Liver biopsy-based validation, confirmation and comparison of the diagnostic performance of established and novel non-invasive non-alcoholic fatty liver disease indexes

Kouvari, Matina; Valenzuela-Vallejo, Laura; Guatibonza-Garcia, Valentina; Polyzos, Stergios A.; Deng, Yixiang; Kokkorakis, Michail; Agraz, Melih; Mylonakis, Sophia C.; Katsarou, Angeliki; Verrastro, Ornella

Published in:
Metabolism

DOI:
10.1016/j.metabol.2023.155666

Publication date:
2023

Metabolism 147 (2023) 155666

Contents lists available at ScienceDirect

Metabolism

journal homepage: www.journals.elsevier.com/metabolism

Liver biopsy-based validation, confirmation and comparison of the diagnostic performance of established and novel non-invasive steatotic liver disease indexes: Results from a large multi-center study

Matina Kouvari [a,1], Laura Valenzuela-Vallejo [a,1], Valentina Guatibonza-Garcia [a,1], Stergios A. Polyzos [b], Yixiang Deng [a], Michail Kokkorakis [a,c], Melih Agraz [a], Sophia C. Mylonakis [a], Angeliki Katsarou [a], Ornella Verrastro [d], Georgios Markakis [e], Mohammed Eslam [f], Georgios Papatheodoridis [e], Jacob George [f,*], Geltrude Mingrone [d], Christos S. Mantzoros [a,g]

[a] Department of Medicine, Beth-Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215, United States of America
[b] First Laboratory of Pharmacology, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece
[c] Department of Clinical Pharmacy and Pharmacology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands
[d] Department of Gastroenterology, Università Cattolica del Sacro Cuore, Rome, Italy
[e] Department of Gastroenterology, Medical School of National and Kapodistrian University of Athens, General Hospital of Athens "Laiko", Athens, Greece
[f] Storr Liver Centre, Westmead Institute for Medical Research, Westmead Hospital and University of Sydney, Sydney, New South Wales, Australia
[g] Department of Medicine, Boston VA Healthcare System, Boston, MA 02130, United States of America

ARTICLE INFO

Keywords:
Metabolic-dysfunction associated steatotic liver disease (MASLD)
Non-alcoholic fatty liver disease (NAFLD)
Non-alcoholic steatohepatitis (NASH)
Liver biopsy
Validation study
Diagnosis

ABSTRACT

Background: Non-invasive tools (NIT) for metabolic-dysfunction associated liver disease (MASLD) screening or diagnosis need to be thoroughly validated using liver biopsies.
Purpose: To externally validate NITs designed to differentiate the presence or absence of liver steatosis as well as more advanced disease stages, to confirm fully validated indexes ($n = 7$ NITs), to fully validate partially validated indexes ($n = 5$ NITs), and to validate for the first time one new index ($n = 1$ NIT).
Methods: This is a multi-center study from two Gastroenterology-Hepatology Departments (Greece and Australia) and one Bariatric-Metabolic Surgery Department (Italy). Overall, $n = 455$ serum samples of patients with biopsy-proven MASLD ($n = 374$, including 237 patients with metabolic-dysfunction associated steatohepatitis (MASH)) and Controls ($n = 81$) were recruited. A complete validation analysis was performed to differentiate the presence of MASLD vs. Controls, MASH vs. metabolic-dysfunction associated steatotic liver (MASL), histological features of MASH, and fibrosis stages.
Results: The index of NASH (ION) demonstrated the highest differentiation ability for the presence of MASLD vs. Controls, with the area under the curve (AUC) being 0.894. For specific histological characterization of MASH, no NIT demonstrated adequate performance, while in the case of specific features of MASH, such as hepatocellular ballooning and lobular inflammation, ION demonstrated the best performance with AUC being close to or above 0.850. For fibrosis (F) classification, the highest AUC was reached by the aspartate aminotransferase to platelet ratio index (APRI) being ~0.850 yet only with the potential to differentiate the severe fibrosis stages (F3, F4) vs. mild or moderate fibrosis (F0–2) with an AUC > 0.900 in patients without T2DM. When we excluded

Abbreviations: ALT, alanine aminotransferase; APRI, aspartate aminotransferase to Platelet Ratio Index; AST, aspartate aminotransferase; AUC, area under the curve; CRN, Clinical Research Network; FIB-4, Fibrosis-4 Index; FLI, Fatty Liver Index; HSI, Hepatic Steatosis Index; ION, index of NASH; LAP, Lipid Accumulation Product; LFS, Liver Fat Score; MASLD, Metabolic-dysfunction associated steatotic liver disease; MAFL, Metabolic-dysfunction associated steatotic liver; MASH, Metabolic-dysfunction associated steatohepatitis; NAFLD, non-alcoholic fatty liver disease; NAS, Non-alcoholic Fatty Liver Disease Activity Score; NFS, Non-alcoholic Fatty Liver Disease Fibrosis Score; NIT, non-invasive tools; NPV, Negative predictive value; PPV, Positive predictive value; ROC, Receiver operating characteristic; T2DM, type 2 diabetes mellitus; TyG, Triglyceride to Glucose index; TyGO, Original Triglyceride to Glucose index.

* Corresponding author at: Sydney Medical School, The Westmead Institute for Medical Research, Westmead Hospital, Sydney West Local Health District, 176 Hawkesbury Road, Westmead, NSW 2145, Australia.
E-mail address: jacob.george@sydney.edu.au (J. George).
[1] These authors contributed equally.

https://doi.org/10.1016/j.metabol.2023.155666
Received 31 March 2023; Accepted 25 July 2023
Available online 30 July 2023

patients with morbid obesity, the differentiation ability of APRI was improved, reaching AUC = 0.802 for differentiating the presence of fibrosis F2–4 vs. F0–1. The recommended by current guidelines index FIB-4 seemed to differentiate adequately between severe (i.e., F3–4) and mild or moderate fibrosis (F0–2) with an AUC = 0.820, yet this was not the case when FIB-4 was used to classify patients with fibrosis F2–4 vs. F0–1. Trying to improve the predictive value of all NITs, using Youden's methodology, to optimize the suggested cut-off points did not materially improve the results.

Conclusions: The validation of currently available NITs using biopsy-proven samples provides new evidence for their ability to differentiate between specific disease stages, histological features, and, most importantly, fibrosis grading. The overall performance of the examined NITs needs to be further improved for applications in the clinic.

1. Introduction

NAFLD - or MASLD which is the latest term very recently suggested in a multi-society Delphi consensus statement [45] - is a chronic liver disease that affects more than a quarter of the global population [1]. The prevalence of MASLD is increasing worldwide, from 43.1 % in 2020 to 50 % by 2040 in North America, and from 43.4 % in 2020 to over 60 % by 2040 in Europe and Asia [2]. The rising burden of MASLD is strongly associated with the pandemic of obesity and its cardio-metabolic consequences [3]. MASH is also in the epicenter of scientific interest [4]; at least 20–30 % of patients with MASLD develop MASH, which can lead to cirrhosis and associated complications, including hepatocellular cancer [5]. Moreover, with the aging of the affected population and longer exposure to the condition, the disease burden from cirrhosis due to MASLD is expected to increase around two to three-fold from 2015 to 2030 in many regions around the globe [6].

Despite these alarming trends, >70 % of MASLD cases are underdiagnosed, and only 3 % of patients with MASLD and at high risk of advanced fibrosis development receive specialized care [7]. The reference standard for the evaluation of MASLD and determination of disease progression is the histological examination of liver tissue through liver biopsy [8]. However, this process is followed by patient discomfort and risk of complications [8]. Simple NITs used at least for the first screening of a patient with a high-risk metabolic profile to determine the presence of liver abnormalities are highly demanded [9]. This approach can be cost-effective for the healthcare system, limiting unneeded clinical examinations and driving at the same time the treatment [10]. Towards this perspective, non-invasive diagnostic scores, which include accessible and widely available clinical, biochemical, and metabolic parameters, have been developed [11–19]. These NITs have been validated using liver imaging or histological assessment of the liver for specific MASLD features. However, to obtain comparable results, the validation process should be performed against the gold standard method of liver biopsy, which has not always been done.

Thus, we used the data from a multi-center, diagnostic accuracy study, to test the performance of thirteen (13) NITs using biopsy as a gold standard. The specific objectives were to examine **(a)** the potential use of the available NITs to differentiate MASLD, NASH, and liver fibrosis, i.e., above and beyond the initial context of use for which these indexes were created; **(b)** the potential use of different cut-off points – compared with the standard cut-off points – for each index after maximizing sensitivity or specificity using Youden's methodology to rule out or rule in the presence of liver pathology.

2. Methods

2.1. Sample and setting

Overall, $n = 455$ data and serum samples of patients with available histological assessment of the liver through biopsy were collected in three medical centers. More specifically, $n = 297$ serum samples (235 MASLD cases; 62 Controls (i.e., absence of MASLD)) were collected by the Bariatric-Metabolic Surgery Department of the Università Cattolica del Sacro Cuore, Rome, Italy (Study 1). Additionally, $n = 36$ samples (all MASLD cases) were collected in Greece by the Gastroenterology-Hepatology Department of the Medical School of National and Kapodistrian University of Athens, General Hospital of Athens “Laiko” (Study 2) as well as $n = 122$ serum samples (103 MASLD cases; 19 Controls (i.e., absence of MASLD)) were collected in Australia, by the Gastroenterology-Hepatology Department of the University of Sydney, Westmead Hospital, and Sydney West Local Health District (Study 3). Biochemical assessment (i.e., chemistry and hematology (e.g., ALT, AST, platelets, etc.)), as well as insulin and lipid profiles in participants' serum, were implemented after participant assignment to the study.

More details on the study centers and the specific criteria according to which patients were assigned for liver biopsy can be found in Table 1.

2.1.1. Bioethics

All studies were done in accordance with the Declaration of Helsinki, Good Clinical Practice, and applicable regulatory requirements. The protocol in Study 1 was approved by the Ethics Committees of Fondazione Policlinico A Gemelli, Policlinico Umberto I, and Azienda Ospedaliera San Camillo-Forlanini, Rome, Italy. The protocol in Study 2 was approved by the Ethical Committee in the General Hospital of Athens “Laiko”. The protocol in Study 3 was approved by the Western Sydney Local Health District. All participants provided their consent form to undergo liver biopsy. The anonymized data were shared with Beth Israel Deaconess Medical Center (BIDMC) without any possibility for deidentification. The BIDMC Institutional Review Board (IRB) provided an exempt ethics approval.

2.1.2. Study 1

Samples from Study 1 were collected in the context of a randomized-controlled clinical trial [20]. In particular, between April 15, 2019, and June 21, 2021, $n = 431$ individuals eligible for bariatric surgery that visited the Bariatric-Metabolic Surgery Department of the Università Cattolica del Sacro Cuore in Rome, Italy were referred to liver biopsy if they met the referral criteria described in Table 1; enrollment in the study was considered upon positive liver biopsy. Candidates without MASLD were enrolled from patients that underwent laparoscopic elective cholecystectomy in the clinic; enrollment in the study was considered upon negative liver biopsy.

