [611]	53.6±9.4	0.519
Muscle-fat analysis			
Weight (kg)	76.9±14.5	75.4±12.9	0.254
Skeletal muscle mass (kg)	30.4±6.5 [588]	29.9±5.8 [149]	0.433
Body fat mass (kg)	22.6±7.7	21.8±7.2	0.235
Obesity diagnosis			
Percent body fat (%)	29.1±6.7	28.7±7.0	0.439
Waist-hip ratio	0.93±0.05 [588]	0.92±0.06 [149]	0.644

Table 1 (*continued*)

Table 1 (*continued*)

Characteristics	Training group (n=613)	Validation group (n=153)	P value
Lean balance			
Right arm (kg)	3.1±0.7	3.1±0.7	0.636
Left arm (kg)	3.0±0.7	3.0±0.7	0.701
Trunk (kg)	24.6±4.5	24.4±4.3	0.608
Right leg (kg)	8.3±1.9	8.1±1.6	0.366
Left leg (kg)	8.2±1.9	8.1±1.6	0.386
Visceral fat area			
Visceral fat area (cm^2)	102.3±27.9 [583]	101.4±27.9 [149]	0.705
Additional data			
Body cell mass (kg)	35.4±6.6 [558]	35.2±6.4 [143]	0.744
Bone mineral content (kg)	3.0±0.6 [583]	3.0±0.5 [147]	0.638
Basal metabolic rate (kcal)	1,541.5±249.3	1,527.7±213.4	0.527
Arm circumference (cm)	33.5±3.2 [557]	33.3±3.0 [142]	0.525
Arm muscle circumference (cm)	27.3±2.5 [557]	27.2±2.6 [142]	0.600
Non-invasive NASH scores			
ION	37.7 (22.8, 63.4) [521]	38.6 (20.9, 72.2) [130]	0.833
HAIR	1.0 (1.0, 2.0) [586]	1.0 (1.0, 2.0) [145]	0.545
NICE	–2.6 (–3.9, –0.8) [428]	–2.4 (–4.1, 0.1) [109]	0.653
Liver histology, n (%)			
Steatosis			0.354
1	236 (38.5)	64 (41.8)	
2	259 (42.3)	55 (35.9)	
3	118 (19.2)	34 (22.2)	
Hepatocyte ballooning			0.131
0	63 (10.3)	23 (15.0)	
1	339 (55.3)	73 (47.7)	
2	211 (34.4)	57 (37.3)	
Lobular inflammation			0.716
0	53 (8.6)	16 (10.5)	
1	358 (58.4)	89 (58.2)	
2	192 (31.3)	44 (28.8)	
3	10 (1.6)	4 (2.6)	

Table 1 (*continued*)

Table 1 (*continued*)

Characteristics	Training group (n=613)	Validation group (n=153)	P value
Fibrosis stage			0.833
0	184 (30.0)	45 (29.4)	
1	290 (47.3)	71 (46.4)	
2	112 (18.3)	32 (20.9)	
3/4	27 (4.4)	5 (3.3)	
NAS score	4.0 (3.0, 5.0)	4.0 (3.0, 5.0)	0.857
Definite NASH, n (%)	418 (68.2)	98 (64.1)	0.329

Data are presented as means and SD; medians and IQR; or proportions. [n]: the numbers of data available. NAFLD, non-alcoholic fatty liver disease; BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; CK-18 M30, cytokeratine-18 neoepitope M30; NAS, NAFLD activity score; NASH, nonalcoholic steatohepatitis.

