Clinical Gastroenterology and Hepatology 2016;14:1463–1472

Noninvasive Detection of Nonalcoholic Steatohepatitis Using Clinical Markers and Circulating Levels of Lipids and Metabolites

You Zhou,[*,‡] Matej Orešič,[§] Marja Leivonen,[‖] Peddinti Gopalacharyulu,[¶] Jenni Hyysalo,[*,#] Johanna Arola,[**] An Verrijken,[‡‡] Sven Francque,[§§] Luc Van Gaal,[‡‡] Tuulia Hyötyläinen,[§] and Hannele Yki-Järvinen[*,#]

*Minerva Foundation Institute for Medical Research, Helsinki, Finland; [‡]Systems Immunity University Research Institute and Division of Infection and Immunity, School of Medicine, Cardiff University, Cardiff, United Kingdom; [§]Steno Diabetes Center, Gentofte, Denmark; [‖]Department of Surgery, **Department of Pathology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland; [¶]Institute for Molecular Medicine Finland, [#]Department of Medicine, University of Helsinki, Helsinki, Finland; [‡‡]Department of Endocrinology, Diabetology and Metabolism, [§§]Department of Gastroenterology and Hepatology, Antwerp University Hospital, University of Antwerp, Antwerp, Belgium

BACKGROUND & AIMS: **Use of targeted mass spectrometry (MS)-based methods is increasing in clinical chemistry laboratories. We investigate whether MS-based profiling of plasma improves noninvasive risk estimates of nonalcoholic steatohepatitis (NASH) compared with routinely available clinical parameters and patatin-like phospholipase domain-containing protein 3 (*PNPLA3*) genotype at rs738409.**

METHODS: **We used MS-based analytic platforms to measure levels of lipids and metabolites in blood samples from 318 subjects who underwent a liver biopsy because of suspected NASH. The subjects were divided randomly into estimation (n = 223) and validation (n = 95) groups to build and validate the model. Gibbs sampling and stepwise logistic regression, which fulfilled the Bayesian information criterion, were used for variable selection and modeling.**

RESULTS: **Features of the metabolic syndrome and the variant in *PNPLA3* encoding I148M were significantly more common among subjects with than without NASH. We developed a model to identify subjects with NASH based on clinical data and *PNPLA3* genotype (NASH Clin Score), which included aspartate aminotransferase (AST), fasting insulin, and *PNPLA3* genotype. This model identified subjects with NASH with an area under the receiver operating characteristic of 0.778 (95% confidence interval, 0.709–0.846). We then used backward stepwise logistic regression analyses of variables from the NASH Clin Score and MS-based factors associated with NASH to develop the NASH ClinLipMet Score. This included glutamate, isoleucine, glycine, lysophosphatidylcholine 16:0, phosphoethanolamine 40:6, AST, and fasting insulin, along with *PNPLA3* genotype. It identified patients with NASH with an area under the receiver operating characteristic of 0.866 (95% confidence interval, 0.820–0.913). The NASH ClinLipMet score identified patients with NASH with significantly higher accuracy than the NASH Clin Score or MS-based profiling alone.**

CONCLUSIONS: **A score based on MS (glutamate, isoleucine, glycine, lysophosphatidylcholine 16:0, phosphoethanolamine 40:6) and knowledge of AST, fasting insulin, and *PNPLA3* genotype is significantly better than a score based on clinical or metabolic profiles alone in determining the risk of NASH.**

Keywords: Liver; Nonalcoholic Fatty Liver Disease; Diagnosis; Prediction; Triglycerides.

Abbreviations used in this paper: **AST, aspartate aminotransferase; AUROC, area under the receiver operating characteristic curve; CI, confidence interval; fP, fasting plasma; fS, fasting serum; GC, gas chromatography; Glu, glutamate; Gly, glycine; HDL, high-density lipoprotein; Ile, isoleucine; IR, insulin resistance; LC, lipid cluster; Lip, lipidomics; Leu, leucine; LysoPC, lysophosphatidylcholines; MET, metabolomic; MetS, metabolic syndrome; MS, mass spectrometry; NAFL, nonalcoholic fatty liver; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; PE, phosphatidylethanolamine; PNPLA3, patatin-like phospholipase domain-containing protein 3; Ser, serine; TG, triacylglycerol; TOFMS, time-of-flight mass spectrometry; Tyr, tyrosine; UPLC, ultra-performance liquid chromatography; Val, valine.**

Most current article

1542-3565/$36.00
http://dx.doi.org/10.1016/j.cgh.2016.05.046

The spectrum of nonalcoholic fatty liver disease (NAFLD) ranges from the nonalcoholic fatty liver (NAFL) to nonalcoholic steatohepatitis (NASH), which increases the risk of cirrhosis and mortality from liver disease.[1] Only a fraction of patients with NAFLD progress to NASH, of which the diagnosis requires a liver biopsy. Thus, there is a need to develop noninvasive tools to identify patients, who might be at risk of having NASH.

Factors such as age, sex, liver enzymes, components of the metabolic syndrome (MetS), as well as circulating markers of inflammation, fibrosis, apoptosis, and extracellular matrix components, have been shown to be associated with NASH.[1] Genetic factors, especially the I148M variant in patatin-like phospholipase domain-containing protein 3 (PNPLA3), also confers susceptibility to NASH.[1]

Ultra-performance liquid chromatography mass spectrometry (UPLC-MS)-based techniques are rapidly entering clinical chemistry laboratories and replacing many conventional techniques.[2] Few data are available regarding such approaches to identify new markers for noninvasive estimation of the risk of NASH. In 24 obese patients with NASH, 11 with steatosis, and 25 lean controls, who did not undergo a liver biopsy, plasma concentrations of glutamate (Glu), isoleucine (Ile), leucine (Leu), tyrosine (Tyr), and valine (Val) were increased significantly, but diagnostic performance compared with routinely available markers was not examined.[3] Barr et al[4] characterized lipids and metabolites in serum of 467 Caucasian patients and found between 9 and 237 metabolites to be markers of NASH, depending on the degree of obesity.

The human liver lipidome differs markedly between subjects with NAFLD associated with insulin resistance (IR NAFLD) compared with those without, and between subjects with NAFLD and the PNPLA3 I148M genotype (PNPLA3 NAFLD) compared with those lacking the gene variant.[5] The liver lipidome is enriched markedly with saturated and monounsaturated triglycerides and free fatty acids in IR NAFLD, and with polyunsaturated triglycerides in PNPLA3 NAFLD.[5] These differences also influence the circulating lipidome in NAFLD in subjects whose liver fat content has been measured using proton magnetic resonance spectroscopy but who have not undergone a liver biopsy.[6] However, there are no studies in patients who have undergone a liver biopsy that would have determined whether knowledge of the PNPLA3 genotype influences biomarkers of NASH as compared with non-NASH (NAFL or normal liver histology) subjects. Furthermore, to our knowledge, no study has analyzed whether MS-based markers significantly improve predictive performance of scores based on routinely available physical and biochemical parameters. In the present study, we developed scores based on the following: (1) routinely available clinical parameters and PNPLA3 genotype, (2) UPLC-MS analyses alone, and (3) all available information for estimation of the risk of NASH. The diagnostic performance of the 3 models then was compared.