2.1.3. Study 2

The Study 2 was performed in a Gastroenterology-Hepatology Department of the University of Athens Laikon Hospital, in Athens Greece. Patients were collected retrospectively using hospital health records (Table 1). Enrollment in the study was considered upon positive liver biopsy. The maximum time gap between liver biopsy and serum sample selection was six months. In total, $n = 69$ patients with biopsy-proven MASLD were examined for potential inclusion in the study. Overall, $n = 33$ patients were excluded due to the time gap between biopsy and serum sample selection above the accepted time limit for a final sample size of $n = 36$ patients with biopsy-proven MASLD.

2.1.4. Study 3

The Study 3 was performed in a Gastroenterology-Hepatology Department of the University of Sydney, Westmead Hospital, Sydney

West Local Health District in Australia. Individuals visiting the clinic with indications of impaired liver function tests – as summarized in Table 1 – were assigned to liver biopsy; positive or negative biopsy resulted in enrollment to the study as MASLD patient or control. Candidates without MASLD (Controls) were also collected from individuals that underwent gallbladder surgery in the clinic; enrollment to the study was considered upon negative liver biopsy.

Due to the similar profile of MASLD patients assigned in Study 2 and Study 3 (i.e., patients visiting Gastroenterology-Hepatology Departments with impaired liver function tests), the respective samples were merged and examined against the participants assigned in Study 1 (i.e., patients with class I-III obesity and T2DM evaluated in the Bariatric-Metabolic Surgery Department).

2.2. Diagnosis of MASLD, MASH, and fibrotic staging

The presence of MASLD and MASH was defined according to the recently suggested algorithm [45]. Percutaneous liver biopsies were performed using ultrasonography. Each biopsy was assessed centrally and subsequently read by two expert hepato-pathologists independently to assess NAS and fibrosis stage (according to NASH-CRN criteria) in a blinded manner (degree of concordance 85 %). Diagnosis of MASLD was defined according to the NASH CRN criteria through biopsy as per standard procedure and after excluding other liver diseases or other causes of secondary fatty liver disease [21,22]. MASH was defined as NAS $\geq$ 4, including a score of 1 or more in each component according to the NAS system – i.e. steatosis, ballooning degeneration, and lobular inflammation, on screening or historic (within the previous 6 months) biopsy as determined by a single central reader [21,22]. Steatosis was graded from 0 to 3. The presence and stages of liver fibrosis were performed using the NASH-CRN system: Stage 0 indicates no fibrosis (F0), Stage 1 indicates centrilobular pericellular fibrosis (F1), Stage 2 indicates centrilobular and periportal fibrosis (F2), Stage 3 indicates bridging fibrosis (F3), and Stage 4 indicates cirrhosis (F4) [12,16].

Table 1
Description of studies.

Region	Study 1		Study 2		Study 3	
	Italy		Greece		Australia	
	Controls $N = 62$	Patients $N = 235$	Controls $N = 0$	Patients $N = 36$	Controls $N = 19$	Patients $N = 103$
Center	Bariatric-Metabolic Surgery Department, Università Cattolica del Sacro Cuore, Rome, Italy		Department of Gastroenterology of the Medical School of National and Kapodistrian University of Athens, General Hospital of Athens "Laiko"		Department of Gastroenterology-Hepatology of the University of Sydney, Westmead Hospital, Sydney West Local Health District	
Number of centers	Single-center		Single-center		Single-center	
Type of center	Bariatric-Metabolic Surgery Department		Gastroenterology-Hepatology Department		Gastroenterology-Hepatology Department	
Type of care	Secondary care		Secondary care		Secondary care	
Recruitment of participants	Prospective		Retrospective		Retrospective	
Serum collection in relation to liver biopsy	Cross-sectional (serum samples selected at the same time as liver biopsy)		Cross-sectional (allowed time gap between serum selection and liver biopsy $\leq$ 6 months)		Cross-sectional (serum samples selected at the same time as liver biopsy)	
Data source	In the context of a randomized controlled clinical trial		In the context of a cohort study utilizing health records		In the context of a cohort study	
Candidate patients referred for liver biopsy	▪ Candidates for bariatric surgery with BMI = 30–55 kg/m^2 and type 2 diabetes **AND** ▪ Confirmed steatosis via ultrasonography **AND** ▪ NAFLD Fibrosis score > −1.455 **AND** ▪ Signed consent form		▪ Abnormal liver function tests **AND/OR** ▪ FibroScan- liver stiffness >8 kPa **AND** ▪ Signed consent form		Criteria before 2008 ▪ Abnormal liver function tests after a 4-month lifestyle intervention Criteria 2008 onwards ▪ FibroScan- liver stiffness >12 kPa **AND** ▪ Signed consent form	
Candidate controls referred for liver biopsy	▪ Laparoscopic elective cholecystectomy **AND** ▪ Signed consent form		*No controls assigned*		Criteria before 2008 ▪ Abnormal liver function tests after a 4-month lifestyle intervention Criteria 2008 onwards ▪ FibroScan- liver stiffness > 12 kPa OR ▪ Gall bladder surgery **AND** ▪ Signed consent form	
Enrollment in the study, patients	Positive liver biopsy		Positive liver biopsy		Positive liver biopsy	
Enrollment in the study, controls	Negative liver biopsy		*No controls assigned*		Negative liver biopsy	

Patients refer to participants with metabolic-dysfunction associated steatotic liver disease (MASLD) confirmed through liver biopsy.
Controls refer to free-of-MASLD participants confirmed through liver biopsy.
Abbreviations: Alanine aminotransferase (ALT); Fibrosis score 4 (FIB-4); Non-alcoholic fatty liver disease (NAFLD).

Hepatocellular ballooning grading 0 (i.e., none), 1 (i.e., few), and 2 (i.e., many), as well as lobular inflammation from grade 0 to grade 3 were assessed [19]; for the scope of the present work, all grades ≥ 1 were merged in one category and examined against no ballooning or no inflammation.

2.3. Non-invasive assessment of MASLD, MASH, and fibrotic stages

Overall, $n = 13$ NITs were examined for their ability to differentiate the presence of MASLD and the specific histological characteristics (simple steatosis, hepatocellular ballooning, lobular inflammation) and grading of liver fibrosis. Each NIT was examined according to all the aforementioned liver outcomes – irrespective of the initially recommended context of use. To this issue, we examined **a**. steatosis-related indexes (i.e., HSI, FLI, LAP, NAFLD-LFS), **b**. fibrosis-related indexes (i.e., FIB-4, APRI, and NFS) and **c.** other MASLD-related indexes (i.e., ION, TyGO and TyG, AST/ALT ratio, ALT/AST ratio), for which the original TyGO is presented and compared with its variation TyG, that has been found to present an incidental modification of the formula, creating confusion in the literature [23]. All the formulas and standard or previously suggested cut-off points are summarized in ***Supplementary Table 1***. Additionally, we proposed an exploratory adjustment for HSI called α-HSI, which has a similar structure to HSI yet reversing the ratio of transaminases (i.e., from ALT/AST to AST/ALT).

2.4. Statistical analyses

Continuous variables are presented as mean (standard error of mean). In the case of continuous, normally distributed variables, *p*-values were obtained through Student's *t*-test for independent samples. The normality of the continuous variables' distribution was tested through the P—P plot and the Shapiro-Wilk test. Post-hoc pairwise comparisons with Bonferroni correction were also performed. In the case of categorical variables, p-values were obtained through Chi-square test. The differentiation abilities of the examined NITs in relation to the presence vs. absence of MASLD presence as well as different MASLD and fibrosis stages was examined by assessing the AUC and the corresponding 95 % Confidence Interval through ROC analysis. Sensitivity (Recall), Specificity, PPV; Precision, and NPV were evaluated using standard formulas to evaluate the performance of the indexes as per suggested cut-off points. Accuracy was also evaluated to depict the overall performance of the index. Two cut-off points were suggested to rule in or rule out participants through maximizing specificity or sensitivity, respectively, at 0.70. All statistical analyses were performed with IBM SPSS Statistics for Windows, Version 28.0 (IBM Corp., Armonk, NY) for Windows, and GraphPad Prism 9.3.1. (GraphPad Software Inc., La Jolla, CA), and R studio platforms (Boston, MA, 02210).

Table 2
Participants' demographic, clinical, and biochemical characteristics according to the presence of MASLD as well as specific MASLD stages, per study center ($n = 455$).

Study center		Controls	All MASLD	MASL	MASH	p-Value	p-Value[a]
Gastroenterology-Hepatology Department	N	19	139	48	91	–	–
Bariatric-Metabolic Surgery Department		62	235	89	146	–	–
Gastroenterology-Hepatology Department	Age, years	51 (3)	56 (1)	54 (2)	57 (1)	0.105	0.100
Bariatric-Metabolic Surgery Department		40 (1)	49 (1)	49 (1)	48 (1)	**<0.001**	0.528
Gastroenterology-Hepatology Department	Female sex, %	74	50	42	55	**0.048**	0.139
Bariatric-Metabolic Surgery Department		71	52	54	51	**0.006**	0.704
Gastroenterology-Hepatology Department	BMI, kg/m^2	31.2 (1.70)	32.6 (0.55)	32.27 (1.04)	32.74 (0.64)	0.434	0.704
Bariatric-Metabolic Surgery Department		31.2 (1.35)	42.8 (0.37)	41.96 (0.63)	43.25 (0.45)	**<0.001**	0.096
Gastroenterology-Hepatology Department	Waist circumference, cm	95 (5)	107 (1)	106 (3)	108 (1)	**0.040**	0.523
Bariatric-Metabolic Surgery Department		121 (3)	131 (1)	129 (2)	132 (1)	**0.007**	0.114
Gastroenterology-Hepatology Department	Waist-to-hip ratio	0.88 (0.02)	0.98 (0.09)	0.96 (0.02)	0.99 (0.01)	**<0.001**	0.088
Bariatric-Metabolic Surgery Department		0.90 (0.03)	1.00 (0.06)	0.98 (0.01)	1.00 (0.07)	**0.002**	0.095
Gastroenterology-Hepatology Department	HOMA-IR, units	2.36 (0.36)	5.58 (0.30)	4.53 (0.52)	6.15 (0.35)	**<0.001**	**0.012**
Bariatric-Metabolic Surgery Department		2.14 (0.13)	6.77 (0.30)	5.19 (0.41)	7.70 (0.39)	**<0.001**	**<0.001**
Gastroenterology-Hepatology Department	Type 2 diabetes, %	16	58	49	62	**<0.001**	0.167
Bariatric-Metabolic Surgery Department		2	44	39	47	**<0.001**	0.235
Gastroenterology-Hepatology Department	CHOL, mg/dL	202 (11)	174 (4)	178 (7)	172.5 (5)	**0.025**	0.550
Bariatric-Metabolic Surgery Department		174 (4)	190 (2)	185 (4)	193 (3)	**<0.001**	0.112
Gastroenterology-Hepatology Department	LDL-C, mg/dL	124 (9)	98 (4)	106 (7)	95 (5)	**0.022**	0.181
Bariatric-Metabolic Surgery Department		104 (3)	115 (2)	116 (4)	115 (3)	**0.001**	0.808
Gastroenterology-Hepatology Department	Triglycerides, mg/dL	109 (13)	148 (6)	135 (10)	155 (7)	**0.009**	0.098
Bariatric-Metabolic Surgery Department		103 (3)	152 (4)	138 (6)	161 (5)	**<0.001**	**0.007**
Gastroenterology-Hepatology Department	HDL-C, mg/dL	61(4)	47 (1)	47 (2)	47 (2)	**<0.001**	0.959
Bariatric-Metabolic Surgery Department		57 (2)	46 (1)	47 (1)	45 (1)	**<0.001**	0.114
Gastroenterology-Hepatology Department	AST, UI/L	40 (5)	46 (2)	39 (3)	50 (2)	0.240	**0.005**
Bariatric-Metabolic Surgery Department		16 (1)	25 (1)	21 (1)	28 (1)	**<0.001**	**<0.001**
Gastroenterology-Hepatology Department	ALT, UI/L	50 (6)	62 (3)	49 (4)	68 (3)	0.090	**0.004**
Bariatric-Metabolic Surgery Department		17 (1)	34 (1)	26 (1)	40 (2)	**<0.001**	**<0.001**
Gastroenterology-Hepatology Department	GGT, UI/L	101 (15)	111 (9)	80 (11)	126 (12)	0.588	**0.006**
Bariatric-Metabolic Surgery Department		21 (2)	36 (1)	27 (2)	41 (2)	**<0.001**	**<0.001**
Gastroenterology-Hepatology Department	Albumin, g/L	42 (1)	42 (0)	43 (1)	42 (0)	0.881	0.109
Bariatric-Metabolic Surgery Department		40 (0)	41 (0)	41 (0)	41 (0)	**<0.001**	0.317
Gastroenterology-Hepatology Department	Platelets, x 10^9/L	246 (18)	219 (6)	228 (10)	214 (8)	0.172	0.138
Bariatric-Metabolic Surgery Department		209 (5)	244 (4)	239 (6)	248 (5)	**<0.001**	0.248