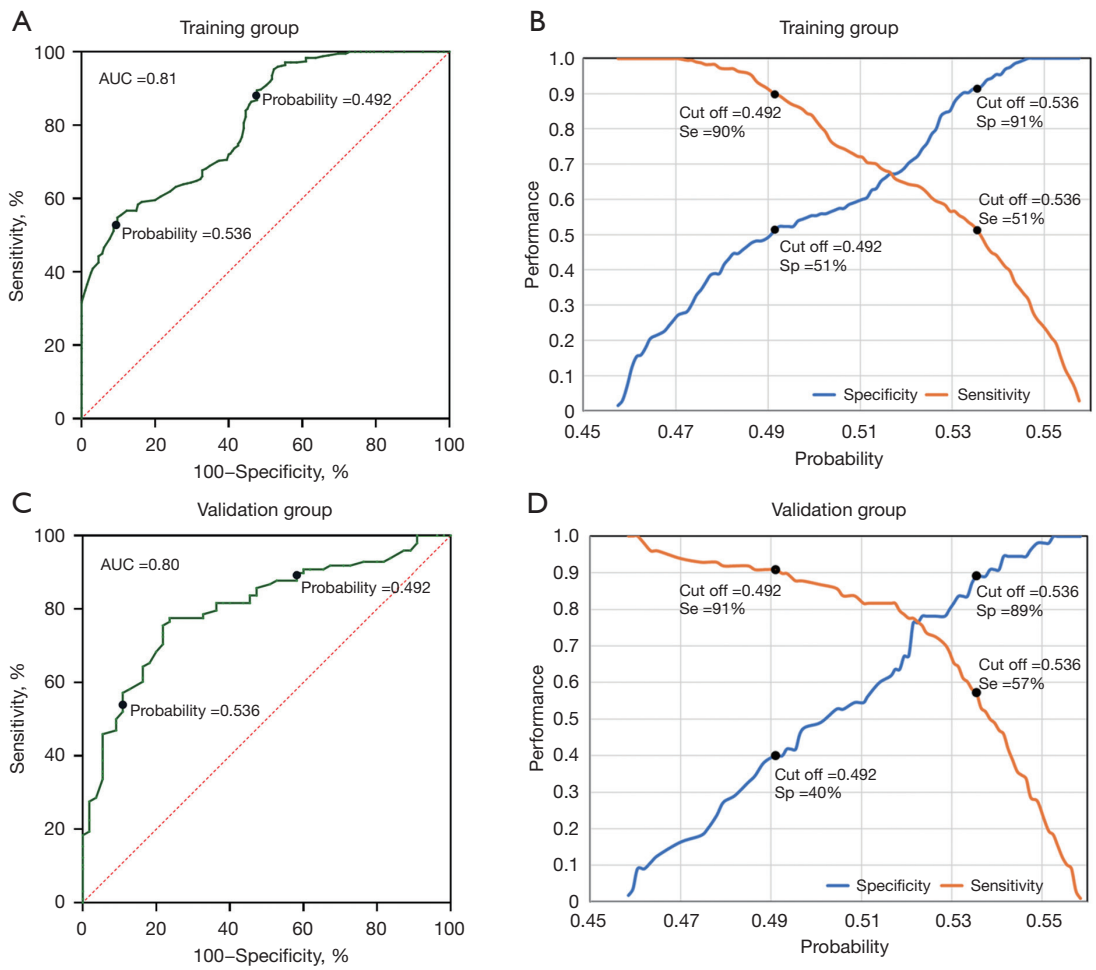

Figure 3 Diagnostic performance of LEARN algorithm and sensitivity, specificity of the dual cut-off values in the training and validation groups. (A,B) Training group. (C,D) Validation group. AUC, area under the curve; LEARN algorithm: deep neural network model for identifying non-alcoholic steatohepatitis; LEARN, bioeLectrical impEdance Analysis foR Nash.

Table 2 Diagnostic performance of the LEARN algorithm

Grouping	AUROC (95% CI)	NASH prevalence (%, n)	Rule-out zone (≤0.492)				Gray zone (0.492, 0.536)	Rule-in zone (≥0.536)			
			n (%)	Sensitivity	Specificity	NPV		n (%)	Sensitivity	Specificity	PPV
Training group	0.81 (0.77, 0.84)	68.2 (418/613)	143 (23.3)	0.90	0.51	0.70	239 (39.0)	231 (37.7)	0.51	0.91	0.93
Validation group	0.80 (0.73, 0.87)	64.1 (98/153)	32 (20.9)	0.91	0.40	0.71	59 (38.6)	62 (40.5)	0.57	0.89	0.90

LEARN, bioeLectrical impEdance Analysis foR Nash; AUROC, area under the receiver operating characteristics; CI, confidence interval; NASH, non-alcoholic steatohepatitis; NPV, negative predictive value; PPV, positive predictive value.

histopathological severity of NAFLD in the training group. We observed that the prediction probability, as calculated by the LEARN algorithm, increased progressively with the histological severity of lobular inflammation, ballooning, steatosis and presence of definite NASH.

Subgroup analyses

We tested the diagnostic performance of the LEARN algorithm in different patient subgroups, in both training and validation groups. As shown in *Table 3*, the LEARN algorithm performed well in all subgroups both in the training group and in the validation group, regardless of sex, age, serum ALT levels and the presence or absence of liver fibrosis, hypertension, diabetes, or obesity (AUROCs ranging from 0.77 to 0.82). In particular, it should be noted that the LEARN algorithm performed well among patients with or without liver fibrosis (AUROCs ranging from 0.77 to 0.83), in both the training and validation groups.

Diagnostic performance of the LEARN algorithm vs. other widely used non-invasive scores or biomarkers for NASH

As shown in *Table 4*, the LEARN algorithm showed a better diagnostic performance for identifying NASH (AUROC: 0.80, 95% CI: 0.77–0.84) compared with other non-invasive NASH scores and the biomarker CK-18 M30. The AUROCs of these non-invasive scores or biomarkers were all less than 0.75 in the whole cohort; in particular, serum CK-18 M30 had an AUROC 0.73 (95% CI: 0.69–0.77); HAIR, 0.63 (95% CI: 0.59–0.67); ION, 0.67 (95% CI: 0.63–0.72); and the NICE model, 0.73 (95% CI: 0.69–0.77).

Diagnostic performance of the LEARN algorithm in the PMBC group

As shown in Table S1, there were 76 NAFLD patients with PMBC in the whole cohort. In order to improve the applicability of the LEARN algorithm in the real world where incomplete BIA data may occur, we adopted the strategy of replacing partially missing data of BIA values by mean values for the group. Table S3 shows that the diagnostic performance of the LEARN algorithm in the PMBC group, and the AUROC was 0.82 (95% CI: 0.72–0.92). As shown in Figure S1, in the PMBC group, the LEARN algorithm showed a greater AUROC compared with ION and HAIR for predicting NASH. The diagnostic performance of the LEARN algorithm in the PMBC group also performed well despite partially missing data.

Discussion

In this large cross-sectional multicenter study, we have developed a novel, fully automatic MLA, referred to as the LEARN algorithm (patent-pending, 2021110501603) to non-invasively diagnose NASH. For patients with biopsy-proven NAFLD who undergo impedance-based measures of body composition and provide simple information on age, sex, diabetes status and hypertension, it is possible to predict their probability of having NASH on histology with acceptable certainty. Our newly developed LEARN algorithm performed well in both the training and validation groups, and across a range of clinically relevant subgroups of patients. To our knowledge, this is the first multicenter study to develop a prediction model based on body composition, for non-invasively identifying NASH.