Materials and Methods

Study Subjects

Metabolic studies were conducted at the University of Helsinki (Finland) and Antwerp University Hospital (Belgium). A total of 318 subjects were recruited among those referred to the Department of Gastroenterology (Finland, n = 54) because of chronically increased serum aminotransferase concentrations and among those referred for bariatric surgery in Belgium (n = 193) and Finland (n = 71). Subjects were eligible if they met the following criteria: (1) age 18 to 75 years; (2) no known acute or chronic disease except for obesity or type 2 diabetes on the basis of medical history, physical examination, and standard laboratory tests (blood counts, serum creatinine, thyroid-stimulating hormone, electrolyte concentrations) and electrocardiogram; and (3) alcohol consumption less than 20 g/d. Hepatitis B surface antigen, transferrin saturation, and antibodies against hepatitis A and C and anti–smooth muscle, antinuclear, and antimitochondrial antibodies were measured in all patients referred to the gastroenterologist because of chronically increased liver function tests using routine methods of local laboratories. Patients were excluded if they used thiazolidinediones or were pregnant. The study protocol was approved by the ethics committees of the Helsinki and the Antwerp University Hospitals. Each participant provided written informed consent.

Metabolic Study

All subjects were invited to a clinical visit 1 week before surgery for metabolic characterization after an overnight fast. After anthropometric measurements (body weight, height, and waist circumference), an intravenous cannula was inserted in an antecubital vein for withdrawal of blood for measurement of HbA_{1C}, glycosylated hemoglobin 1c, serum insulin and adiponectin, plasma glucose, low-density lipoprotein and high-density lipoprotein (HDL) cholesterol, triglyceride, total blood counts, albumin, aspartate aminotransferase (AST), alanine aminotransferase, alkaline phosphatase, γ-glutamyl transpeptidase, and albumin concentrations, and for genotyping of *PNPLA3* at rs738409 as described.[7] Blood sampling was performed before intake of any medications.

Histologic Assessment

Immediately at the beginning of the surgery, wedge biopsy specimens of the liver were obtained. The biopsy specimens from Belgium were sent to Finland, where

they were assessed simultaneously with the Finnish samples by an experienced liver pathologist in a blinded fashion according to the criteria proposed by Brunt et al.[8] Liver fat was quantified as the percentage of hepatocytes with macrovesicular steatosis.

Mass Spectrometry–Based Profiling

Lipidomic analysis. An unthawed plasma sample was used from all subjects and extracted for lipidomic and metabolomic analysis (see later). An established platform based on acquity UPLC quadruple time-of-flight mass spectrometry (Hertfordshire, UK) was used to analyze the plasma samples. The data were processed by using MZmine 2 software (Espoo, Finland)[9] and the lipid identification was based on an internal spectral library (see Supplementary Methods section for more detail).

Metabolomic analysis. Polar metabolites were analyzed using comprehensive 2-dimensional gas chromatography combined with time-of-flight mass spectrometry (GC × GC-TOFMS) (see Supplementary Methods section for more detail).

Other Analytic Procedures and Measurements

Body weight, waist circumference, blood pressure, and fasting concentrations of plasma glucose, serum-free insulin, lipids (HDL and low-density lipoprotein cholesterol and triacylglycerols [TGs]), and liver enzyme (AST, alanine aminotransferase, and γ-glutamyl transpeptidase) concentrations were measured as previously described.[9] The MetS was defined and PNPLA3 at rs739409 was genotyped as described.[9]

Statistical Analyses

Assessment of abundances of triacylglycerol species. After $\log_2$ transformation, the average abundances of TG molecules were compared between the NASH and non-NASH groups by Student *t* tests. Multiple comparisons were corrected by using the Benjamini–Hochberg method[10] (see the Supplementary Methods for more detail).

Cluster analysis of lipidomics data. Lipids were grouped by using Bayesian model-based clustering as previously described.[11]

Diagnostic model. The biopsy subjects were divided randomly into estimation (n = 223) and validation (n = 95) groups to build and validate the model, respectively (see later). All study subjects (n = 318) were used as the second validation group. The Shapiro–Wilk test was used to test the normality of the distribution. The unpaired *t* test or the Wilcoxon rank-sum test was used to compare the differences between the estimation and validation groups. Normally distributed data are shown as means ± SEM and non-normally distributed data are shown as the median followed by the 25th and 75th quartiles. Multiple hypotheses testing was performed by using the Benjamini–Hochberg[10] false-discovery rate method to calculate q-values. In data with a large number of variables measured by UPLC, the Gibbs sampling algorithm was used for simulation.[12] After sampling 10,000 times, the variables were selected among the models based on Bayesian information criterion. Logistic regression including all of the selected variables was used to build the scores. Variables in the scores finally were assessed by backward stepwise regression to identify the optimal NASH score. The area under the receiver operating characteristic curve (AUROC) was used to describe the diagnostic accuracy of the scores. The optimal cut-off point was calculated using the Youden index. The sensitivity, specificity, positive predictive values, and negative predictive values for relevant cut-off values were calculated as described.[6] The AUROCs were compared using the generalized U-statistics.[13] One-way analysis of variance was used to compare 3 groups. The Tukey Honestly Significant Differences test was used for post hoc analyses. A 2-sided *P* value of less than .05 was considered statistically significant. The statistical analyses were performed using R version 3.0.1 (http://www.r-project.org/).

Results

Characteristics of the Study Groups

Comparison of nonalcoholic steatohepatitis and non-nonalcoholic steatohepatitis groups. Characteristics of the NASH and non-NASH groups are shown in Table 1. Liver fat content and all liver enzyme levels were significantly higher in the NASH than in the non-NASH group (Table 1). Features of the MetS (hyperglycemia, hypertriglyceridemia, hypertension, and low HDL cholesterol) were significantly more common in the NASH than in the non-NASH group. The NASH group had a significantly increased prevalence of the PNPLA I148M variant compared with the non-NASH group ($P < .001$). These significances remained significant after adjusting for age (data not shown).

Comparison of estimation and validation groups. The estimation and validation groups were comparable with respect to clinical and biochemical features such as age, sex, body mass index, components of the MetS, prevalence of NASH, liver fat, liver function tests, as well as PNPLA3 genotype (Supplementary Table 1). Within the estimation and validation cohorts (Supplementary Tables 2 and 3), the NASH group showed similar abnormalities compared with the non-NASH group, as was observed in the entire group (Table 1).