Continuous variables are presented as mean (SEM), and categorical variables as frequencies (%). P-values for continuous variables were obtained through Student's *t*-test for independent samples. P-values for categorical variables were obtained through Chi-squared test. Bold indicates statistical significant differences (p-value<0.05).

Abbreviations: Alanine transaminase (ALT); Aspartate aminotransferase (AST); Body mass index (BMI); Gamma-glutamyl transferase (GGT); High-density lipoprotein cholesterol (HDL-C); Homeostatic Model Assessment for Insulin Resistance (HOMA-IR); Low-density lipoprotein cholesterol (LDL-C); Metabolic-dysfunction associated steatotic liver disease (MASLD); Metabolic-dysfunction associated steatohepatitis (MASH), Total cholesterol (CHOL).

[a] p-values for the comparisons between MASL and MASH.

3. Results

3.1. Demographic, clinical, and biochemical characteristics of the study participants

Overall, $n = 455$ subjects, including $n = 374$ cases with MASLD (63 % NASH diagnosis) were recorded. Demographic, clinical, and biochemical characteristics (including glucose and lipid metabolism, and liver function markers) of participants per study center are summarized in Table 2. MASLD subjects had BMI within the range of obesity (78 % subjects with obesity), and 53 % had morbid obesity (i.e., Class III) (*data not shown in Table*). More specifically, MASLD patients from the Bariatric-Metabolic Surgery Department had around 10 points higher BMI (42.8 kg/m^2) than the MASLD patients recruited from a Gastroenterology-Hepatology Department (32.6 kg/m^2), which is justified given the eligibility criteria for bariatric surgery (as per protocol). Central obesity was recorded in almost all MASLD patients. Investigating further the profile of Controls, the BMI was similar in both studied samples and close to 31 kg/m^2; however, almost all Controls from the Bariatric-Metabolic Surgery Department had central obesity, with the respective frequency in Controls from the Gastroenterology-Hepatology Department being 47 %. Finally, renal function was within normal age-appropriate ranges for all groups (*data not shown in Table*).

3.2. The differentiation ability of the examined NITs

Irrespective of the original scope for which the examined NITs were created or previously validated, we examined their overall differentiation ability against many different outcomes, from the presence of overall MASLD, MASL, or MASH, to histological characteristics of NASH and liver fibrosis stages. All results are illustrated in Fig. 1. Overall, ION presented the highest ability to differentiate the presence vs. absence of MASLD, with the AUC being 0.894, as well as to differentiate the presence of MASH among MASL patients, yet the AUC did not exceed 0.800 (AUC = 0.747). In terms of histologic characterization of MASLD, again, ION presented the highest differentiation ability; for ballooning and inflammation, AUC = 0.839 and AUC = 0.856, respectively. Moreover, APRI presented the highest differentiation ability regarding fibrosis F2–4 vs. F0–1, yet the AUC did not exceed 0.800 (AUC = 0.735), as well as for F3–4 vs. F0–2 being close to 0.850 (AUC = 0.845), followed by – the recommended in guidelines – FIB-4 (AUC = 0.820).

We validated for the first time using liver biopsies, the differentiation abilities of FLI and LAP in relation to fibrosis F2–4 vs. F0–1, yet these were very low (i.e., AUC < 0.510). Additionally, ION and HSI presented the most promising differentiation ability in relation to the presence or absence of histological characteristics of MASH (all AUCs > 0.800). The α-HSI we suggested herein did not show significant differences from the standard HSI, while no additive value compared with other indexes was observed. Furthermore, the TyGO and the TyG indices presented equal AUCs for all comparisons.

3.2.1. Results from sensitivity analyses

Sensitivity analyses were performed examining the differentiation ability of the examined NITs separately in the sample recruited in the Gastroenterology-Hepatology Departments compared with the sample recruited in the Bariatric-Metabolic Surgery Department, and results are summarized in Table 3. Differences from the observations above were seen in terms of differentiating fibrosis F2–4 vs. F0–1; in particular, LFS seemed to reach an AUC = 0.860 for patients visiting Gastroenterology-Hepatology Departments – with more significant impaired liver function tests – while ION had the best performance for the participants in the Bariatric-Metabolic Surgery Department (AUC = 0.664). Additionally, in the evaluation of hepatocellular ballooning and lobular inflammation, LFS showed the best diagnostic performance in the Bariatric-Metabolic Surgery Department (AUC = 0.894 and 0.873, respectively), which were very close to the AUCs obtained with ION. No other major differences were revealed.

Sensitivity analyses limited only to participants without T2DM or excluding participants with morbid obesity (BMI > 40 cm/kg^2) were performed and are shown in Table 4. The only difference noted from the analyses within the complete study sample was in the differentiation of lobular inflammation when excluding patients with T2DM. More specifically, LFS showed the highest AUC (0.848) but was very close to ION AUC (0.825).

3.3. The accuracy of the standard cut-off points in the examined NITs, and their maximization

The overall accuracy of the standard cut-off points of the examined NITs was assessed, and results are summarized in Tables 5–7.

Starting with the differentiation of MASLD patients vs. Controls or the presence of MASH among patients with MASLD, results are provided in Table 5. Using the standard cut-off points, FLI has the highest accuracy (0.90) for the comparison MASLD vs. Controls (Sensitivity: 0.99; PPV: 91 %). This index was followed by LFS, LAP, and HSI with accuracy scores >0.86; in this case, most of the suggested cut-off points seemed to be useful for ruling-out participants at a sensitivity > 0.92, yet the specificities were very low.

	HSI	α-HSI	ALT/AST	FLI	LAP	TyGO	TyG	ION	FIB-4	APRI	NFS	LFS
MASLD vs. Controls	0.795 (0.732-0.857)	0.709 (0.633-0.785)	0.722 (0.663-0.780)	0.701 (0.610-0.793)	0.736 (0.643-0.829)	0.814 (0.766-0.862)	0.814 (0.766-0.862)	**0.894 (0.854-0.934)**	0.645 (0.580-0.710)	0.678 (0.615-0.742)	0.764 (0.698-0.829)	0.871 (0.798-0.944)
MASH vs. MASL	0.566 (0.502-0.630)	0.504 (0.440-0.569)	0.606 (0.544-0.668)	0.613 (0.542-0.685)	0.609 (0.539-0.679)	0.630 (0.570-0.690)	0.630 (0.570-0.690)	**0.747 (0.682-0.812)**	0.566 (0.504-0.627)	0.645 (0.586-0.705)	0.527 (0.464-0.591)	0.717 (0.643-0.792)
F2-4 vs. F0-1	0.547 (0.491-0.603)	0.500 (0.444-0.557)	0.605 (0.551-0.659)	0.456 (0.391-0.521)	0.509 (0.445-0.572)	0.666 (0.615-0.717)	0.666 (0.615-0.717)	0.696 (0.638-0.754)	0.672 (0.621-0.724)	**0.735 (0.687-0.782)**	0.640 (0.586-0.694)	0.686 (0.619-0.753)
F3-4 vs. F0-2	0.393 (0.331-0.456)	0.397 (0.332-0.463)	0.523 (0.453-0.593)	0.382 (0.307-0.456)	0.421 (0.352-0.491)	0.635 (0.572-0.698)	0.635 (0.572-0.698)	0.636 (0.572-0.701)	0.820 (0.769-0.872)	**0.845 (0.802-0.887)**	0.666 (0.600-0.733)	0.653 (0.580-0.726)
Ballooning, yes/no	0.819 (0.764-0.874)	0.780 (0.719-0.842)	0.645 (0.582-0.708)	0.768 (0.684-0.853)	0.782 (0.707-0.858)	0.749 (0.694-0.804)	0.749 (0.694-0.804)	**0.839 (0.788-0.891)**	0.581 (0.515-0.647)	0.572 (0.507-0.636)	0.748 (0.686-0.810)	0.806 (0.722-0.890)
Inflammation, yes/no	0.827 (0.768-0.885)	0.784 (0.715-0.852)	0.674 (0.610-0.737)	0.748 (0.657-0.840)	0.734 (0.624-0.843)	0.750 (0.689-0.811)	0.750 (0.689-0.811)	**0.856 (0.807-0.905)**	0.577 (0.506-0.649)	0.567 (0.496-0.637)	0.764 (0.697-0.830)	0.789 (0.680-0.898)

Fig. 1. The differentiation ability of 13 non-invasive tools of MASLD vs. Controls, MASH vs. MASL, presence vs. absence of histological liver features, severe (F3–4) vs. mild/moderate (F0–2) liver fibrosis and moderate/severe (F2–4) vs. mild liver fibrosis (F0–1).
Bold and white indicates the index that reached the highest AUC per outcome.
Abbreviations: Aspartate aminotransferase to platelet ratio index (APRI); Hepatic Steatosis Index (HSI); Fatty Liver Index (FLI); Fibrosis-4 (FIB-4); Index of Non-alcoholic steatohepatitis (ION); Lipid Accumulation Product (LAP); Liver Fat Score (LFS); Metabolic-dysfunction associated steatotic liver disease (MASLD); Metabolic-dysfunction associated steatohepatitis (MASH); Non-alcoholic fatty liver disease fibrosis score (NFS); Original Triglycerides-Glucose index (TyGO); Triglycerides-Glucose index (TyG).

Table 3
The differentiation ability of 13 non-invasive tools of MASLD vs. Controls, MASH vs. MASL, presence vs. absence of histological liver features, severe (F3–4) vs. mild/moderate (F0–2) liver fibrosis and moderate/severe (F2–4) vs. mild liver fibrosis (F0–1), according to the study center ($n = 455$).