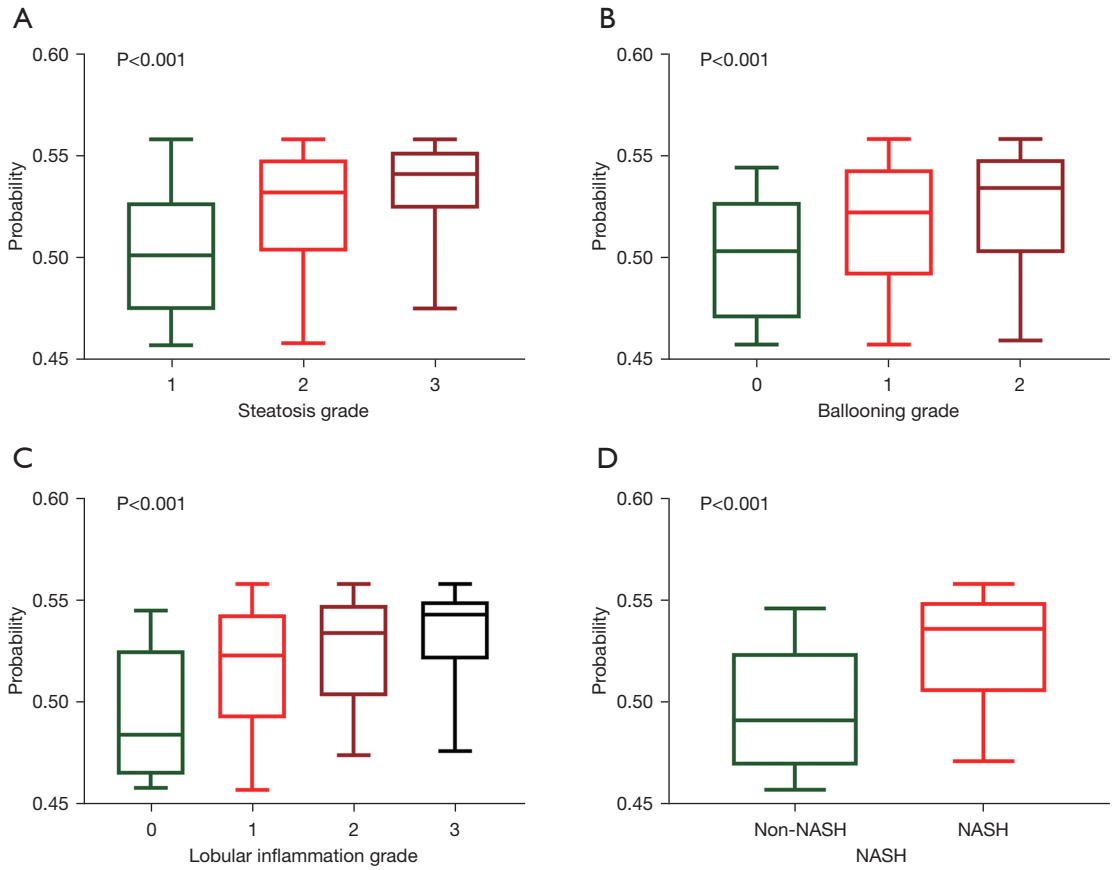

Figure 4 Boxplot of the LEARN algorithm versus histopathological severity of NAFLD in the training group: (A) steatosis grade, (B) ballooning grade, (C) lobular inflammation grade, and (D) presence of definite NASH. NASH, non-alcoholic steatohepatitis; LEARN algorithm: deep neural network model for identifying non-alcoholic steatohepatitis; LEARN, bioeLectrical impEdance Analysis foR Nash; NAFLD, non-alcoholic fatty liver disease.

Besides clinical features of the metabolic syndrome (including hypertension and diabetes), there are other risk factors for a faster progression of NAFLD to NASH. A large number of studies suggested that body composition may be different in NAFLD from that in people without NAFLD, and that there is metabolic dysfunction in NAFLD (16-19). Increased dietary calorie intake and lack of physical exercise may increase the amount of adipose tissue, and accumulation of fat mass may induce insulin resistance and exacerbate liver damage in NAFLD (26-28). Otgonsuren *et al.* (29) showed that anthropometric measures, such as arm circumference and body fat percentage, were significantly higher in NAFLD than in non-steatotic controls. Ko *et al.* (30) found that ultrasound-detected NAFLD was associated with higher BMI, larger waist circumference, and greater body fat mass, through a large sample analysis involving 2,759 participants. Idilman *et al.* (31) showed that visceral adipose tissue alone could be a modest risk factor for predicting NASH (AUROC: 0.64). In addition, arm circumference, percent body fat, BMI, waist circumference, visceral adipose tissue, skeletal muscle mass (sarcopenia) may be also risk factors for greater NAFLD severity (32). Filip *et al.* (33) reported that osteoporosis (as measured by bone mineral content) may also increase the risk of NAFLD. In our study, the LEARN algorithm highlights the utility of body composition measurements for the diagnosis of NASH and this algorithm may help in reducing the number of unnecessary liver biopsies for diagnosing NASH. The value of impedance-based measurements of body composition may also be even greater if the full cost of liver biopsies is to be taken into account (allowing for biopsy-associated complications).

Table 3 Diagnostic performance of LEARN algorithm in different patient subgroups

Subgroup	Training group			Validation group		
	n	AUROC	(95% CI)	n	AUROC	(95% CI)
Sex						
Men	440	0.82	(0.78, 0.86)	108	0.80	(0.72, 0.88)
Women	173	0.82	(0.75, 0.88)	45	0.72	(0.51, 0.92)
Age						
≥50 years	166	0.83	(0.75, 0.91)	35	0.80	(0.65, 0.96)
<50 years	447	0.79	(0.75, 0.83)	118	0.81	(0.73, 089)
Hypertension						
Yes	137	0.87	(0.81, 0.93)	37	0.72	(0.54, 0.90)
No	476	0.79	(0.75, 0.83)	116	0.83	(0.75, 0.90)
ALT level*						
Normal	223	0.79	(0.73, 0.85)	63	0.82	(0.71, 0.93)
Abnormal	390	0.80	(0.76, 0.85)	90	0.77	(0.67, 0.87)
Obesity**						
Yes	209	0.84	(0.79, 0.89)	58	0.79	(0.66, 0.93)
No	404	0.80	(0.76, 0.85)	95	0.77	(0.67, 0.86)
Diabetes						
Yes	147	0.83	(0.76, 0.89)	43	0.79	(0.65, 0.93)
No	466	0.80	(0.76, 0.84)	110	0.80	(0.71, 0.89)
With or without liver fibrosis						
Without fibrosis	184	0.83	(0.78, 0.89)	45	0.82	(0.67, 0.96)
With fibrosis	429	0.78	(0.73, 0.84)	108	0.77	(0.67, 0.86)

*, normal ALT (<40 IU/L); **, obesity (BMI ≥28 kg/m^2). LEARN, bioeLectrical impEdance Analysis foR Nash. NASH, non-alcoholic steatohepatitis; AUROC, area under the receiver operating characteristics; CI, confidence interval; ALT, alanine aminotransferase; BMI, body mass index.

Prediction models lacking transparency and predictability have the potential to cause harm. Our research overcomes this shortcoming. Choosing machine-learning models with high transparency rather than black box models, with high decision-making risk is preferred. Our study uses the classical algorithm of deep learning to develop a new deep neural network for processing big data that includes impedance-based measurements of body composition. In the LEARN algorithm, residual networks were used to prevent gradient disappearance and strengthen the ability of the deep neural network to extract features, thus improving the classification performance of the deep neural network; the dropout method was used to address the over-fitting issue and improve the generalization ability of the deep neural network; fully connected layers were added to analyze the relationship between features more clearly and intuitively, and reduce the influence of feature position on classification. Finally, we re-cycled the residual network to further improve the non-linear expression ability and complexity of the deep neural network. By following this design approach, the LEARN algorithm was optimized.