Development of a Model to Predict Nonalcoholic Steatohepatitis

We first developed a model based on clinical parameters and the PNPLA3 genotype alone, then models

Table 1. Clinical Characteristics of the Study Subjects According to Liver Fat (%) and a Proposal by Brunt et al[7]

Total	Non-NASH (n = 249)	Non-NAFLD (n = 132)	NAFL (n = 117)	NASH (n = 69)
Liver fat/steatosis, %	5 (0–15)	0 (0–5)	15 (10–30)	40 (30–60)[a,b]
Microscopic steatosis, %	10 (0–30)	0 (0–10)	30 (11–40)	40 (20–40)[a,b]
Grade, 0/1/2/3	249/0/0/0	132/0/0/0	117/0/0/0	0/57/11/1[a]
Ballooning	0	0	0	12[a,c]
Inflammation	0	0	0	69[a]
Fibrosis stage, 0/1/2/3/4	210/31/6/0/2	117/10/3/0/2	93/21/3/0/0	6/47/10/5/1[a]
Fibrosis stage > 0, %	15.7	11.4	20.5	91.3[a]
Age, *y*	45.4 ± 0.8	45.1 ± 1.1	45.9 ± 1.1	49.4 ± 1.3
Sex, n (% women)	160 (64.2)	96 (72.7)	64 (54.7)	37 (53.6)
BMI, kg/m^2	39.6 ± 0.6	39.0 ± 0.8	40.3 ± 0.8	41.1 ± 1.0
Waist circumference, *cm*	118 ± 1	114 ± 1	122 ± 1	126 ± 2[a,b,c]
fS insulin, *mU/L*	12.7 (8.0–18.1)	11.0 (7.5–15.9)	14.3 (9.5–18.9)	20.6 (14.3–28.7)[a,b,d]
fP glucose, *mmol/L*	4.9 (4.4–5.7)	4.8 (4.2–5.6)	5.1 (4.6–5.8)	5.7 (5.0–6.4)[a,b,e]
HbA_{1C}, %	5.6 (5.4–5.9)	5.5 (5.3–5.8)	5.7 (5.5–6.0)	6.0 (5.7–6.6)[a,b,d]
HOMA-IR, *mmol/L* × *mU/L*	2.8 (1.9–4.0)	2.4 (1.5–3.5)	3.2 (2.2–4.4)	5.0 (3.1–8.5)[a,b,e]
fP triglycerides, *mmol/L*	1.35 (1.03–1.92)	1.28 (0.97–1.78)	1.50 (1.04–1.95)	1.76 (1.26–2.54)[a,b,f]
fP HDL cholesterol, *mmol/L*	1.25 (1.03–1.51)	1.27 (1.09–1.51)	1.19 (0.96–1.46)	1.07 (0.94–1.27)[a,b]
fP LDL cholesterol, *mmol/L*	2.90 ± 0.06	2.83 ± 0.08	2.92 ± 0.09	2.99 ± 0.12
P-AST, *IU/L*	28 (24–38)	26 (22–33)	32 (27–42)	42 (29–58)[a,b,e]
P-ALT, *IU/L*	37 (30–51)	34 (27–42)	46 (33–62)	54 (40–89)[a,b,f]
P-GGT, *U/L*	28 (21–48)	27 (21–49)	30 (22–48)	47 (29–73)[a,b]
P-albumin, *g/L*	42.4 ± 0.3	42.1 ± 0.5	42.7 ± 0.5	41.7 ± 0.7
B-platelets, $\times 10^9/L$	259 (219–311)	264 (219–310)	258 (222–310)	240 (197–302)
PNPLA3 (CC/CG/GG), n	133/91/10	79/44/2	54/47/8	25/30/11[g]
Use of lipid-lowering drugs, %	17.7	12.1	23.9	20.3
Type 2 diabetes, n	40	14	26	30***

NOTE. Data are shown as n (%), means ± SEM, or median (25th–75th percentile), as appropriate. The HOMA-IR was calculated as follows: fS-insulin (mU/L) × fP-glucose (mmol/L)/22.5.

ALT, alanine aminotransferase; B, blood; BMI, body mass index; fP, fasting plasma; GGT, γ-glutamyl transpeptidase; HbA_{1C}, glycosylated hemoglobin 1c; HOMA-IR, homeostasis model assessment of insulin resistance; LDL, low-density lipoprotein; P, plasma.

[a]$P < .001$ for comparison with the non-NASH.

[b]$P < .001$ for 1-way analysis of variance.

[c]$P < .05$ for comparison with the non-NASH.

[d]$P < .01$ for the Tukey honestly significant differences test compared with the NAFL.

[e]$P < .001$ for the Tukey honestly significant differences test compared with the NAFL.

[f]$P < .05$ for the Tukey honestly significant differences test compared with the NAFL.

[g]$P < .01$ for comparison with the non-NASH.

based on MS-based profiling, and, finally, a model using all data.

Model based on clinical parameters and I148M variant in PNPLA3: nonalcoholic steatohepatitis clinical score. To build the NASH Clin Score, we used variables differing significantly between NASH and non-NASH groups in univariate analysis in the estimation group (Supplementary Table 2). The model was developed using multivariate logistic regression analysis based on clinical variables and PNPLA3 genotype. The final model included the same predictors as the NASH score, which was developed recently in a group of 296 Finnish patients and validated in a cohort of 380 Italian patients[12] (ie, fasting insulin level, AST level, and PNPLA3 genotype). The NASH Clin Score was calculated as follows: -3.05 + 0.562 × PNPLA3 genotype (CC = 1/GC = 2/GG = 3) - 0.0092 × fasting serum (fS)-insulin level (mU/L) + 0.0023 × AST (IU/L) + 0.0019 × (fS-insulin × AST).

The AUROC for the NASH Clin Score in the entire group (n = 318) was 0.778 (95% confidence interval [CI], 0.709–0.846).

Lipidomics and metabolomics data. By using the 2 MS-based analytic platforms, a total of 597 molecular lipids and metabolites were measured and 168 were identified. Total fS-TG measured enzymatically was correlated closely with the sum of plasma TGs identified by UPLC-MS (r = 0.92; $P < .001$).