Outcome to discriminate	Study center	HSI	α-HSI	AST/ALT	ALT/AST	FLI	LAP	TyGO	TyG	ION	FIB-4	APRI	NFS	LFS
MASLD vs. Controls	Gastroenterology-Hepatology Department	0.588 (0.446, 0.731)	0.600 (0.438, 0.762)	0.475 (0.327, 0.623)	0.525 (0.377, 0.673)	0.694 (0.522, 0.866)	0.803 (0.672, 0.933)	0.751 (0.630, 0.873)	0.751 (0.630, 0.873)	**0.879 (0.796, 0.962)**	0.693 (0.543, 0.843)	0.668 (0.528, 0.808)	0.735 (0.581, 0.888)	0.858 (0.718, 0.999)
	Bariatric-Metabolic Surgery Department	0.864 (0.804, 0.925)	0.753 (0.671, 0.835)	0.187 (0.138, 0.236)	0.813 (0.764, 0.862)	0.735 (0.636, 0.834)	0.712 (0.596, 0.828)	0.833 (0.782, 0.884)	0.833 (0.782, 0.884)	**0.904 (0.860, 0.949)**	0.598 (0.518, 0.678)	0.661 (0.587, 0.736)	0.779 (0.706, 0.852)	0.893 (0.834, 0.952)
MASH vs. MASL	Gastroenterology-Hepatology Department	0.571 (0.453, 0.690)	0.557 (0.439, 0.675)	0.460 (0.350, 0.571)	0.540 (0.429, 0.650)	0.609 (0.478, 0.740)	0.626 (0.504, 0.747)	0.666 (0.566, 0.765)	0.666 (0.566, 0.765)	**0.788 (0.673, 0.903)**	0.670 (0.565, 0.774)	0.669 (0.560, 0.778)	0.598 (0.483, 0.713)	0.711 (0.577, 0.846)
	Bariatric-Metabolic Surgery Department	0.625 (0.549, 0.701)	0.515 (0.436, 0.593)	0.356 (0.283, 0.430)	0.644 (0.570, 0.717)	0.657 (0.571, 0.743)	0.622 (0.538, 0.706)	0.607 (0.531, 0.682)	0.607 (0.531, 0.682)	**0.735 (0.656, 0.813)**	0.505 (0.428, 0.583)	0.639 (0.567, 0.712)	0.494 (0.417, 0.570)	0.715 (0.625, 0.806)
Fibrosis stage 2–4 vs. stage 0–1	Gastroenterology-Hepatology Department	0.626 (0.506, 0.746)	0.553 (0.427, 0.678)	0.371 (0.255, 0.486)	0.629 (0.514, 0.745)	0.671 (0.527, 0.814)	0.714 (0.599, 0.829)	0.701 (0.603, 0.798)	0.701 (0.603, 0.798)	0.851 (0.768, 0.935)	0.666 (0.558, 0.775)	0.740 (0.643, 0.837)	0.705 (0.580, 0.830)	**0.860 (0.753, 0.967)**
	Bariatric-Metabolic Surgery Department	0.649 (0.586, 0.712)	0.607 (0.542, 0.673)	0.397 (0.328, 0.466)	0.603 (0.534, 0.672)	0.550 (0.470, 0.630)	0.552 (0.474, 0.630)	0.656 (0.594, 0.719)	0.656 (0.594, 0.719)	**0.664 (0.588, 0.739)**	0.584 (0.516, 0.652)	0.646 (0.581, 0.712)	0.639 (0.574, 0.704)	0.636 (0.548, 0.724)
Fibrosis stage 3–4 vs. stage 0–2	Gastroenterology-Hepatology Department	0.528 (0.426, 0.630)	0.548 (0.448, 0.647)	0.508 (0.409, 0.607)	0.492 (0.393, 0.591)	0.571 (0.458, 0.684)	0.586 (0.479, 0.692)	0.598 (0.503, 0.693)	0.598 (0.503, 0.693)	0.696 (0.590, 0.802)	0.720 (0.633, 0.807)	**0.721 (0.635, 0.806)**	0.690 (0.596, 0.784)	0.674 (0.559, 0.788)
	Bariatric-Metabolic Surgery Department	0.662 (0.552, 0.773)	0.646 (0.510, 0.783)	0.456 (0.298, 0.613)	0.544 (0.387, 0.702)	0.659 (0.522, 0.795)	0.596 (0.470, 0.722)	0.719 (0.614, 0.824)	0.719 (0.614, 0.824)	0.690 (0.535, 0.845)	0.769 (0.661, 0.878)	**0.793 (0.703, 0.882)**	0.773 (0.646, 0.900)	0.760 (0.641, 0.880)
Hepatocellular ballooning, yes/no	Gastroenterology-Hepatology Department	0.630 (0.525, 0.736)	0.669 (0.563, 0.775)	0.508 (0.401, 0.615)	0.492 (0.385, 0.599)	0.698 (0.583, 0.812)	0.728 (0.621, 0.835)	0.668 (0.573, 0.764)	0.668 (0.573, 0.764)	**0.861 (0.778, 0.944)**	0.682 (0.580, 0.784)	0.648 (0.545, 0.751)	0.695 (0.586, 0.803)	0.829 (0.721, 0.937)
	Bariatric-Metabolic Surgery Department	0.870 (0.805, 0.936)	0.786 (0.704, 0.869)	0.213 (0.157, 0.268)	0.787 (0.732, 0.843)	0.689 (0.511, 0.867)	0.739 (0.603, 0.875)	0.827 (0.771, 0.882)	0.827 (0.771, 0.882)	0.866 (0.804, 0.928)	0.599 (0.515, 0.682)	0.634 (0.555, 0.714)	0.790 (0.718, 0.861)	**0.894 (0.826, 0.961)**
Lobular inflammation, yes/no	Gastroenterology-Hepatology Department	0.626 (0.510, 0.742)	0.639 (0.516, 0.762)	0.461 (0.345, 0.577)	0.539 (0.423, 0.655)	0.664 (0.526, 0.801)	0.712 (0.583, 0.841)	0.699 (0.586, 0.813)	0.699 (0.586, 0.813)	**0.872 (0.792, 0.952)**	0.620 (0.494, 0.747)	0.609 (0.477, 0.741)	0.664 (0.540, 0.787)	0.808 (0.673, 0.944)
	Bariatric-Metabolic Surgery Department	**0.879 (0.816, 0.943)**	0.812 (0.728, 0.895)	0.217 (0.159, 0.276)	0.783 (0.724, 0.841)	0.691 (0.581, 0.801)	0.579 (0.342, 0.817)	0.781 (0.712, 0.851)	0.781 (0.712, 0.851)	0.863 (0.801, 0.924)	0.587 (0.501, 0.674)	0.566 (0.483, 0.650)	0.822 (0.750, 0.895)	0.873 (0.755, 0.992)

Abbreviations: Aspartate aminotransferase to platelet ratio index (APRI); Hepatic Steatosis Index (HSI); False negative (FN); False Positive (FP); Fatty Liver Index (FLI); Fibrosis-4 (FIB-4); Index of Non-alcoholic steatohepatitis (ION); Lipid Accumulation Product (LAP); Liver Fat Score (LFS); Metabolic-dysfunction associated liver disease (MASLD); Metabolic-dysfunction associated steatohepatitis (MASH); Non-alcoholic fatty liver disease Fibrosis Score (NFS); Negative Predictive Value (NPV); Positive Predictive Value; (PPV); Original Triglycerides-Glucose index (TyGO); Triglycerides-Glucose index (TyG).
Bold indicates the index that reached the highest AUC per outcome.

Table 4
Sensitivity analyses excluding all patients with the diagnosis of type 2 diabetes mellitus and morbid obesity (BMI > 40 cm/kg^2).

Outcome to discriminate	Sensitivity Analyses	HSI	α-HSI	AST/ALT	ALT/AST	FLI	LAP	TyGO	TyG	ION	FIB-4	APRI	NFS	LFS
MASLD vs. Controls	Type 2 diabetes excluded	0.788 (0.721, 0.856)	0.708 (0.628, 0.789)	0.279 (0.211, 0.347)	0.721 (0.653, 0.789)	0.722 (0.621, 0.822)	0.749 (0.661, 0.836)	0.777 (0.720, 0.834)	0.777 (0.720, 0.834)	**0.881 (0.833, 0.929)**	0.614 (0.540, 0.688)	0.664 (0.594, 0.734)	0.705 (0.628, 0.783)	0.854 (0.785, 0.923)
	Morbid obesity excluded	0.854 (0.796, 0.911)	0.755 (0.662, 0.848)	0.277 (0.207, 0.346)	0.723 (0.654, 0.793)	0.701 (0.565, 0.837)	0.744 (0.616, 0.873)	0.824 (0.767, 0.881)	0.824 (0.767, 0.881)	**0.920 (0.868, 0.971)**	0.770 (0.703, 0.837)	0.758 (0.689, 0.826)	0.824 (0.763, 0.885)	0.861 (0.750, 0.971)
MASH vs. MASL	Type 2 diabetes excluded	0.551 (0.464, 0.638)	0.501 (0.413, 0.589)	0.423 (0.336, 0.509)	0.577 (0.491, 0.664)	0.599 (0.499, 0.699)	0.583 (0.483, 0.682)	0.594 (0.510, 0.678)	0.594 (0.510, 0.678)	**0.748 (0.659, 0.836)**	0.594 (0.509, 0.680)	0.626 (0.543, 0.709)	0.526 (0.437, 0.614)	0.744 (0.637, 0.850)
	Morbid obesity excluded	0.629 (0.546, 0.713)	0.513 (0.421, 0.605)	0.379 (0.295, 0.464)	0.621 (0.536, 0.705)	0.658 (0.570, 0.747)	0.650 (0.559, 0.740)	0.667 (0.589, 0.746)	0.667 (0.589, 0.746)	**0.774 (0.683, 0.866)**	0.599 (0.515, 0.684)	0.681 (0.597, 0.764)	0.517 (0.429, 0.606)	0.723 (0.619, 0.827)
Fibrosis stage 2–4 vs. stage 0–1	Type 2 diabetes excluded	0.577 (0.506, 0.648)	0.520 (0.447, 0.593)	0.375 (0.301, 0.448)	0.625 (0.552, 0.699)	0.445 (0.356, 0.535)	0.493 (0.405, 0.581)	0.623 (0.553, 0.693)	0.623 (0.553, 0.693)	0.713 (0.638, 0.788)	0.654 (0.581, 0.727)	**0.740 (0.675, 0.805)**	0.595 (0.523, 0.668)	0.683 (0.591, 0.776)
	Morbid obesity excluded	0.677 (0.607, 0.746)	0.570 (0.494, 0.647)	0.331 (0.262, 0.401)	0.669 (0.599, 0.738)	0.546 (0.455, 0.637)	0.567 (0.478, 0.657)	0.730 (0.666, 0.793)	0.730 (0.666, 0.793)	0.769 (0.694, 0.843)	0.751 (0.688, 0.813)	**0.802 (0.746, 0.857)**	0.734 (0.668, 0.800)	0.709 (0.610, 0.808)
Fibrosis stage 3–4 vs. stage 0–2	Type 2 diabetes excluded	0.348 (0.262, 0.433)	0.325 (0.228, 0.422)	0.399 (0.284, 0.514)	0.601 (0.486, 0.716)	0.304 (0.193, 0.416)	0.334 (0.229, 0.440)	0.559 (0.452, 0.666)	0.559 (0.452, 0.666)	0.722 (0.645, 0.799)	0.845 (0.753, 0.937)	**0.909 (0.858, 0.960)**	0.578 (0.457, 0.699)	0.703 (0.608, 0.798)
	Morbid obesity excluded	0.524 (0.449, 0.600)	0.499 (0.421, 0.577)	0.444 (0.362, 0.527)	0.556 (0.473, 0.638)	0.499 (0.402, 0.596)	0.483 (0.397, 0.569)	0.652 (0.579, 0.726)	0.652 (0.579, 0.726)	0.708 (0.631, 0.784)	0.837 (0.780, 0.895)	**0.863 (0.815, 0.911)**	0.735 (0.662, 0.809)	0.667 (0.578, 0.756)
Hepatocellular ballooning, yes/no	Type 2 diabetes excluded	0.826 (0.766, 0.885)	0.791 (0.722, 0.860)	0.360 (0.287, 0.433)	0.640 (0.567, 0.713)	0.791 (0.694, 0.888)	0.811 (0.732, 0.890)	0.733 (0.669, 0.797)	0.733 (0.669, 0.797)	**0.835 (0.776, 0.895)**	0.581 (0.506, 0.656)	0.567 (0.492, 0.642)	0.745 (0.673, 0.817)	0.821 (0.743, 0.898)
	Morbid obesity excluded	0.845 (0.790, 0.900)	0.804 (0.735, 0.874)	0.360 (0.286, 0.435)	0.640 (0.565, 0.714)	0.758 (0.666, 0.850)	0.755 (0.659, 0.850)	0.753 (0.688, 0.819)	0.753 (0.688, 0.819)	**0.859 (0.797, 0.921)**	0.692 (0.619, 0.765)	0.655 (0.581, 0.728)	0.775 (0.706, 0.844)	0.786 (0.684, 0.889)
Lobular inflammation, yes/no	Type 2 diabetes excluded	0.817 (0.750, 0.885)	0.788 (0.713, 0.863)	0.343 (0.269, 0.417)	0.657 (0.583, 0.731)	0.762 (0.647, 0.876)	0.766 (0.651, 0.881)	0.733 (0.666, 0.800)	0.733 (0.666, 0.800)	0.825 (0.762, 0.887)	0.581 (0.502, 0.660)	0.562 (0.485, 0.639)	0.756 (0.681, 0.831)	**0.848 (0.758, 0.937)**
	Morbid obesity excluded	0.862 (0.805, 0.919)	0.826 (0.752, 0.900)	0.342 (0.270, 0.415)	0.658 (0.585, 0.730)	0.668 (0.536, 0.799)	0.683 (0.548, 0.819)	0.758 (0.686, 0.830)	0.758 (0.686, 0.830)	**0.885 (0.827, 0.942)**	0.680 (0.600, 0.760)	0.662 (0.580, 0.743)	0.776 (0.704, 0.848)	0.776 (0.642, 0.910)