It is important to underline that in our study we included not only NAFLD patients with elevated serum transaminase levels, but also those without normal serum transaminases who had evidence of hepatic steatosis at recruitment as diagnosed by imaging techniques. For the LEARN

Table 4 Pairwise comparisons between AUROCs for the LEARN algorithm and other non-invasive NASH scores or biomarkers for identifying NASH

Variable(s)	n	AUROC	95% CI	P value
LEARN	766	0.80	0.77–0.84	Reference
CK-18 M30	537	0.73	0.69–0.77	<0.001
HAIR	731	0.63	0.59–0.67	<0.001
ION	651	0.67	0.63–0.72	<0.001
NICE	537	0.73	0.69–0.77	<0.001

The HAIR score for each patient (0–3) was calculated by adding hypertension =1, ALT >40 U/L =1, and HOMA-IR index >5.0 =1. The index of NASH (ION) was calculated as follows: 1.33 waist-to hip ratio + 0.03 × triglycerides (mg/dL) + 0.18 × ALT (U/L) + 8.53 × HOMA-IR – 13.93 in men; 0.02 × triglycerides (mg/dL) + 0.24 × ALT (U/L) + 9.61 × HOMA-IR – 13.99 in women. The NICE model was calculated according to the following equation: –5.654 + 3.780E-02 × ALT (IU/L) + 2.215E-03 × CK18 fragment (IU/L) + 1.825 × (presence of metabolic syndrome =1). AUROC, area under the receiver operating characteristics; LEARN, bioeLectrical impEdance Analysis foR Nash; NASH, non-alcoholic steatohepatitis; CI, confidence interval; CK-18 M30, cytokeratine-18 neoepitope M30; ALT, alanine aminotransferase; HOMA-IR, homeostasis model assessment was used to estimate insulin resistance.

algorithm, we used double cut-off points to identify NASH, as shown in *Figure 3* and *Table 2*. For the purpose of excluding NASH, a lower cut-off value was chosen. For diagnosing NASH, a higher cut-off point was selected. For the LEARN algorithm, the lower cut-off value of 0.492 showed a high sensitivity (90%) and a NPV of 0.70, while the upper cut-off value of 0.536 showed similar specificity (90%) and a PPV of 0.93 in the training and validation groups.

The choice of cut-off values conducive to optimum sensitivity or specificity depends on the purpose of detection. Shown in *Table 2*, there is a gray zone of 39% when dual cut-offs were used to identify NASH. However, it should be noted that there is always a "gray zone" for all non-invasive tests that use two cutoff thresholds (34). On the other hand, approximately 60% of NAFLD patients were able to avoid liver biopsies using our LEARN algorithm when dual cut-offs were chosen.

Currently, treatment of NASH is a major focus of drug development worldwide (35,36). Early, non-invasive identification of NASH for possible drug treatment will be an important medical challenge in the next few years. However, patients with NAFLD, especially those with normal serum ALT levels, and those who are nonobese or do not have diabetes are often ignored in further assessment of NAFLD severity. Therefore, we have also analyzed the diagnostic performance of our newly proposed LEARN algorithm in identifying NASH in different patient subgroups, stratified by obesity, diabetes or serum ALT levels (*Table 3*). Interestingly, our LEARN algorithm performed well in the non-obese, non-diabetic, or serum normal or abnormal ALT (ALT >40 U/L) subgroups, in both the training and validation groups. Importantly, NASH patients with or without fibrosis did not influence the diagnostic performance of the "LEARN" algorithm. Both in the training and the validation groups, the diagnostic performance of the "LEARN" algorithm was above 0.75 among patients with or without liver fibrosis. In the whole cohort, we also compared the diagnostic performance of LEARN algorithm, serum CK-18 M30 level, HAIR, ION and NICE models in identifying NASH, and found that these latter non-invasive scores had moderate accuracy in our cohort, although this finding might be partially affected by differences in the prevalence of NASH among different study populations (24,25,37,38). As shown in *Table 4*, the diagnostic performance of the LEARN algorithm in identifying NASH had an AUROC of 0.80, which is significantly better than other non-invasive NASH scores mentioned above.

In our study, PASS15 was used to estimate the sample size. Each patient was examined for body composition by professionally trained personnel at each center in accordance with uniform operating instructions, and the data were extracted at each of the 6 participating sites by trained data collectors and compiled into spread sheets. Then an experienced AI team used neural network algorithms to build a prediction model that provided clinicians with an individual's probability of having NASH as described above.

Our BIA data were extracted at each of the 6 participating sites by trained data collectors and compiled into spread sheets. The data collection process was checked repeatedly,

to reduce the chance of document errors. However, some of our patient's body composition examination reports were incomplete and reported between 12 and 19 measurements, which we referred to as subjects with 'partial missing BIA data'. We did not exclude these subjects from the analysis in order to improve the utility of our LEARN algorithm in the real world. In the process of building the LEARN algorithm, the AI algorithm has used an average value to replace missing data for those patients who had 'partial missing BIA data'. Specially, we adopted the strategy of replacing PBMC group values by mean values, as this can improve the utility of the LEARN algorithm in the real world where missing data is relatively common. In the PMBC group, the diagnostic performance for identifying NASH performed well with an AUROC of 0.82.

There are some important limitations that should be mentioned. Firstly, the participants were all Chinese of Han ethnicity, so our results might not be applicable to other ethnic groups. Secondly, when comparing results from our cohort with other published studies that used non-invasive scores or biomarkers for diagnosing NASH, the heterogeneity between studies might at least in part contribute to the different diagnostic performances of these non-invasive tests for NASH (e.g., serum CK-18 M30 level, HAIR, ION, NICE models). Specially, the HAIR score system is a non-invasive score for predicting NASH based on hypertension, ALT levels and insulin resistance. When the score is ≥2, the AUROC for predicting NASH is 0.9, and the sensitivity and specificity are 80% and 89%, respectively. However, this model is currently only applicable to patients with BMI >35 kg/m^2 (23). So the applicability is not widespread. In our study, the Han-population was mainly included, and BMI generally concentrated in 24–29 kg/m^2. Therefore, HAIR had a low AUROC in this study. Finally, there was a "gray zone" and 39% patients couldn't be identified with NASH or NAFL. This latter problem is a common limitation for all non-invasive tests where two cut-off thresholds are used (39). In addition, a two-step approach has been also recently reported. By using this two-step approach, patients in the "gray zone" were re-evaluated in combination with other non-invasive diagnostic tests and the need for liver biopsy was reduced significantly without much effect on the percentage of misclassifications (40). In future studies, we will evaluate whether the combination of our LEARN algorithm with other non-invasive NASH scores contributes to the improved stratification of severity of NAFLD.