Cluster analysis of lipidomics data. We analyzed the global lipidome by clustering the data into a subset of clusters using Bayesian model-based clustering. The lipidomic platform data were decomposed into 8 lipid clusters (LCs), which to a large extent adhered to different lipid functional or structural group. Data on each cluster and representative lipids are shown in Table 2. In the NASH, compared with the non-NASH, group, significant differences were found in 3 lipid clusters (LC3, LC4, and LC6) (Supplementary Figure 1). We found NASH to be associated significantly with increased concentrations of saturated and mono-unsaturated TGs (LC4). In contrast, concentrations of sphingomyelins (LC3) and lysophosphatidylcholines (lysoPC) (LC6) were significantly lower in the NASH

Table 2. Composition of Circulating Lipid Clusters

Cluster name	Number of lipids	Representative members
LC1	18	TG(16:0/18:2/18:1); TG(18:1/16:1/18:2)+TG(18:2/18:2/16:0); TG(18:1/18:2/18:1); TG(18:1/18:1/18:1); TG(54:5)
LC2	24	PC(34:2); PC(36:2); PC(34:1); PC(36:3); PC(38:3)
LC3	32	SM(d18:1/24:1); SM(d18:1/16:0); SM(d18:1/22:0); SM(d18:1/24:0); SM(d18:1/18:0); SM(d18:1/20:0); SM(d18:1/23:0); SM(d18:0/16:0); SM(d18:0/20:4)
LC4	23	TG(14:0/16:0/18:0)+TG(16:0/16:0/16:0); TG(16:0/16:0/18:0); TG(14:0/16:0/16:0)+TG(16:0/18:0/12:0); TG(44:0); TG(16:0/18:0/18:0); TG(44:1); TG(54:1)
LC5	15	PC(38:6); PC(40:6); PC(36:5); PE(40:6); PS(38:1); PS(36:1); PC(38:5); PE(38:5); PE(40:6)
LC6	18	LysoPC(16:0); LysoPC(18:2); LysoPC(18:0); LysoPC(18:1); LysoPC(18:3); LysoPC(20:3)
LC7	14	PC(38:7); PC(40:7); PE(38:4); PE(40:7); PE(40:6); PE(40:4)
LC8	16	PC(34:1e)+PE(37:1e); PC(33:2)+PE(36:2); PC(31:1)+PE(34:1); PC(33:1)+PE(36:1); PC(33:2)+PE(36:2)

PC, phosphatidylcholine; SM, sphingomyelin.

than in the non-NASH group (Supplementary Figure 1, Table 2).

Absolute and relative concentrations of triacylglycerols. Absolute concentrations of circulating TGs between the NASH and non-NASH groups are compared in a heat map (Figure 1, left panel). In the NASH, as compared with the non-NASH, groups, the absolute concentrations of especially saturated and monounsaturated TGs such as TG(46:0), TG(48:0), TG(50:0), TG(46:1), and TG(51:1) were increased significantly.

The relative distribution of TGs (the concentration of an individual TG divided by total TGs measured by UPLC-MS) between the NASH and non-NASH groups is shown in Figure 1 (right panel). The relative concentrations of saturated and monounsaturated TGs were increased in the NASH as compared with the non-NASH group (Figure 1, right). Consistent with an increase in TGs containing saturated and monounsaturated TGs, the fold-changes (NASH/non-NASH) of absolute ($r = -0.75$; $P < .0001$) (Figure 2, left panel) and relative ($r = -0.75$; $P < .0001$) (Figure 2, right panel) concentrations of TGs were correlated inversely with the number of double bonds.

Lipidomics-based model. By using the lipidomics data from the estimation cohort, we derived a logistic regression model (see the Methods section for more detail) for NASH. The final score based on lipidomics data (NASH Lip Score) included 3 selected molecular lipids, TG(48:0), phosphatidylethanolamine (PE)(40:6), and LysoPC(16:0), and was calculated as follows: $2.531 + 2.334 \times \log_{10}(\text{TG[48:0]})$ (μmol/L) $+ 1.555 \times \log_{10}(\text{PE[40:6]})$ (μmol/L) $- 4.081 \times \log_{10}(\text{LysoPC[16:0]})$ (μmol/L).

In the estimation group, the AUROC was 0.767 (95% CI, 0.687–0.847). For the validation group ($n = 95$), the AUROC was 0.809 (95% CI, 0.714–0.905) and in the entire data set ($n = 318$) was 0.779 (95% CI, 0.717-0.841). The negative predictive values, positive predictive values, sensitivity, and specificity of the entire data set are shown in Table 3.

Metabolomics-based model. We also derived a logistic regression model for NASH using the metabolomics data from the estimation cohort. The score based on metabolomics (NASH Met Score) included 5 selected molecular metabolites: Glu, Ile, Tyr, glycine (Gly), and serine (Ser). The NASH Met Score was calculated as follows: $-10.701 + 1.852 \times \log_{10}(\text{Glu})$ (μmol/L) $+ 6.461 \times \log_{10}(\text{Ile})$ (μmol/L) $+ 3.556 \times \log_{10}(\text{Tyr})$ (μmol/L) $- 3.908 \times \log_{10}(\text{Gly})$ (μmol/L) $- 2.822 \times \log_{10}(\text{Ser})$ (μmol/L).

The model had an AUROC of 0.729 (95% CI, 0.649–0.808) in the estimation cohort. In the validation group ($n = 95$), the AUROC was 0.710 (95% CI, 0.604–0.816). In the entire group, the AUROC was 0.719 (95% CI, 0.655–0.782) (Table 3).

Model based on all data. By applying backward stepwise logistic regression analyses of the variables from all of the aforementioned models, we developed the NASH ClinLipMet Score (NASH score based on clinical variables, PNPLA3 genotype, lipidomics, and metabolomics data, corrected for the number of variables included in the model, which was calculated as follows: $-8.167 + 0.954 \times$ PNPLA3 genotype (CC = 1/GC = 2/GG = 3) $+ 0.0451 \times$ AST (IU/L) $+ 0.0667 \times$ fS-insulin (mU/L) $- 3.151 \times \log_{10}(\text{LysoPC[16:0]})$ (μmol/L) $+ 2.617 \times \log_{10}(\text{PE[40:6]})$ (μmol/L) $+ 2.357 \times \log_{10}(\text{Glu})$ (μmol/L) $+ 7.813 \times \log_{10}(\text{Ile})$ (μmol/L) $- 6.102 \times \log_{10}(\text{Gly})$ (μmol/L).

The AUROC was 0.882 (95% CI, 0.827–0.938) in the estimation and 0.856 (95% CI, 0.774–0.938) in the validation cohort. In the entire group, the AUROC was 0.866 (95% CI, 0.820–0.913). The sensitivity was 85.5% and the specificity was 72.1% (Table 3).