Abbreviations: Alanine transaminase (AST), Aspartate aminotransferase (AST), Aspartate aminotransferase to platelet ratio index (APRI); Body mass index (BMI), Hepatic Steatosis Index (HSI); False negative (FN); False Positive (FP); Fatty Liver Index (FLI); Fibrosis-4 (FIB-4); Index of Non-alcoholic steatohepatitis (ION); Lipid Accumulation Product (LAP); Liver Fat Score (LFS); Metabolic-dysfunction associated liver disease (MASLD); Metabolic-dysfunction associated steatohepatitis (MASH); Non-alcoholic fatty liver disease Fibrosis Score (NFS); Negative Predictive Value (NPV); Positive Predictive Value; (PPV); Original Triglycerides-Glucose index (TyGO); Triglycerides-Glucose index (TyG); Type 2 diabetes mellitus (T2DM).
Bold indicates the index that reached the highest AUC per outcome.

Table 5
Performance of NITs to differentiate the presence vs. absence of MASLD or the stage of MASLD (MASH or NAFL) within the NAFLD patients.

		Sensitivity	Specificity	PPV,%	NPV,%	Accuracy		Sensitivity	Specificity	PPV,%	NPV,%	Accuracy
	NITs	Outcome to differentiate: MASLD vs. controls N = 455|Cases = 374					NITs	Outcome to differentiate: MASH vs. MASL N = 374|Cases = 237				
Steatosis-related	HSI						HSI					
NITs	<30	1.00	0.05	83	100	0.83	<30	1.00	0.00	64	0	0.64
	≥36	0.96	0.44	89	70	0.86	≥36	0.96	0.05	64	43	0.63
	FLI						FLI					
	<30	0.99	0.10	91	50	0.90	<30	0.99	0.02	67	67	0.67
	≥60	0.96	0.23	92	37	0.88	≥60	0.97	0.07	68	50	0.67
	LAP						LAP					
	≥38.050	0.92	0.30	92	30	0.86	≥38.050	0.96	0.15	69	65	0.69
	≥33.400	0.94	0.21	91	29	0.87	≥33.400	0.96	0.10	68	59	0.68
	LFS						LFS					
	<−0.640	0.92	0.46	94	39	0.88	<−0.640	0.97	0.18	71	76	0.71
	<−1.413	0.96	0.29	93	44	0.89	<−1.413	0.98	0.08	69	67	0.69
	≥1.257	0.70	0.96	99	25	0.72	≥1.257	0.79	0.51	77	54	0.70
Fibrosis-related	FIB-4						FIB-4					
NITs	<1.300	0.31	0.92	95	23	0.42	<1.300	0.35	0.75	70	41	0.50
	≥2.670	0.07	0.97	92	19	0.23	≥2.670	0.09	0.98	86	39	0.42
	APRI						APRI					
	<0.5	0.22	0.95	95	21	0.35	<0.5	0.28	0.88	79	42	0.50
	≥1.0	0.05	0.97	89	18	0.21	≥1.0	0.05	0.96	69	38	0.39
	NFS						NFS					
	<−1.453	0.80	0.64	91	42	0.77	<−1.453	0.80	0.19	64	35	0.58
	≥0.675	0.17	0.93	92	20	0.31	≥0.675	0.20	0.87	74	38	0.44
Other MASLD-related	TyGO						TyGO					
NITs	≥8.500	0.98	0.05	83	36	0.81	≥8.500	0.99	0.04	63	71	0.63
	TyG						TyG					
	≥4.680	0.70	0.86	96	38	0.73	≥4.680	0.76	0.42	69	51	0.63
	ION						ION					
	≥26	0.82	0.76	94	47	0.81	≥26	0.93	0.40	74	77	0.75
	≥50	0.45	0.98	99	28	0.55	≥50	0.59	0.79	84	51	0.66

Higher values in indexes indicate positive cases. Sensitivity (Recall) = TP/(TP + FN); Specificity = TN/(TN + FP); PPV (precision) = TP/(TP + FP); NPV = TN/(TN + FN).

Abbreviations: Aspartate aminotransferase to platelet ratio index (APRI); Hepatic Steatosis Index (HSI); False negative (FN); False Positive (FP); Fatty Liver Index (FLI); Fibrosis-4 (FIB-4); Index of Non-alcoholic steatohepatitis (ION); Lipid Accumulation Product (LAP); Liver Fat Score (LFS); Metabolic-dysfunction associated liver disease (MASLD); Metabolic-dysfunction associated steatohepatitis (MASH); Non-alcoholic fatty liver disease Fibrosis Score (NFS); Negative Predictive Value (NPV); Positive Predictive Value; (PPV); Original Triglycerides-Glucose index (TyGO); Triglycerides-Glucose index (TyG); True negative (TN); True positive (TP).

Regarding the differentiation of MASH among MASLD patients, all the standard cut-off points presented accuracy scores of <0.76. To provide a hierarchical approach, we show that ION with a cut-off point of 26, followed by LFS with a cut-off point of −0.64, presented the highest level of accuracy (0.75 and 0.71, respectively). Thus, a participant could be ruled out for MASH with a sensitivity of 0.93 and a PPV of 74 % with ION, while with LFS with a sensitivity of 0.97 and a PPV of 71 %.

Moving on to the fibrosis classification, differentiation from severe fibrosis or mild/moderate fibrosis was examined, and the results are summarized in Table 6. We observed that APRI, followed by FIB-4 had the highest accuracies (0.82), with specificities >0.90 at cut-off points of 0.5 for APRI and 2.670 for FIB-4 (NPV > 80 %). However, their sensitivities were < 0.55. On the other hand, to differentiate moderate/severe fibrosis (F2–4), LFS presented the highest accuracy (0.69) at the cut-off point of 1.257, with a sensitivity of 0.75 and a specificity of 0.53. Regarding the analyses performed for the first time, examining FLI and LAP against fibrosis classification, their standard cut-off points presented accuracies between 0.70 and 0.75 for the comparison F3–4 vs. F0–2, with specificities >0.90 yet very low sensitivity.

The standard cut-off points of the examined NITs were examined against the presence vs. absence of histological characteristics of MASLD, i.e., hepatocellular ballooning and lobular inflammation. Results are presented in Table 7. Steatosis indexes, i.e., HSI, LAP, and FLI were examined for the first time in relation to these liver features. The same was for the FIB-4. FLI showed the highest accuracy of 0.88, yet this was mostly attributed to a very high sensitivity with very low specificity, implying its potential use in terms of ruling out patients. FLI was followed by LFS and LAP with accuracies >0.85. Contrarily, FIB-4 did not seem to work well against these liver features, with the standard cut-off points reaching accuracies of <0.45.

Estimation of cut-off points for the indexes that do not have a standardized clinically used cut-off point, i.e., AST/ALT, ASL/AST, as well as exploring cut-off points for all the NITs by maximizing their specificity or sensitivity, were performed, and results are summarized in ***Supplementary Tables 2 and 3***. In particular, we proposed for the first time, one rule in and one rule out cut-off point (i.e., maximizing specificity and sensitivity, respectively, at 0.70) for α-HSI, AST/ALT, and ALT/AST; however, these did not seem to be sensitive nor specific enough for MASLD screening or for differentiating specific disease stages and histological characteristics. A similar analysis was performed for the rest NITs, with the best constricted combination seen in the case of ION for MASLD vs. Controls differentiation, with a rule out cut-off point of 19.540 (Sensitivity = 0.894; Specificity = 0.709); while LFS had the best constricted combination for MASLD vs. Controls with a rule in cut-off point of 1.210 (Sensitivity = 0.709; Specificity = 0.958). For the suggested α-HSI, the best-constricted combinations were seen for differentiation for ballooning and inflammation; rule out cut-off point: 43.950 (Sensitivity = 0.735; Specificity = 0.712) and rule in cut-off point: 44.740 (Sensitivity = 0.703; Specificity = 0.726).

4. Discussion

In the context of a multi-center analysis using liver biopsies as the diagnostic gold standard, we validated available and new NITs in terms of their ability to define the presence of MASLD, MASH, and histological

Table 6
Performance of NITs to differentiate liver fibrosis stages.