In conclusion, we have developed a fully automatic LEARN algorithm utilizing impedance-based measurements of body composition along with age, sex, and prior history of hypertension or diabetes, which shows good predictive ability for non-invasively identifying NASH in a large multi-center study across China. Our results suggest that routine measurement of body composition for the assessment of patients with NAFLD may be helpful in staging severity of liver disease and identification of NASH.

Acknowledgments

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), Project of New Century 551 Talent Nurturing in Wenzhou. GT is supported in part by grants from the University School of Medicine of Verona, Verona, Italy. CDB is supported in part by the Southampton NIHR Biomedical Research Centre (IS-BRC-20004), UK.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-523/rc

Data Sharing Statement: Available at https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-523/dss

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-523/coif). MHZ serves as the unpaid editorial board member of *Hepatobiliary Surgery and Nutrition*. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Review Board of the First Affiliated Hospital of Wenzhou Medical University (2016-246) and informed consent was taken from all individual participants.

Open Access Statement: This is an Open Access article

References

1. Younossi Z, Tacke F, Arrese M, et al. Global Perspectives on Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. Hepatology 2019;69:2672-82.
2. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. Hepatology 2018;67:328-57.
3. Brunt EM, Janney CG, Di Bisceglie AM, et al. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol 1999;94:2467-74.
4. Diehl AM, Day C. Cause, Pathogenesis, and Treatment of Nonalcoholic Steatohepatitis. N Engl J Med 2017;377:2063-72.
5. Alkhouri N, Tincopa M, Loomba R, et al. What Does the Future Hold for Patients With Nonalcoholic Steatohepatitis: Diagnostic Strategies and Treatment Options in 2021 and Beyond? Hepatol Commun 2021;5:1810-23.
6. Younossi ZM, Stepanova M, Rafiq N, et al. Pathologic criteria for nonalcoholic steatohepatitis: interprotocol agreement and ability to predict liver-related mortality. Hepatology 2011;53:1874-82.
7. Zhou YJ, Zheng KI, Targher G, et al. Non-invasive diagnosis of non-alcoholic steatohepatitis and liver fibrosis. Lancet Gastroenterol Hepatol 2021;6:9-10.
8. Zhou YJ, Wong VW, Zheng MH. Consensus scoring systems for nonalcoholic fatty liver disease: an unmet clinical need. Hepatobiliary Surg Nutr 2021;10:388-90.
9. Rios RS, Zheng KI, Targher G, et al. Non-invasive fibrosis assessment in non-alcoholic fatty liver disease. Chin Med J (Engl) 2020;133:2743-5.
10. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. N Engl J Med 2016;375:1216-9.
11. Beam AL, Kohane IS. Big Data and Machine Learning in Health Care. JAMA 2018;319:1317-8.
12. Fralick M, Colak E, Mamdani M. Machine Learning in Medicine. N Engl J Med 2019;380:2588-9.
13. Chen K, Nie Y, Park S, et al. Development and Validation of Machine Learning-based Model for the Prediction of Malignancy in Multiple Pulmonary Nodules: Analysis from Multicentric Cohorts. Clin Cancer Res 2021;27:2255-65.
14. Feng G, Zheng KI, Li YY, et al. Machine learning algorithm outperforms fibrosis markers in predicting significant fibrosis in biopsy-confirmed NAFLD. J Hepatobiliary Pancreat Sci 2021;28:593-603.
15. Pietrobelli A, Rubiano F, St-Onge MP, et al. New bioimpedance analysis system: improved phenotyping with whole-body analysis. Eur J Clin Nutr 2004;58:1479-84.
16. Ariya M, Koohpayeh F, Ghaemi A, et al. Assessment of the association between body composition and risk of non-alcoholic fatty liver. PLoS One 2021;16:e0249223.
17. Miyake T, Miyazaki M, Yoshida O, et al. Relationship between body composition and the histology of non-alcoholic fatty liver disease: a cross-sectional study. BMC Gastroenterol 2021;21:170.
18. Schmitz SM, Schooren L, Kroh A, et al. Association of Body Composition and Sarcopenia with NASH in Obese Patients. J Clin Med 2021;10:3445.
19. Samala N, Desai A, Vilar-Gomez E, et al. Decreased Quality of Life Is Significantly Associated With Body Composition in Patients With Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2020;18:2980-8.e4.
20. Li G, Rios RS, Wang XX, et al. Sex influences the association between appendicular skeletal muscle mass to visceral fat area ratio and non-alcoholic steatohepatitis in patients with biopsy-proven non-alcoholic fatty liver disease. Br J Nutr 2022;127:1613-20.
21. Zhou YJ, Ye FZ, Li YY, et al. Individualized risk prediction of significant fibrosis in non-alcoholic fatty liver disease using a novel nomogram. United European Gastroenterol J 2019;7:1124-34.
22. Newsome PN, Sasso M, Deeks JJ, et al. FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. Lancet Gastroenterol Hepatol 2020;5:362-73.
23. Dixon JB, Bhathal PS, O'Brien PE. Nonalcoholic fatty liver disease: predictors of nonalcoholic steatohepatitis and liver fibrosis in the severely obese. Gastroenterology 2001;121:91-100.
24. Anty R, Iannelli A, Patouraux S, et al. A new composite model including metabolic syndrome, alanine