Diagnostic Performances in the Subgroups

Because bariatric patients might differ from non-bariatric patients, we excluded 54 patients not undergoing bariatric surgery and measured the performance of all scores in the specific group with bariatric patients. The AUROCs of the NASH Clin Score, NASH Lip Score, and NASH Met Score in the bariatric surgery patients were as follows: 0.774 (95% CI, 0.696–0.852), 0.789 (95% CI, 0.720–0.858), and 0.738 (95% CI, 0.672–0.804),

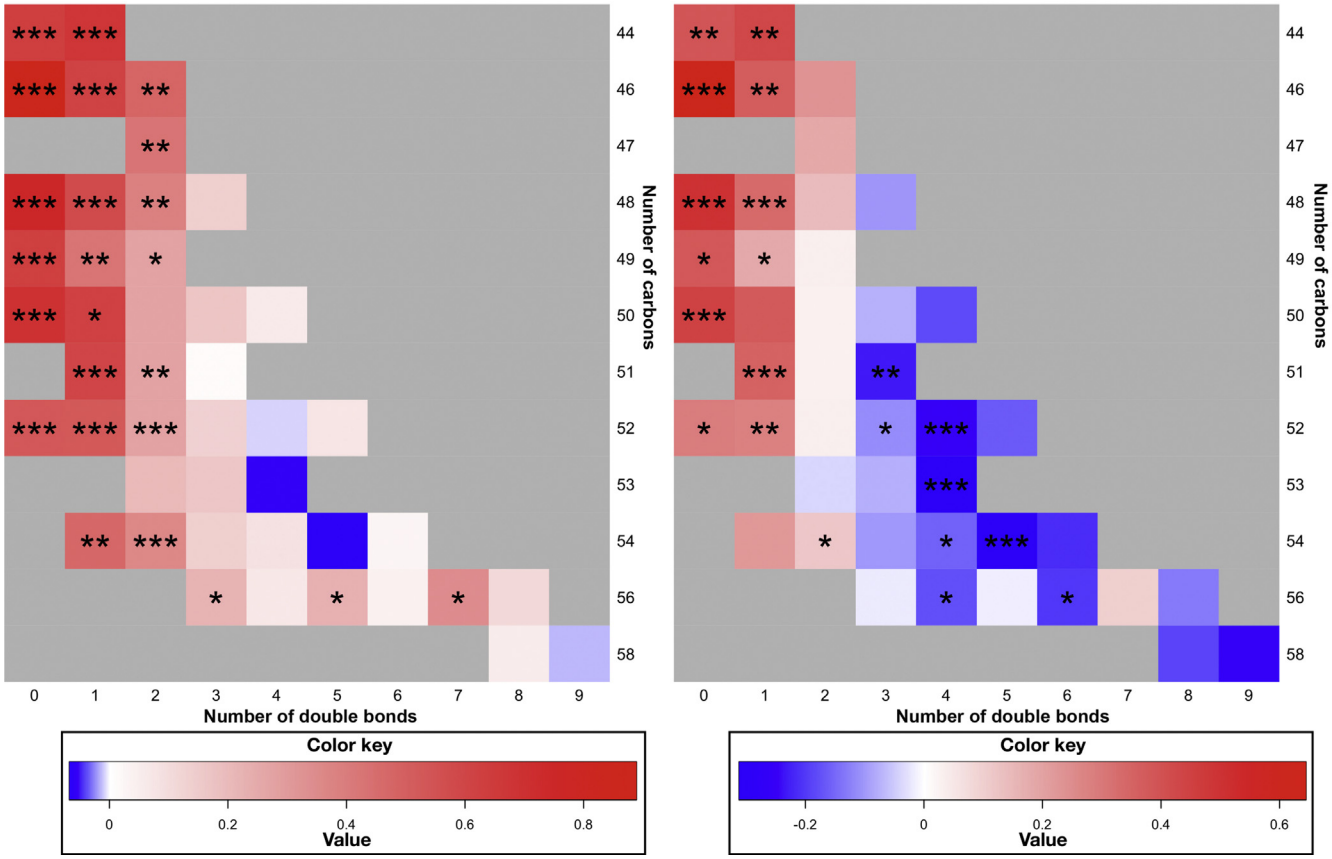


Figure 1. Comparison of concentrations of TGs between NASH and non-NASH groups. The color code denotes the $\log_2$ of the ratio between the means of the groups for an individual TG (*left*: absolute concentrations of TG; *right*: relative concentrations of TG). The y-axes denote the number of carbons, and the x-axes show the number of double bonds. Blue represents a decrease in NASH as compared with non-NASH. Significance for the comparisons are marked as follows: $^{*}P < .05$, $^{**}P < .005$, $^{***}P < .0005$.

respectively. The NASH ClinLipMet Score had an AUROC of 0.865 (95% CI, 0.812–0.918). The AUROCs of the NASH ClinLipMet Score did not differ significantly between the bariatric surgery group and the entire cohort ($P = .961$).

To assess potential confounding effect of lipid-lowering medications, we excluded 58 patients who received lipid medications and re-analyzed the diagnostic performance of all scores. In subjects not using lipid-lowering medications, the AUROCs of the NASH Clin Score, NASH Lip Score, NASH Met Score, and NASH ClinLipMet Score were 0.799 (95% CI, 0.725–0.873), 0.816 (95% CI, 0.754–0.878), 0.731 (95% CI, 0.660–0.801), and 0.889 (95% CI, 0.844–0.934), respectively. The AUROC of

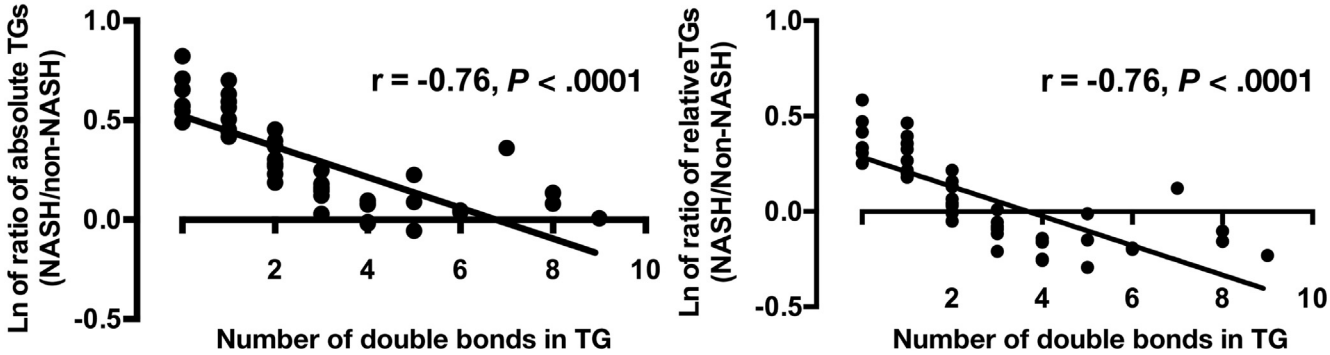


Figure 2. Relationships between fold-changes of mean concentrations of individual TGs (NASH/non-NASH) and numbers of double bonds contained in each TG. Each *dot* represents a TG molecule. The y-axis denotes the $\log_2$ of the ratio of concentrations of TGs between NASH and non-NASH patients, and the x-axis shows the number of double bonds in TGs. *Left*: $\log_2$ of fold-changes of absolute TG abundances plotted against the number of double bonds. *Right*: $\log_2$ of fold changes of relative TG abundances plotted against the number of double bonds.