		Sensitivity	Specificity	PPV, %	NPV, %	Accuracy		Sensitivity	Specificity	PPV, %	NPV, %	Accuracy
	NITs	Outcome to differentiate: Fibrosis F3–4 vs. F0–2 N = 455 | Cases = 109					NITs	Outcome to differentiate: Fibrosis F2–4 vs. F0–1 N = 455 | Cases = 243				
Steatosis-related	HSI						HSI					
NITs	<30	0.00	0.99	0	77	0.77	<30	1.00	0.02	52	100	0.53
	≥36	0.10	0.88	19	77	0.70	≥36	0.94	0.18	55	74	0.57
	FLI						FLI					
	<30	0.01	0.98	17	75	0.74	<30	0.01	0.97	33	44	0.43
	≥60	0.08	0.94	32	76	0.73	≥60	0.06	0.93	53	44	0.44
	LAP						LAP					
	≥38.050	0.12	0.91	30	75	0.70	≥38.050	0.92	0.13	58	55	0.58
	≥33.400	0.08	0.93	29	74	0.71	≥33.400	0.94	0.09	58	54	0.58
	LFS						LFS					
	<−0.640	0.97	0.14	27	93	0.35	<−0.640	0.96	0.21	62	79	0.64
	<−1.413	0.98	0.08	26	94	0.31	<−1.413	0.97	0.11	60	75	0.61
	≥1.257	0.84	0.44	33	89	0.54	≥1.257	0.75	0.53	69	62	0.66
Fibrosis-related	FIB-4						FIB-4					
NITs	<1.300	0.64	0.83	52	89	0.79	<1.300	0.41	0.88	77	58	0.64
	≥2.670	0.22	0.99	83	82	0.82	≥2.670	0.10	0.99	92	51	0.54
	APRI						APRI					
	<0.5	0.53	0.91	61	87	0.82	<0.5	0.32	0.95	88	56	0.62
	≥1.0	0.16	0.99	83	81	0.81	≥1.0	0.07	0.99	89	50	0.51
	NFS						NFS					
	<−1.453	0.86	0.32	25	89	0.43	<−1.453	0.82	0.38	58	66	0.60
	≥0.675	0.29	0.88	40	82	0.76	≥0.675	0.19	0.89	65	51	0.53
Other MASLD-related	TyGO						TyGO					
NITs	≥8.500	0.99	0.03	23	91	0.25	≥8.500	0.98	0.03	53	64	0.53
	TyG						TyG					
	≥4.680	0.78	0.45	30	87	0.53	≥4.680	0.71	0.53	63	62	0.63
	ION						ION					
	≥26	0.94	0.35	28	96	0.47	≥26	0.84	0.42	60	72	0.64
	≥50	0.50	0.66	28	83	0.62	≥50	0.51	0.76	68	60	0.63

Higher values in indexes indicate positive cases. Sensitivity (Recall) = TP/(TP + FN); Specificity = TN/(TN + FP); PPV (precision) = TP/(TP + FP); NPV = TN/(TN + FN).

Abbreviations: Aspartate aminotransferase to platelet ratio index (APRI); Hepatic Steatosis Index (HSI); False negative (FN); False Positive (FP); Fatty Liver Index (FLI); Fibrosis-4 (FIB-4); Index of Non-alcoholic steatohepatitis (ION); Lipid Accumulation Product (LAP); Liver Fat Score (LFS); Metabolic-dysfunction associated liver disease (MASLD); Non-alcoholic fatty liver disease Fibrosis Score (NFS); Negative Predictive Value (NPV); Positive Predictive Value; (PPV); Original Triglycerides-Glucose index (TyGO); Triglycerides-Glucose index (TyG); True negative (TN); True positive (TP).

characteristics of MASH as well as fibrosis stages. Currently, identifying MASH and fibrosis F ≥ 2 through non-invasive methods remains an unmet clinical need, especially for treatment initiation and patient inclusion in clinical trials [24]. Here, we observed that the available NITs are inadequate to correctly classify participants that had been diagnosed with liver fibrosis F ≥ 2 through liver biopsy. However, APRI seemed to have the best differentiation ability (AUC = 0.845), even after excluding patients with morbid obesity (AUC = 0.863), and T2DM (AUC = 0.909). The recommended by guidelines FIB-4 seemed to differentiate severe fibrosis (i.e., F3–4 vs. F0–2) with an AUC = 0.820, yet this was lower when used to classify patients with fibrosis F ≥ 2. Interestingly, among the examined NITs, ION presented the higher differentiation ability for the presence vs. absence of MASLD (AUC = 0.894) as well as MASH vs. MASL (AUC = 0.747), and histological characteristics of MASH (hepatocellular ballooning and lobular inflammation) with AUCs = 0.839 and 0.856, respectively. Finally, suggesting two cut-off points to rule in and rule out patients seemed to work for the indexes that presented the highest overall performance (AUC > 0.800); yet still, the accuracies of the suggested cut-off points compared with the standard ones were not improved in a meaningful way.

Most patients with MASLD are seen in primary care or endocrine clinics. Initial screening of the general population for the presence of MASLD is rather important to prevent disease progression [25]. Additionally, the detection of MASH is crucial since this condition can progress to cirrhosis and severe liver and systemic complications [26]. In our study, ION had the best ability to discriminate MASLD and MASH, as well as the presence of hepatocellular ballooning and lobular inflammation, highlighting its performance not only as a screening test for MASLD, but also for identifying MASH histological characteristics. Of note, all the AUC for differentiating MASH patients from the ones with NAFL, were <0.80. Although ION was not initially proposed as a screening test, it was proposed as a tool to distinguish MASH within the MASLD population with a cut-off of 50 (92 % specificity, 60 % sensitivity, AUC = 0.880) [27]. However, it could be a not useful tool for MASH diagnosis in subjects without obesity (AUC = 0.687), and its predictive properties did not improve even when combined with fibrosis markers [28]. Furthermore, the potential of other indices to distinguish the presence of MASH from simple steatosis has been previously examined with less successful results, such as FLI [27] and AST/ALT [29], as replicated in our analysis.

Hepatic fibrosis is also a crucial determinant of liver and non-liver outcomes in patients with MASLD. Therefore, identifying patients with clinically significant hepatic fibrosis (F ≥ 2) is important for targeted efforts at preventing disease progression. A recent study found that screening for MASLD followed by intensive lifestyle interventions or pioglitazone was cost-effective in patients with T2DM diagnosed with fibrosis F ≥ 2 [30,31]. Most Phase III clinical trials in MASH also target patients with fibrosis F ≥ 2 with the potential to be translated to treatment options. Thus, this makes subjects with F ≥ 2 an important group to identify [31,32]. In our study, APRI was the best NIT in identifying F ≥ 2 (AUC = 0.845); however, for identifying F ≥ 1, none of the evaluated NITs reached an AUC > 0.80. In accordance with our results, APRI has shown an ability to detect F3–4 from mild fibrosis stages (AUC = 0.923) and seems to be an appropriate substitute for FibroScan for

Table 7
Performance of NITs to differentiate the presence vs. absence of histological characteristics of NASH.

		Sensitivity	Specificity	PPV,%	NPV,%	Accuracy		Sensitivity	Specificity	PPV,%	NPV,%	Accuracy
	NITs	Outcome to differentiate: Hepatocellular ballooning, Yes vs. no N = 455 | Cases = 352					NITs	Outcome to differentiate: Lobular inflammation, Yes vs. no N = 455 | Cases = 373				
Steatosis-related NITs	HSI						HSI					
	<30	1.00	0.04	78	100	0.79	<30	1.00	0.05	83	100	0.83
	≥36	0.97	0.41	85	81	0.85	≥36	0.96	0.47	89	72	0.87
	FLI						FLI					
	<30	0.99	0.10	88	67	0.88	<30	0.99	0.11	92	50	0.91
	≥60	0.96	0.23	90	47	0.87	≥60	0.95	0.22	93	32	0.89
	LAP						LAP					
	≥38.050	0.94	0.34	90	45	0.86	≥38.050	0.93	0.38	94	33	0.88
	≥33.400	0.95	0.25	89	46	0.86	≥33.400	0.94	0.24	93	29	0.88
	LFS						LFS					
	<−0.640	0.93	0.44	92	50	0.87	<−0.640	0.92	0.50	95	36	0.89
	<−1.413	0.96	0.22	89	44	0.86	<−1.413	0.96	0.30	94	38	0.90
	≥1.257	0.70	0.84	97	30	0.72	≥1.257	0.67	0.85	98	19	0.69
Fibrosis-related NITs	FIB-4						FIB-4					
	<1.300	0.29	0.82	85	25	0.41	<1.300	0.30	0.85	90	21	0.39
	≥2.670	0.07	0.97	88	23	0.27	≥2.670	0.06	0.96	88	18	0.22
	APRI						APRI					
	<0.5	0.21	0.89	86	25	0.37	<0.5	0.20	0.86	86	19	0.32
	≥1.0	0.04	0.95	72	23	0.25	≥1.0	0.04	0.95	78	18	0.20
	NFS						NFS					
	<−1.453	0.81	0.60	88	47	0.76	<−1.453	0.80	0.65	91	41	0.77
	≥0.675	0.18	0.94	92	25	0.35	≥0.675	0.18	0.96	95	20	0.32
Other MASLD-related NITs	TyGO						TyGO					
	≥8.500	0.98	0.04	78	36	0.77	≥8.500	0.98	0.05	82	36	0.81
	TyG						TyG					
	≥4.680	0.70	0.74	90	42	0.71	≥4.680	0.68	0.78	93	36	0.70
	ION						ION					
	≥26	0.82	0.69	91	49	0.79	≥26	0.81	0.75	94	45	0.80
	≥50	0.46	0.94	97	31	0.56	≥50	0.45	0.96	98	26	0.54

Higher values in indexes indicate positive cases. Sensitivity (Recall) = TP/(TP + FN); Specificity = TN/(TN + FP); PPV (precision) = TP/(TP + FP); NPV = TN/(TN + FN).

Abbreviations: Aspartate aminotransferase to platelet ratio index (APRI); Hepatic Steatosis Index (HSI); False negative (FN); False Positive (FP); Fatty Liver Index (FLI); Fibrosis-4 (FIB-4); Index of Non-alcoholic steatohepatitis (ION); Lipid Accumulation Product (LAP); Liver Fat Score (LFS); Metabolic-dysfunction associated liver disease (MASLD); Non-alcoholic fatty liver disease Fibrosis Score (NFS); Non-invasive tool (NIT); Negative Predictive Value (NPV); Positive Predictive Value; (PPV); Original Triglycerides-Glucose index (TyGO); Triglycerides-Glucose index (TyG); True negative (TN); True positive (TP).