aminotransferase and cytokeratin-18 for the diagnosis of non-alcoholic steatohepatitis in morbidly obese patients. Aliment Pharmacol Ther 2010;32:1315-22.
25. Younes R, Rosso C, Petta S, et al. Usefulness of the index of NASH - ION for the diagnosis of steatohepatitis in patients with non-alcoholic fatty liver: An external validation study. Liver Int 2018;38:715-23.
26. Després JP, Lemieux I. Abdominal obesity and metabolic syndrome. Nature 2006;444:881-7.
27. Kuk JL, Katzmarzyk PT, Nichaman MZ, et al. Visceral fat is an independent predictor of all-cause mortality in men. Obesity (Silver Spring) 2006;14:336-41.
28. Nobarani S, Alaei-Shahmiri F, Aghili R, et al. Visceral Adipose Tissue and Non-alcoholic Fatty Liver Disease in Patients with Type 2 Diabetes. Dig Dis Sci 2022;67:1389-98.
29. Otgonsuren M, Stepanova M, Gerber L, et al. Anthropometric and clinical factors associated with mortality in subjects with nonalcoholic fatty liver disease. Dig Dis Sci 2013;58:1132-40.
30. Ko YH, Wong TC, Hsu YY, et al. The Correlation Between Body Fat, Visceral Fat, and Nonalcoholic Fatty Liver Disease. Metab Syndr Relat Disord 2017;15:304-11.
31. Idilman IS, Low HM, Gidener T, et al. Association between Visceral Adipose Tissue and Non-Alcoholic Steatohepatitis Histology in Patients with Known or Suspected Non-Alcoholic Fatty Liver Disease. J Clin Med 2021;10:2565.
32. Habig G, Smaltz C, Halegoua-DeMarzio D. Presence and Implications of Sarcopenia in Non-alcoholic Steatohepatitis. Metabolites 2021;11:242.
33. Filip R, Radzki RP, Bieńko M. Novel insights into the relationship between nonalcoholic fatty liver disease and osteoporosis. Clin Interv Aging 2018;13:1879-91.
34. Boursier J, Guillaume M, Leroy V, et al. New sequential combinations of non-invasive fibrosis tests provide an accurate diagnosis of advanced fibrosis in NAFLD. J Hepatol 2019;71:389-96.
35. Newsome PN, Buchholtz K, Cusi K, et al. A Placebo-Controlled Trial of Subcutaneous Semaglutide in Nonalcoholic Steatohepatitis. N Engl J Med 2021;384:1113-24.
36. Dufour JF, Caussy C, Loomba R. Combination therapy for non-alcoholic steatohepatitis: rationale, opportunities and challenges. Gut 2020;69:1877-84.
37. Wieckowska A, Zein NN, Yerian LM, et al. In vivo assessment of liver cell apoptosis as a novel biomarker of disease severity in nonalcoholic fatty liver disease. Hepatology 2006;44:27-33.
38. Poynard T, Ratziu V, Charlotte F, et al. Diagnostic value of biochemical markers (NashTest) for the prediction of non alcoholo steato hepatitis in patients with non-alcoholic fatty liver disease. BMC Gastroenterol 2006;6:34.
39. Zhou YJ, Gao F, Liu WY, et al. Screening for compensated advanced chronic liver disease using refined Baveno VI elastography cutoffs in Asian patients with nonalcoholic fatty liver disease. Aliment Pharmacol Ther 2021;54:470-80.
40. 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;35:1804-12.

Cite this article as: Li G, Zheng TL, Chi XL, Zhu YF, Chen JJ, Xu L, Shi JP, Wang XD, Zhao WG, Byrne CD, Targher G, Rios RS, Huang OY, Tang LJ, Zhang SJ, Geng S, Xiao HM, Chen SD, Zhang R, Zheng MH. LEARN algorithm: a novel option for predicting non-alcoholic steatohepatitis. HepatoBiliary Surg Nutr 2023;12(4):507-522. doi: 10.21037/hbsn-21-523

Supplementary

Table S1 Baseline characteristics of patients with or without partial missing data of body composition

Characteristics	Without missing BIA (n=690)	Partial missing BIA data (n=76)	P value
Demographics			
Age (years)	41.6±12.5	36.1±11.4	<0.001
Male sex, n (%)	499 (72.3)	49 (64.5)	0.150
BMI (kg/m^2)	26.9±3.8	29.4±4.1	<0.001
Type 2 diabetes, n (%)	178 (25.8)	12 (15.8)	0.055
Hypertension, n (%)	162 (23.5)	12 (15.8)	0.129
Biochemical measurements			
Albumin (g/L)	46.0±3.9	46.8±3.6	0.091
Platelet count (×10^9)	244.5±62.5	238.0±58.3	0.387
AST (U/L)	34.0 (25.0, 52.0)	37.0 (25.0, 56.2)	0.485
ALT (U/L)	49.5 (30.0, 87.8)	67.1 (34.8, 104.0)	0.016
ALP (U/L)	82.0 (67.2, 99.8)	80.0 (69.0, 94.6)	0.526
Glucose (mmol/L)	5.3 (4.9, 6.3)	5.3 (4.8, 5.9)	0.112
Total cholesterol (mmol/L)	5.1 (4.4, 5.9)	5.1 (4.3, 6.0)	0.383
Triglycerides (mmol/L)	1.9 (1.4, 2.9)	1.7 (1.2, 2.5)	0.058
HDL-C (mmol/L)	1.0±0.3	1.1±0.2	0.002
LDL-C (mmol/L)	3.0±0.9	3.3±0.9	0.004
CK-18 M30 (U/L)	156.1 (74.8, 345.2)	–	–
Body composition			
Body composition analysis			
Intracellular water (e)	24.6±4.7	24.5±4.4 [74]	0.769
Extracellular water (e)	14.8±2.6	16.0±2.4 [74]	<0.001
Total body water (e)	39.5±7.3	40.4±6.7 [74]	0.301
Soft lean mass (g)	50.6±9.7	52.9±9.6 [74]	0.051
Fat free mass (kg)	53.8±10.1	56.6±10.2 [74]	0.022
Muscle-fat analysis			
Weight (kg)	75.9±14.0	82.6±14.7	<0.001
Skeletal muscle mass (kg)	30.1±6.0	32.1±9.9 [47]	0.038
Body fat mass (kg)	22.1±7.4	25.9±9.2	<0.001
Obesity diagnosis			
Percent body fat (%)	28.8±6.6	31.0±7.8	0.008
Waist-hip ratio	0.9±0.0	0.9±0.1 [47]	<0.001