Table 3. Comparison of the Performances of the Scores for Diagnosing NASH in 318 Biopsy Patients

Test scores	AUROC (95% CI)	Cut-off value	Sens, %	Spec, %	PPV, %	NPV, %
NASH ClinLipMet score	0.866 (0.820–0.913)	0.134	85.5	72.1	45.3	94.8
NASH Lip Score	0.779 (0.717–0.841)	0.148	88.4	53.8	34.7	94.4
NASH Met Score	0.719 (0.655–0.782)	0.203	65.2	69.1	36.9	87.8
NASH Clin Score = NASH score[6]	0.792 (0.726–0.859)	-1.354	77.4	70.7	41.7	92.0

NOTE. See main text for statistical comparisons of AUROCs.
NPV, negative predictive value; PPV, positive predictive value; Sens, sensitivity; Spec, specificity.

the NASH ClinLipMet Score was not significantly different between the entire cohort and the group not using statins ($P = .496$).

Comparison of Area Under the Receiver Operating Characteristics

The AUROC of the NASH ClinLipMet Score ($P < .001$) was significantly higher than that of the NASH Lip Score ($P < .05$), the NASH Met Score ($P < .001$), and the NASH Clin Score ($P < .01$) (Figure 3). The performance of each score is summarized in Table 3.

Comparison of Nonalcoholic Steatohepatitis With Nonalcoholic Fatty Liver

We also determined whether the MS-based markers in the NASH group were specific to NASH or also observed between NAFL as compared with non-NAFLD subjects. The clinical characteristics of NASH, NAFL, and non-NAFLD groups are shown in Table 1.

A comparison of lipid concentrations between the 3 groups are shown in Supplementary Figure 1. Of the 3 lipids entering the final lipidomics model, TG48:0 (shown as TG[14:0/16:0/18:0] + TG[16:0/16:0/16:0] in

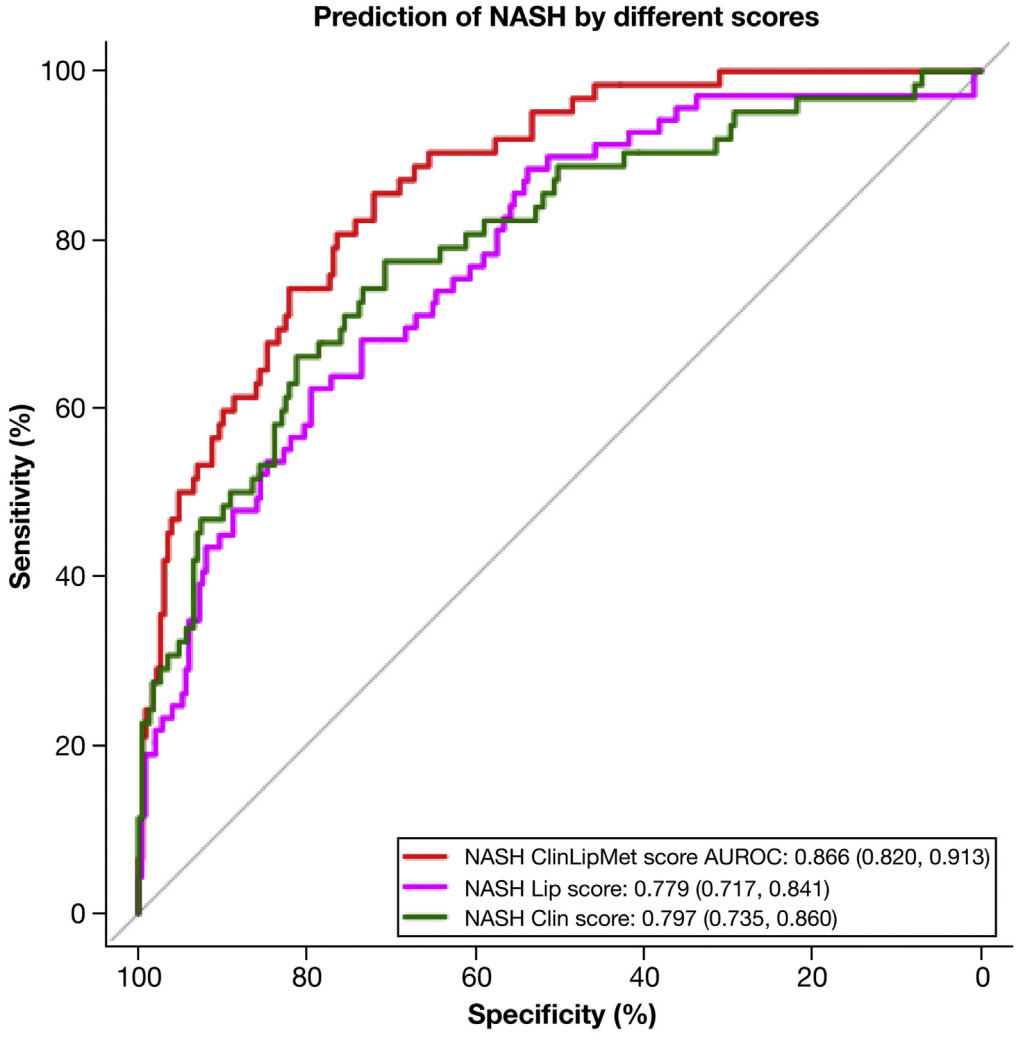


Figure 3. ROC curves of the 3 scores to predict NASH in the entire biopsy cohort. The AUROCs are compared in Table 3. Please see the text for definitions of scores.

Supplementary Figure 2) differed significantly between NASH, and NAFL and NAFL and non-NAFLD. PE40:6 and LysoPC16:0 differed significantly between NASH and NAFL, but not between NAFL and controls (Supplementary Figure 2).

Discussion

This study developed an MS-based model and score for NASH and compared its diagnostic performance with scores based on routinely available data and on PNPLA3 genotype at rs738409. We identified a set of lipids and metabolites that significantly were associated with NASH in a liver biopsy cohort of 318 subjects. We performed Gibbs sampling and backward stepwise logistic regression to select variables that fulfilled Bayesian information criterion. A model that included AST, the PNPLA3 genotype, fasting insulin, LysoPC(16:0), PE(40:6), Glu, Ile, and Gly best predicted NASH (the NASH ClinLipMet Score). The AUROC of this score was 0.86, which was significantly higher than that of the NASH Lip score, NASH Met Score, NASH Clin Score,[6] NASH Liver Fat Score,[6,9] and the NAFLD lipid triplet score.[14] These data show that MS-based profiling combined with clinical variables may help in the development of a noninvasive diagnosis of NASH.