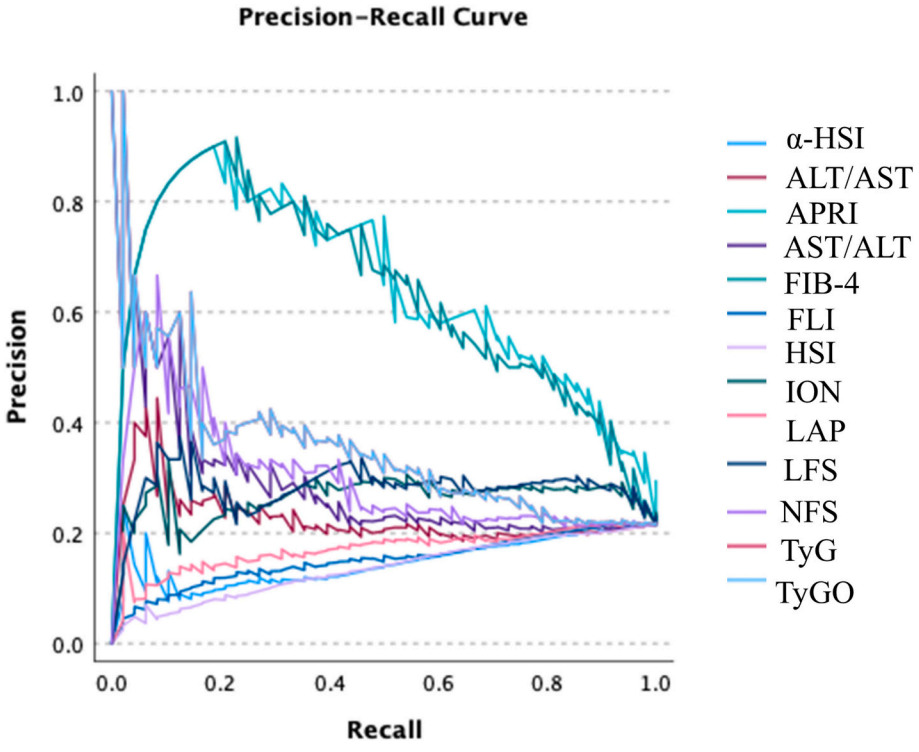

Fig. 2. Precision-recall curve of 13 non-invasive tools for patients' classification according to liver fibrosis grade i.e. fibrosis F2–4 vs. F0–1.
Abbreviations: Aspartate aminotransferase to platelet ratio index (APRI); Hepatic Steatosis Index (HSI); Fatty Liver Index (FLI); Fibrosis-4 (FIB-4); Index of Non-alcoholic steatohepatitis (ION); Lipid Accumulation Product (LAP); Liver Fat Score (LFS); Non-alcoholic fatty liver disease Fibrosis Score (NFS); Original Triglycerides-Glucose index (TyGO); Triglycerides-Glucose index (TyG).

severe fibrosis detection over FIB-4 ability [33]. Similarly, among the other liver fibrosis-related indexes, the diagnostic performance of the -guidelines recommended- FIB-4 index [25] was close to APRI's ability for differentiating severe vs. mild/moderate fibrosis [33] (precision-recall curve is provided for all NITs against fibrosis $F \geq 2$ in Fig. 2), yet this was not the case in terms of differentiating against fibrosis $F \geq 2$.

Our results are in line with other studies that have reported considerable percentages of false-negative results oriented by FIB-4 standard cut-off points [34,35], potentially leading to underdiagnosis [36,37]. In contrast to our results, other analyses have shown that FIB-4 is superior to APRI as a predictor of advanced fibrosis with a validated AUC of 0.802 [38]. Moreover, a cross-sectional study reported that NFS was able to detect any significant fibrosis stage, as well as advanced fibrosis, with AUC of 0.700, and 0.761, respectively, over APRI and FIB-4 [39,40]; yet this was not confirmed herein. Considering that the samples from the present work belong to a high-risk population seen in secondary care clinics underlines the low specificity of available liver fibrosis NITs to identify the fibrosis stage, which remains an unmet need [24].

Furthermore, we observed improvement in the diagnostic ability of LFS to differentiate fibrosis when evaluating only patients from the Gastroenterology-Hepatology Departments (i.e., MASLD patients with chronically altered liver function tests). On the other hand, LFS diagnostic performance also improved in terms of the evaluation of hepatocellular ballooning and lobular inflammation when including only patients from the Bariatric-Metabolic Surgery Department. Even though in our study LFS had not the best diagnostic ability for any evaluation within the complete population, it was among the highest ones. LFS has not been suggested as an index that can predict the specific stages of the disease; [15] however, here we report that it is helpful in correctly identifying lobular inflammation/hepatocellular ballooning, which comes in line with previous results [18]. LFS was initially developed in a study with 470 Finnish individuals with MASLD, where the liver fat content of proton magnetic resonance spectroscopy was used as the reference method [15]. This study predicted increased liver fat content using the cut-off point of −0.640 with a sensitivity of 0.86 and a specificity of 0.71.

Finally, we studied non-specific MASLD-related indices initially developed to determine the presence of other health conditions, such as insulin resistance, as is the case of the TyGO index [41]. The TyGO index was first proposed by Simental-Mendía, et al. as an alternative to the homeostasis model assessment of insulin resistance (HOMA-IR) index as a screening tool for insulin resistance [41]. However, in the following manuscripts, the TyGO formula presented a -probably inadvertent- modification that has been implemented in some of the manuscripts, thus creating differences in their cut-off values and confusion in the literature [23,41–44]. Although we calculated and used both formulas, similar diagnostic performance was observed in all comparisons. Thus, when using this index, it should be first confirmed that the correct formula and the respective cut-off values are being considered.

4.1. Limitations and strengths

This work has several limitations that need to be reported for a better interpretation of the outcomes. First, the samples of the present study were primarily collected retrospectively (Study 2 and Study 3); however, both the liver biopsies and serum samples had been collected without the investigators being aware of the aim of this study, and thus, there was no bias introduced in the analysis. Second, participants were recruited from secondary care centers, which may justify the low performance of almost all examined NITs that have not been developed to be that sensitive or specific for administration in a high-risk population. Additionally, this limits the generalization of the findings. However, the results herein allow us to report how sensitive or specific NITs can be in the context of a specialized hepatology gastroenterology or bariatric metabolic surgery clinic.

The above limitations are compensated by several strengths. This is a multi-center study with diverse populations, including patients eligible for bariatric surgery as well as patients with clinical evidence that implied the presence of MASLD, MASH and/or advanced fibrosis, which necessitated the referral of subjects to the Gastroenterology-Hepatology Department. A liver biopsy was also performed on participants assigned as Controls herein, which is scarce in the relevant literature. Additionally, this study is one of only a few that have validated a large set of NITs in relation to any liver health-related outcome within the MASLD spectrum using the gold standard of liver biopsy and the only study that fully validated newer indexes that had not been validated to date.

4.2. Conclusions

The increasing prevalence of MASLD, combined with the slow, asymptomatic disease progression which underlies its underdiagnosis until the disease has advanced significantly may result in significant pressures on national healthcare systems. Effective screening and differentiation of specific disease stages, histological features, and, most importantly, fibrosis grading remain an unmet clinical need. Despite the promising results we report herein, the overall performance of the examined NITs remains suboptimal for clinical implementation. Advancing the existing NITs or suggesting new ones to address the currently unmet clinical needs is necessary. In this context, the use of novel molecules or multi-omics markers should be considered.

Data sharing

Any other data that support the findings of this study may become available from the corresponding author upon request. Source data are provided with this paper.

Author contributions

CSM generated the research hypothesis and supervised the study design, analysis, and interpretation of the outcomes. LVV, VGG, and MaK wrote the manuscript and interpreted the results (equal contribution). MA and YA analyzed the data with MaK, CSM, AK, LVV, VGG, SM, and MiK contributions. GM supervised the data selection from the Bariatric-Metabolic Surgery Department in Italy, JG supervised the data selection from the Gastroenterology-Hepatology Department in Australia, and GP supervised the data selection from the Gastroenterology-Hepatology Department in Greece. OV, ME, and GM contributed to the data selection per center, respectively. All authors critically reviewed the manuscript.

CRediT authorship contribution statement

Matina Kouvari: Data curation, Investigation, Formal analysis, Writing - original draft. **Laura Valenzuela-Vallejo:** Data curation, Investigation, Writing - original draft. **Valentina Guatibonza-Garcia:** Data curation, Investigation, Writing - original draft. **Stergios A. Polyzos:** Investigation, Writing - review & editing. **Yixiang Deng:** Formal analysis, Writing - review & editing. **Michail Kokkorakis:** Investigation, Writing – review & editing. **Melih Agraz:** Investigation, Writing – review & editing. **Sophia C. Mylonakis:** Investigation, Writing – review & editing. **Angeliki Katsarou:** Investigation, Writing – review & editing. **Ornella Verrastro:** Data curation, Writing - review & editing. **Georgios Markakis:** Data curation, Investigation, Writing - review & editing. **Mohammed Eslam:** Data curation, Investigation, Writing - review & editing. **Georgios Papatheodoridis:** Supervision, Writing - review & editing. **Jacob George:** Supervision, Writing – review & editing. **Geltrude Mingrone:** Supervision, Writing - review & editing. **Christos S. Mantzoros:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

CSM reports grants through his institution from Merck, Massachusetts Life Sciences Center, and Boehringer Ingelheim, has been a shareholder of and has received grants through his institution and personal consulting fees from Coherus Inc., and AltrixBio; he reports personal consulting fees from Novo Nordisk, reports personal consulting fees and collaborative research from Ansh Inc., collaborative research support from LabCorp Inc., reports personal consulting fees from Genfit, Lumos, Amgen, Corcept, Intercept, 89 Bio, Madrigal, and Regeneron, reports educational activity meals through his institution or national conferences from Esperion, Merck, Boehringer Ingelheim and travel support and fees from TMIOA, Elsevier, and the Cardio Metabolic Health Conference. None is related to the work presented herein. GP has received fees for advisory board meetings and lectures from Abbvie, Albireo, Amgen, Dicerna, Gilead, GlaxoSmithKline, Ipsen, Janssen, Merck Sharp & Dohme, Novo Nordisk, Roche, Takeda and has received research grants from Abbvie and Gilead. All other authors have no competing interest to declare.

Acknowledgments

Supported in part by the Bits to Bytes grants for scientific projects, Massachusetts Life Sciences Center, and a research grant from the Investigators Studies Research Program of Merck Sharp & Dohme Corp, and a Scientific Advancement Grant from Boehringer Ingelheim. The opinions expressed in this paper are those of the investigators and do not necessarily represent those of Merck Sharp & Dohme Corp.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.metabol.2023.155666.

References

[1] Younossi Z, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. Nat Rev Gastroenterol Hepatol 2018;15:11–20. https://doi.org/10.1038/nrgastro.2017.109.

[2] Le MH, Yeo YH, Zou B, Barnet S, Henry L, Cheung R, et al. Forecasted 2040 global prevalence of nonalcoholic fatty liver disease using hierarchical bayesian approach. Clin Mol Hepatol 2022;28:841–50. https://doi.org/10.3350/cmh.2022.0239.

[3] Godoy-Matos AF, Silva Júnior WS, Valerio CM. NAFLD as a continuum: from obesity to metabolic syndrome and diabetes. Diabetol Metab Syndr 2020;12:60. https://doi.org/10.1186/s13098-020-00570-y.

[4] Younossi ZM, Yilmaz Y, Yu M-L, Wai-Sun Wong V, Fernandez MC, Isakov VA, et al. Clinical and patient-reported outcomes from patients with nonalcoholic fatty liver disease across the world: data from the global non-alcoholic steatohepatitis (NASH)/non-alcoholic fatty liver disease (NAFLD) registry. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc 2022;20:2296–2306.e6. https://doi.org/10.1016/j.cgh.2021.11.004.

[5] Chalasani N, Younossi Z, Lavine JE, Charlton M, Cusi K, Rinella M, 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. https://doi.org/10.1002/hep.29367.

[6] Estes C, Razavi H, Loomba R, Younossi Z, Sanyal AJ. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. Hepatol Baltim Md 2018;67:123–33. https://doi.org/10.1002/hep.29466.

[7] Dokmak A, Lizaola-Mayo B, Trivedi HD. The impact of nonalcoholic fatty liver disease in primary care: a population health perspective. Am J Med 2021;134:23–9. https://doi.org/10.1016/j.amjmed.2020.08.010.