Table S1 (*continued*)

https://dx.doi.org/10.21037/hbsn-21-523

Table S1 (*continued*)

Characteristics	Without missing BIA (n=690)	Partial missing BIA data (n=76)	P value
Lean balance			
Right arm (kg)	3.1±0.7	3.1±1.0	0.882
Left arm (kg)	3.0±0.7	3.0±1.0	0.966
Trunk (kg)	24.6±4.3	25.0±5.7	0.366
Right leg (kg)	8.1±1.7	9.3±2.7	<0.001
Left leg (kg)	8.1±1.7	9.3±2.7	<0.001
Visceral fat area			
Visceral fat area (cm^2)	102.2±27.5	101.0±34.5 [42]	0.701
Additional data			
Body cell mass (kg)	35.3±6.6	38.6±6.9 [11]	0.097
Bone mineral content (kg)	3.0±0.6	3.1±0.6 [40]	0.082
Basal metabolic rate (kcal)	1,537.0±240.7	1,555.1±259.0	0.538
Arm circumference (cm)	33.4±3.1	34.9±1.6 [9]	0.152
Arm muscle circumference (cm)	27.3±2.5	28.2±2.2 [9]	0.259
Non-invasive NASH scores			
ION	38.5 (23.1, 66.6) [625]	29.4 (14.1, 40.6) [26]	0.043
HAIR	1.0 (1.0, 2.0) [670]	1.0 (1.0, 2.0) [61]	0.922
NICE	−2.6 (−3.9, −0.8) [532]	–	–
Liver histology, n (%)			
Steatosis			0.020
1	280 (40.6)	20 (26.3)	
2	272 (39.4)	42 (55.3)	
3	138 (20.0)	14 (18.4)	
Hepatocyte ballooning			0.013
0	85 (12.3)	1 (1.3)	
1	369 (53.5)	43 (56.6)	
2	236 (34.2)	32 (42.1)	
Lobular inflammation			0.004
0	67 (9.7)	2 (2.6)	
1	388 (56.2)	59 (77.6)	
2	222 (32.2)	14 (18.4)	
3	13 (1.9)	1 (1.3)	

Table S1 (*continued*)

https://dx.doi.org/10.21037/hbsn-21-523

Table S1 (*continued*)

Characteristics	Without missing BIA (n=690)	Partial missing BIA data (n=76)	P value
Fibrosis stage			0.126
0	213 (30.9)	16 (21.1)	
1	325 (47.1)	36 (47.4)	
2	123 (17.8)	21 (27.6)	
3/4	29 (4.2)	3 (3.9)	
NAS	4.0 (3.0, 5.0)	4.5 (4.0, 5.0)	0.089
NASH, n (%)	452 (65.5)	64 (84.2)	<0.001

[n]: the numbers of data available. BIA, bioelectrical impedance analysis; BMI, body mass index; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALP, alkaline phosphatase; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; CK-18 M30, cytokeratine-18 neoepitope M30; NASH, non-alcoholic steatohepatitis; NAS, NAFLD activity score.

https://dx.doi.org/10.21037/hbsn-21-523

Table S2 Baseline characteristics of patients, stratified by NASH or non-NASH on histology in the training or validation groups

Characteristics	Training group (n=613)			Validation group (n=153)		
	Non-NASH (n=195)	NASH (n=418)	P value	Non-NASH (n=55)	NASH (n=98)	P value
Demographics						
Age (years)	43.9±11.6	40.2±12.9	<0.001	43.1±9.8	38.2±12.9	0.017
Male sex, n (%)	149 (76.4)	291 (69.6)	0.082	44 (80.0)	64 (65.3)	0.056
BMI (kg/m^2)	26.0 (24.1, 27.7)	27.3 (24.8, 29.7)	<0.001	25.3 (23.4, 27.7)	27.1 (25.0, 29.8)	0.003
Type 2 diabetes, n (%)	47 (24.1)	100 (23.9)	0.961	17 (30.9)	26 (26.5)	0.563
Hypertension, n (%)	48 (24.6)	89 (21.3)	0.358	13 (23.6)	24 (24.5)	0.906
Biochemical measurements						
Albumin (g/L)	45.7±3.6	46.3±4.0	0.031	44.9±4.0	46.8±3.8	0.002
Platelet count (×10^9)	240.9±61.0	244.4±62.5	0.467	251.6±66.6	243.3±60.8	0.588
AST (U/L)	27.0 (22.0, 35.5)	39.0 (27.0, 58.0)	<0.001	26.0 (20.5, 34.0)	47.0 (30.0, 64.8)	<0.001
ALT (U/L)	38.0 (25.0, 57.5)	61.0 (38.0, 104.0)	<0.001	34.0 (21.0, 56.5)	68.5 (36.6, 130.8)	<0.001
ALP (U/L)	80.0 (66.0, 98.5)	83.0 (69.3, 99.0)	0.118	74.0 (60.5, 88.0)	83.0 (68.2, 103.5)	0.019
Glucose (mmol/L)	5.2 (4.8, 6.0)	5.4 (4.9, 6.4)	0.017	5.4 (4.9, 6.8)	5.3 (4.9, 6.5)	0.905
Total cholesterol (mmol/L)	4.9 (4.0, 5.5)	5.2 (4.4, 6.0)	<0.001	4.9 (4.3, 5.5)	5.4 (4.6, 6.0)	0.024
Triglycerides (mmol/L)	1.8 (1.2, 2.5)	2.0 (1.5, 3.0)	<0.001	1.9 (1.4, 3.0)	1.7 (1.3, 2.5)	0.302
HDL-C (mmol/L)	1.0±0.3	1.0±0.3	0.106	1.0 (0.2)	1.1 (0.2)	0.015
LDL-C (mmol/L)	2.9±0.8	3.1±0.9	0.002	3.0 (0.9)	3.2 (1.0)	0.259
CK-18 M30 (U/L)	105.3 (55.0, 175.0)	214.8 (108.5, 468.5)	<0.001	107.6 (54.8, 155.9)	345.0 (148.8, 678.1)	<0.001
Body composition						
Body composition analysis						
Intracellular water (e)	24.6±4.1	24.6±5.0	0.865	24.8±4.1	24.4±4.6	0.524
Extracellular water (e)	14.9±2.4	15.0±2.8	0.684	15.0±2.3	14.7±2.6	0.348
Total body water (e)	39.6±6.7	39.7±7.6	0.799	39.8±6.4	39.0±7.1	0.440
Soft lean mass (g)	50.9±8.8	50.9±10.3	0.835	51.0±8.6	50.2±9.2	0.529
Fat free mass (kg)	53.9±8.9	54.3±11.0	0.654	54.3±8.8	53.2±9.8	0.454
Muscle-fat analysis						
Weight (kg)	74.4±12.6	78.1±15.1	0.002	72.7±10.6	77.0±13.8	0.102
Skeletal muscle mass (kg)	30.2±5.4	30.5±6.9	0.715	30.3±5.5	29.7±6.0	0.531
Body fat mass (kg)	20.4±6.8	23.6±7.9	<0.001	18.3±5.1	23.7±7.5	<0.001
Obesity diagnosis						
Percent body fat (%)	27.2±6.5	30.1±6.6	<0.001	25.2±5.8	30.6±6.9	<0.001
Waist-hip ratio	0.9±0.0	0.9±0.1	0.084	0.9±0.0	0.9±0.1	0.012