The NASH group, compared with the non-NASH group, had an absolute and relative excess of saturated and monounsaturated TGs in their circulating lipidomics profile (Figure 1). TGs containing saturated fatty acids and monounsaturated fatty acid were shown previously to be overproduced in a study involving 9 subjects by the splanchnic area.[15] De novo lipogenesis produces saturated fatty acids exclusively.[16] Stable isotope studies tracing the origin of intrahepatocellular TGs suggest that de novo lipogenesis is prominent and perhaps the only abnormal pathway in patients with NAFLD.[17] Hence, circulating TGs containing saturated fatty acids and monounsaturated fatty acids might reflect increased de novo lipogenesis. Individual TGs did not, however, remain significant independent predictors of NASH in the final model including both clinical and MS-profiling–based parameters (NASH ClinLipMet). This is most likely because of multicolinearity, that is, saturated and monounsaturated TGs were correlated closely with features of IR such as fS insulin and thus more markers of IR and steatosis than NASH.

Circulating LysoPC16:0 deficiency was associated with NASH. This metabolite as well as other lysoPCs and phosphatidylcholines (Supplementary Figure 2), which also were deficient in NASH, are found mostly in the HDL lipoprotein fraction, which was low in the NASH as compared with other groups.[18] LysoPC16:0 recently was found to be the most deficient metabolite when comparing 180 metabolites between 20 insulin-resistant and 20 insulin-sensitive morbidly obese subjects with NAFLD.[19] In 14 subjects who underwent a liver biopsy, lysoPC16:0 levels were higher in insulin-resistant subjects with a trend toward higher inflammation in their liver.[19] Low lysoPC16:0 concentrations also were observed in preadipocyte cultures from 10 metabolically unhealthy as compared with 10 metabolically healthy obese subjects.[20]

The metabolite data are consistent with data reported in several small studies. Branch chain amino acids (BCAA) and essential amino acids are increased in obese/insulin-resistant subjects.[21] BCAA also promote IR induced by high-fat feeding. Increases in BCAA are accompanied by increases in C3 and C5 acylcarnitines, which are BCAA metabolites in the liver and in skeletal muscle.[22] The increase in the BCAA Ile, and in Glu, which is the first step of BCAA catabolism, therefore could be attributed to obesity/IR, which is associated with NASH. Increases in Glu previously have been found in studies that included 24[3] and 16[23] patients with NASH. Very recently, a genome-scale metabolomics model was constructed to interpret liver transcriptome data in NASH patients. Altered Glu metabolism was predicted to be the single most abnormal site of metabolism in NASH.[24] The second most common abnormality was predicted to be Ser deficiency, which is known to characterize patients with NASH and insulin-resistant as compared with insulin-sensitive subjects.[3,25] Consistently, Ser deficiency also characterized the patients with NASH in the present study. Gly is formed from Ser in a reaction catalyzed by SHMT1, an enzyme leading to Ser formation that was predicted to be down-regulated in NASH.[24] Thus, the observed changes in amino acid concentrations in the NASH group compared with the non-NASH group reflect previously described pathophysiologic changes in human beings and in experimental animals.

Use of the predictive equation developed in this study requires set-up of an assay specifically measuring each component. This is feasible given that an increasing number of analytic methods in clinical chemistry laboratories use targeted MS-based methods.[26] Once established and automated, such a method requires a minute amount of plasma and is less time consuming and cumbersome for clinicians than assessment of the different components, especially scores that necessitate inclusion of parameters from physical examination such as waist circumference.[27] Regarding the cost and reproducibility of the "omics" technology, it is important to establish whether screening using the NASH ClinLipMet Score is cost effective. This cannot be performed based on the present study, which is a first step and shows that it is possible to improve the diagnostic accuracy of a predictive score for NASH using MS-based analytic platforms in morbidly obese patients with a high prevalence of NASH.

Limitations of the present study should be considered when interpreting the results. The score was derived from a cohort including a large number of obese patients, which may hamper its application to the general population. Although the histologic criteria for NASH are

similar irrespective of obesity, it is important to validate the NASH ClinLipMet Score in a cohort that is not morbidly obese. Treatment with lipid-lowering drugs may influence their plasma lipid levels and act as a potential confounder. However, the performance of the NASH Clin Score, NASH Lip Score, NASH Met Score, and the NASH ClinLipMet Score was not influenced by use of lipid-lowering medications, which thus suggest that the scores are robust and that use of lipid-lowering medications does not limit the usefulness of these scores. The study was cross-sectional and thus the term *predictor* merely denotes a factor that is associated with the risk of NASH. Scores ideally should be validated in a longitudinal study, but such a study is challenging because it is ethically unacceptable to obtain repeated liver biopsy specimens from individuals with no indication for such a procedure. Although the NASH ClinLipMet Score had the highest AUROC of 0.86 in diagnosing NASH among the formulae tested with sensitivities and specificities of 80.6% and 75.3%, respectively, the diagnosis of NASH will be missed in 19.4% of those with NASH and 24.7% will be diagnosed incorrectly as having the disease. The score developed in Finnish and Belgian Caucasian subjects may not be valid in other ethnic groups.

In conclusion, use of MS-based methods helps in improving a noninvasive diagnosis of NASH compared with scores relying on routinely available clinical data and PNPLA3 genotype at rs738409. In the present study, the findings of increases of the saturated TG 48:0, Glu, Ile, and decreases in lysoPC 16:0, Ser, and Gly in a relatively large cohort of patients with NASH are consistent with the known pathophysiology of NASH.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at http://dx.doi.org/10.1016/j.cgh.2016.05.046.