[8] Siddiqui MS, Yamada G, Vuppalanchi R, Van Natta M, Loomba R, Guy C, et al. Diagnostic accuracy of noninvasive fibrosis models to detect change in fibrosis stage. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc 2019;17:1877–1885.e5. https://doi.org/10.1016/j.cgh.2018.12.031.

[9] Long MT, Gandhi S, Loomba R. Advances in non-invasive biomarkers for the diagnosis and monitoring of non-alcoholic fatty liver disease. Metabolism 2020;111S:154259. https://doi.org/10.1016/j.metabol.2020.154259.

[10] Serra-Burriel M, Graupera I, Torán P, Thiele M, Roulot D, Wai-Sun Wong V, et al. Transient elastography for screening of liver fibrosis: cost-effectiveness analysis from six prospective cohorts in Europe and Asia. J Hepatol 2019;71:1141–51. https://doi.org/10.1016/j.jhep.2019.08.019.

[11] Perakakis N, Stefanakis K, Mantzoros CS. The role of omics in the pathophysiology, diagnosis and treatment of non-alcoholic fatty liver disease. Metabolism 2020;111S:154320. https://doi.org/10.1016/j.metabol.2020.154320.

[12] Davison BA, Harrison SA, Cotter G, Alkhouri N, Sanyal A, Edwards C, et al. Suboptimal reliability of liver biopsy evaluation has implications for randomized clinical trials. J Hepatol 2020;73:1322–32. https://doi.org/10.1016/j.jhep.2020.06.025.

[13] Castera L, Friedrich-Rust M, Loomba R. Noninvasive assessment of liver disease in patients with nonalcoholic fatty liver disease. Gastroenterology 2019;156:1264–1281.e4. https://doi.org/10.1053/j.gastro.2018.12.036.

[14] Zhang S, Du T, Zhang J, Lu H, Lin X, Xie J, et al. The triglyceride and glucose index (TyG) is an effective biomarker to identify nonalcoholic fatty liver disease. Lipids Health Dis 2017;16:15. https://doi.org/10.1186/s12944-017-0409-6.

[15] Kotronen A, Peltonen M, Hakkarainen A, Sevastianova K, Bergholm R, Johansson LM, et al. Prediction of non-alcoholic fatty liver disease and liver fat using metabolic and genetic factors. Gastroenterology 2009;137:865–72. https://doi.org/10.1053/j.gastro.2009.06.005.

[16] Reddy YK, Marella HK, Jiang Y, Ganguli S, Snell P, Podila PSB, et al. Natural history of non-alcoholic fatty liver disease: a study with paired liver biopsies. J Clin Exp Hepatol 2020;10:245–54. https://doi.org/10.1016/j.jceh.2019.07.002.

[17] Bedogni G, Kahn HS, Bellentani S, Tiribelli C. A simple index of lipid overaccumulation is a good marker of liver steatosis. BMC Gastroenterol 2010;10:98. https://doi.org/10.1186/1471-230X-10-98.

[18] Fedchuk L, Nascimbeni F, Pais R, Charlotte F, Housset C, Ratziu V, et al. Performance and limitations of steatosis biomarkers in patients with nonalcoholic fatty liver disease. Aliment Pharmacol Ther 2014;40:1209–22. https://doi.org/10.1111/apt.12963.

[19] Kim JW, Lee CH, Kim B-H, Lee Y-S, Hwang S-Y, Park BN, et al. Ultrasonographic index for the diagnosis of non-alcoholic steatohepatitis in patients with non-alcoholic fatty liver disease. Quant Imaging Med Surg 2022;12:1815–29. https://doi.org/10.21037/qims-21-895.

[20] Verrastro O, Panunzi S, Castagneto-Gissey L, Gaetano AD, Lembo E, Capristo E, et al. Bariatric–metabolic surgery versus lifestyle intervention plus best medical care in non-alcoholic steatohepatitis (BRAVES): a multicentre, open-label, randomised trial. Lancet 2023;401:1786–97. https://doi.org/10.1016/S0140-6736(23)00634-7.

[21] Juluri R, Vuppalanchi R, Olson J, Ünalp A, Van Natta ML, Cummings OW, et al. Generalizability of the NASH CRN histological scoring system for nonalcoholic fatty liver disease. J Clin Gastroenterol 2011;45:55–8. https://doi.org/10.1097/MCG.0b013e3181dd1348.

[22] Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatol Baltim Md 2005;41:1313–21. https://doi.org/10.1002/hep.20701.

[23] Alizargar J, Hsieh N-C, Wu S-FV. The correct formula to calculate triglyceride-glucose index (TyG). J Pediatr Endocrinol Metab JPEM 2020;33:945–6. https://doi.org/10.1515/jpem-2019-0579.

[24] Kanwal F, Shubrook JH, Younossi Z, Natarajan Y, Bugianesi E, Rinella ME, et al. Preparing for the NASH epidemic: a call to action. Gastroenterology 2021;161:1030–1042.e8. https://doi.org/10.1053/j.gastro.2021.04.074.

[25] Kanwal F, Shubrook JH, Adams LA, Pfotenhauer K, Wong VW-S, Wright E, et al. Clinical care pathway for the risk stratification and management of patients with nonalcoholic fatty liver disease. Gastroenterology 2021;161:1657–69. https://doi.org/10.1053/j.gastro.2021.07.049.

[26] Harrison SA, Ratziu V, Boursier J, Francque S, Bedossa P, Majd Z, et al. A blood-based biomarker panel (NIS4) for non-invasive diagnosis of non-alcoholic steatohepatitis and liver fibrosis: a prospective derivation and global validation study. Lancet Gastroenterol Hepatol 2020;5:970–85. https://doi.org/10.1016/S2468-1253(20)30252-1.

[27] Otgonsuren M, Estep MJ, Hossain N, Younossi E, Frost S, Henry L, et al. Single non-invasive model to diagnose non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH). J Gastroenterol Hepatol 2014;29:2006–13. https://doi.org/10.1111/jgh.12665.

[28] Younes R, Rosso C, Petta S, Cucco M, Marietti M, Caviglia GP, 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 Off J Int Assoc Study Liver 2018;38:715–23. https://doi.org/10.1111/liv.13612.

[29] Sorbi D, Boynton J, Lindor KD. The ratio of aspartate aminotransferase to alanine aminotransferase: potential value in differentiating nonalcoholic steatohepatitis from alcoholic liver disease. Am J Gastroenterol 1999;94:1018–22. https://doi.org/10.1111/j.1572-0241.1999.01006.x.

[30] Anstee QM, Targher G, Day CP. Progression of NAFLD to diabetes mellitus, cardiovascular disease or cirrhosis. Nat Rev Gastroenterol Hepatol 2013;10:330–44. https://doi.org/10.1038/nrgastro.2013.41.

[31] Dulai PS, Singh S, Patel J, Soni M, Prokop LJ, Younossi Z, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: systematic review and meta-analysis. Hepatol Baltim Md 2017;65:1557–65. https://doi.org/10.1002/hep.29085.

[32] Taylor RS, Taylor RJ, Bayliss S, Hagström H, Nasr P, Schattenberg JM, et al. Association between fibrosis stage and outcomes of patients with nonalcoholic fatty liver disease: a systematic review and meta-analysis. Gastroenterology 2020;158:1611–1625.e12. https://doi.org/10.1053/j.gastro.2020.01.043.

[33] Amernia B, Moosavy SH, Banookh F, Zoghi G. FIB-4, APRI, and AST/ALT ratio compared to FibroScan for the assessment of hepatic fibrosis in patients with non-alcoholic fatty liver disease in Bandar Abbas, Iran. BMC Gastroenterol 2021;21:453. https://doi.org/10.1186/s12876-021-02038-3.

[34] Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: clinical prediction rules and blood-based biomarkers. J Hepatol 2018;68:305–15. https://doi.org/10.1016/j.jhep.2017.11.013.
[35] Rigor J, Diegues A, Presa J, Barata P, Martins-Mendes D. Noninvasive fibrosis tools in NAFLD: validation of APRI, BARD, FIB-4, NAFLD fibrosis score, and Hepamet fibrosis score in a Portuguese population. Postgrad Med 2022;134:435–40. https://doi.org/10.1080/00325481.2022.2058285.
[36] Graupera I, Thiele M, Serra-Burriel M, Caballeria L, Roulot D, Wong GL-H, et al. Low accuracy of FIB-4 and NAFLD fibrosis scores for screening for liver fibrosis in the population. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc 2022;20:2567–76 (e6), https://doi.org/10.1016/j.cgh.2021.12.034.
[37] Viganò M, Pugliese N, Cerini F, Turati F, Cimino V, Ridolfo S, et al. Accuracy of FIB-4 to detect elevated liver stiffness measurements in patients with non-alcoholic fatty liver disease: a cross-sectional study in referral centers. Int J Mol Sci 2022;23:12489. https://doi.org/10.3390/ijms232012489.
[38] Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ, et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc 2009;7:1104–12. https://doi.org/10.1016/j.cgh.2009.05.033.
[39] Bernstein D, Kovalic AJ. Noninvasive assessment of fibrosis among patients with nonalcoholic fatty liver disease [NAFLD]. Metab Open 2022;13:100158. https://doi.org/10.1016/j.metop.2021.100158.
[40] Younes R, Caviglia GP, Govaere O, Rosso C, Armandi A, Sanavia T, et al. Long-term outcomes and predictive ability of non-invasive scoring systems in patients with non-alcoholic fatty liver disease. J Hepatol 2021;75:786–94. https://doi.org/10.1016/j.jhep.2021.05.008.
[41] Simental-Mendía LE, Rodríguez-Morán M, Guerrero-Romero F. The product of fasting glucose and triglycerides as surrogate for identifying insulin resistance in apparently healthy subjects. Metab Syndr Relat Disord 2008;6:299–304. https://doi.org/10.1089/met.2008.0034.
[42] Guerrero-Romero F, Villalobos-Molina R, Jiménez-Flores JR, Simental-Mendia LE, Méndez-Cruz R, Murguía-Romero M, et al. Fasting triglycerides and glucose index as a diagnostic test for insulin resistance in young adults. Arch Med Res 2016;47:382–7. https://doi.org/10.1016/j.arcmed.2016.08.012.
[43] Locateli JC, Lopes WA, Simões CF, de Oliveira GH, Oltramari K, Bim RH, et al. Triglyceride/glucose index is a reliable alternative marker for insulin resistance in South American overweight and obese children and adolescents. J Pediatr Endocrinol Metab JPEM 2019;32:1163–70. https://doi.org/10.1515/jpem-2019-0037.
[44] Guerrero-Romero F, Simental-Mendía LE, González-Ortiz M, Martínez-Abundis E, Ramos-Zavala MG, Hernández-González SO, et al. The product of triglycerides and glucose, a simple measure of insulin sensitivity. Comparison with the euglycemic-hyperinsulinemic clamp. J Clin Endocrinol Metab 2010;95:3347–51. https://doi.org/10.1210/jc.2010-0288.
[45] Rinella ME, Lazarus JV, Ratziu V, et al. A multi-society Delphi consensus statement on new fatty liver disease nomenclature [published online ahead of print, 2023 Jun 24] Hepatology 2023. https://doi.org/10.1097/HEP.0000000000000520.