Table S2 (*continued*)

https://dx.doi.org/10.21037/hbsn-21-523

Table S2 (*continued*)

Characteristics	Training group (n=613)			Validation group (n=153)		
	Non-NASH (n=195)	NASH (n=418)	P value	Non-NASH (n=55)	NASH (n=98)	P value
Lean balance						
Right arm (kg)	3.1±0.6	3.1±0.8	0.541	3.1±0.7	3.0±0.7	0.286
Left arm (kg)	3.1±0.6	3.0±0.8	0.884	3.1±0.7	3.0±0.7	0.338
Trunk (kg)	24.7±3.8	24.6±4.7	0.965	24.8±4.3	24.2±4.3	0.306
Right leg (kg)	8.3±1.6	8.3±2.0	0.607	8.2±1.6	8.1±1.6	0.985
Left leg (kg)	8.2±1.6	8.3±2.0	0.658	8.1±1.6	8.1±1.6	0.955
Visceral fat area						
Visceral fat area (cm^2)	97.4±24.4	104.7±29.2	0.008	92.3±23.6	106.4±28.9	0.002
Additional data						
Body cell mass (kg)	35.2±5.8	35.5±7.0	0.666	35.5±6.0	35.0±6.7	0.631
Bone mineral content (kg)	2.9±0.5	3.0±0.6	0.250	2.9±0.5	3.0±0.6	0.896
Basal metabolic rate (kcal)	1,535.9±204.0	1,544.2±268.0	0.630	1,551.9±217.7	1,514.1±210.9	0.326
Arm circumference (cm)	32.9±2.5	33.8±3.4	0.002	32.7±2.6	33.6±3.2	0.110
Arm muscle circumference (cm)	27.2±2.2	27.4±2.6	0.229	27.5±2.6	27.1±2.7	0.343
Non-invasive NASH scores						
LEARN	0.50±0.03	0.53±0.02	<0.001	0.50±0.03	0.53±0.03	<0.001
ION	29.0 (15.4, 45.8)	43.9 (27.8, 71.1)	<0.001	25.1 (15.5, 37.1)	44.6 (33.5, 83.1)	<0.001
HAIR	1.0 (0.0, 1.0)	1.0 (1.0, 2.0)	<0.001	1.0 (0.0, 1.0)	1.0 (1.0, 2.0)	<0.001
NICE	–3.4 (–4.5, –2.1)	–2.0 (–3.2, 0.3)	<0.001	–3.6 (–4.5, –2.4)	–0.9 (–3.2, 2.6)	<0.001
Liver histology, n (%)						
Fibrosis stage			<0.001			<0.001
0	111 (56.9)	73 (17.5)		31 (56.4)	14 (14.3)	
1	67 (34.4)	223 (53.3)		21 (38.2)	50 (51.0)	
2	11 (5.6)	101 (24.2)		1 (1.8)	31 (31.6)	
3/4	6 (3.1)	21 (5.0)		2 (3.6)	3 (3.1)	

NASH, non-alcoholic steatohepatitis; BMI, body mass index; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALP, alkaline phosphatase; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; CK-18 M30, cytokeratine-18 neoepitope M30; LEARN, bioeLectrical impEdance Analysis foR Nash.

https://dx.doi.org/10.21037/hbsn-21-523

Table S3 Diagnostic performance of LEARN algorithm in groups of patients with or without partial missing data of body composition

Grouping	n	AUROC	95% CI
With or without partial missing BIA data			
WMBC group	690	0.80	(0.77, 0.83)
PMBC group	76	0.82	(0.72, 0.92)

LEARN, bioeLectrical impEdance Analysis foR Nash; AUROC, area under the receiver operating characteristics; CI, confidence interval; BIA, bioelectrical impedance analysis; WMBC, without partial missing data of body composition; PMBC, partial missing data of body composition.

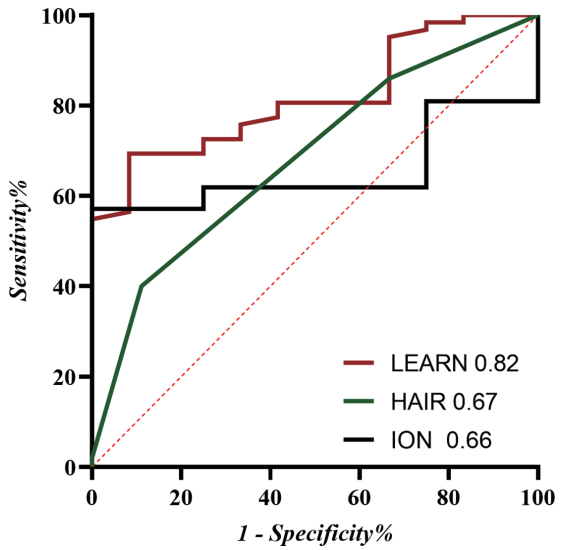

Figure S1 Pairwise comparison of ROC curves between the deep neural network model for identifying NASH (LEARN algorithm) and ION and HAIR models in the group of patients with partial missing data of body composition.

https://dx.doi.org/10.21037/hbsn-21-523