References

1. Yki-Jarvinen H. Non-alcoholic fatty liver disease as a cause and a consequence of metabolic syndrome. Lancet Diabetes Endocrinol 2014;2:901–910.
2. Pitt JJ. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. Clin Biochem Rev 2009;30:19–34.
3. Kalhan SC, Guo L, Edmison J, et al. Plasma metabolomic profile in nonalcoholic fatty liver disease. Metabolism 2011;60:404–413.
4. Barr J, Caballeria J, Martinez-Arranz I, et al. Obesity-dependent metabolic signatures associated with nonalcoholic fatty liver disease progression. J Proteome Res 2012;11:2521–2532.
5. Luukkonen PK, Zhou Y, Sadevirta S, et al. Ceramides dissociate steatosis and insulin resistance in the human liver in non-alcoholic fatty liver disease. J Hepatol 2016;64:1167–1175.
6. Hyysalo J, Mannisto VT, Zhou Y, et al. A population-based study on the prevalence of NASH using scores validated against liver histology. J Hepatol 2014;60:839–846.
7. Kotronen A, Peltonen M, Hakkarainen A, et al. Prediction of non-alcoholic fatty liver disease and liver fat using metabolic and genetic factors. Gastroenterology 2009;137:865–872.
8. Brunt EM, Janney CG, Di Bisceglie AM, et al. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol 1999;94:2467–2474.
9. Pluskal T, Castillo S, Villar-Briones A, et al. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics 2010;11:395.
10. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med 1990;9:811–818.
11. Zhou Y, Llaurado G, Oresic M, et al. Circulating triacylglycerol signatures and insulin sensitivity in NAFLD associated with the E167K variant in TM6SF2. J Hepatol 2015;62:657–663.
12. George EI, Robert EM. Variable selection via Gibbs sampling. J Am Stat Assoc 1993;88:881–889.
13. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–845.
14. Oresic M, Hyotylainen T, Kotronen A, et al. Prediction of non-alcoholic fatty-liver disease and liver fat content by serum molecular lipids. Diabetologia 2013;56:2266–2274.
15. Westerbacka J, Kotronen A, Fielding BA, et al. Splanchnic balance of free fatty acids, endocannabinoids, and lipids in subjects with nonalcoholic fatty liver disease. Gastroenterology 2010;139:1961–1971.e1.
16. Aarsland A, Wolfe RR. Hepatic secretion of VLDL fatty acids during stimulated lipogenesis in men. J Lipid Res 1998;39:1280–1286.
17. Lambert JE, Ramos-Roman MA, Browning JD, et al. Increased de novo lipogenesis is a distinct characteristic of individuals with nonalcoholic fatty liver disease. Gastroenterology 2014;146:726–735.
18. Kotronen A, Velagapudi VR, Yetukuri L, et al. Serum saturated fatty acids containing triacylglycerols are better markers of insulin resistance than total serum triacylglycerol concentrations. Diabetologia 2009;52:684–690.
19. Lehmann R, Franken H, Dammeier S, et al. Circulating lysophosphatidylcholines are markers of a metabolically benign nonalcoholic fatty liver. Diabetes Care 2013;36:2331–2338.
20. Bohm A, Halama A, Meile T, et al. Metabolic signatures of cultured human adipocytes from metabolically healthy versus unhealthy obese individuals. PLoS One 2014;9:e93148.
21. Felig P, Wahren J, Hendler R, et al. Splanchnic glucose and amino acid metabolism in obesity. J Clin Invest 1974;53:582–590.
22. Newgard CB. Interplay between lipids and branched-chain amino acids in development of insulin resistance. Cell Metab 2012;15:606–614.
23. Li H, Wang L, Yan X, et al. A proton nuclear magnetic resonance metabolomics approach for biomarker discovery in nonalcoholic fatty liver disease. J Proteome Res 2011;10:2797–2806.
24. Mardinoglu A, Agren R, Kampf C, et al. Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. Nat Commun 2014;5:3083.
25. Thalacker-Mercer AE, Ingram KH, Guo F, et al. BMI, RQ, diabetes, and sex affect the relationships between amino acids and

clamp measures of insulin action in humans. Diabetes 2014; 63:791–800.

26. Kenar E, Franken H, Forcisi S, et al. Automated label-free quantification of metabolites from liquid chromatography-mass spectrometry data. Mol Cell Proteomics 2014;13:348–359.
27. Grandison GA, Angulo P. Can NASH be diagnosed, graded, and staged noninvasively? Clin Liver Dis 2012;16:567–585.

Reprint requests
Address requests for reprints to: Hannele Yki-Järvinen, MD, University of Helsinki, Haartmaninkatu 8, 00290 Helsinki, Finland. e-mail: hannele.yki-jarvinen@helsinki.fi; fax: +358 9 4717 1896.

Acknowledgments
The authors thank Anne Salo, Aila Karioja-Kallio, and Pentti Pölönen for their excellent technical assistance.

Conflicts of interest
The authors disclose no conflicts.

Funding
This study was supported by research grants from the Academy of Finland (H.Y.-J. and 265966 to P.G.), EU/European Federation of Pharmaceutical Industries and Associations Innovative Medicines Initiative Joint Undertaking (European Medical Information Framework 115372 to H.Y.-J., M.O., T.H.), EU H2020 Elucidating Pathways of Steatohepatitis 634413 (H.Y.-J., M.O., T.H.), the Sigrid Juselius (H.Y.-J.), EVO (H.Y.-J.), the Novo Nordisk Foundation (H.Y.-J.), the Orion-Farmos Research Foundation (Y.Z.) and the Clinical Chemistry Research Foundation (Y.Z.).

Supplementary Methods

Lipidomic Analysis by Ultra-Performance Liquid Chromatography–Quadruple Time-of-Flight Mass Spectrometry

An aliquot (10 μL) of the plasma sample was diluted with 10 μL of 0.15 mol/L (0.9%) sodium chloride, and 10 μL of internal standard mixture 1A was added. This mixture contained phosphatidylcholine (17:0/0:0) and (17:0/17:0), PE(17:0/17:0), phosphatidylglycerol (17:0/17:0), Cer(d18:1/17:0), PS(17:0/17:0), and PA(17:0/17:0) (Avanti Polar Lipids, Inc, Alabaster, AL), as well as monoacylglycerol (17:0/0:0/0:0), diacylglycerol (17:0/17:0/0:0), and TG(17:0/17:0/17:0). The lipids were extracted using a mixture of high-performance liquid chromatography–grade chloroform and methanol (2:1; 100 μL). The lower phase (60 μL) was collected and 10 μL of an internal standard mixture containing labeled phosphatidylcholine (16:1/0:0-D_3), phosphatidylcholine (16:1/16:1-D_6), and TG(16:0/16:0/16:0-^{13}C3) was added.

The extracts were analyzed on a Waters Q-Tof Premier mass spectrometer combined with an Acquity Ultra Performance LC. The column (at 50°C) was an Acquity UPLC BEH C18 2.1 × 100 mm with 1.7-μm particles. The solvent system included (A) ultrapure water including 1% 1 mol/L NH_4Ac and 0.1% HCOOH) and (B) LC/MS-grade acetonitrile/isopropanol (1:1, 1% 1 mol/L NH_4Ac, 0.1% HCOOH). The gradient started from 65% A/35% B, reached 80% B in 2 minutes, 100% B in 7 minutes, and remained there for 7 minutes. The flow rate was 0.400 mL/min and the injected amount was 2.0 μL (Acquity Sample Organizer, at 10°C). Reserpine was used as the lock spray reference compound. The lipid profiling was performed using the electrospray ionization mode and the data were collected at a mass range of m/z 300 to 1200 with a scan duration of 0.2 seconds.

The data processing included alignment of peaks, peak integration, normalization, and identification. Lipids were identified using an internal spectral library. The data were normalized using 1 or more internal standards representative of each class of lipid present in the samples: the intensity of each identified lipid was normalized by dividing it with the intensity of its corresponding standard and multiplying it by the concentration of the standard. All monoacyl lipids except cholesterol esters, such as monoacylglycerols and monoacylglycerophospholipids, were normalized with phosphatidylcholine (17:0/0:0), all diacyl lipids except ethanolamine phospholipids were normalized with phosphatidylcholine (17:0/17:0), all ceramides with Cer(d18:1/17:0), all diacyl ethanolamine phospholipids with PE(17:0/17:0), and TG and cholesterol esters with TG(17:0/17:0/17:0). Other (unidentified) molecular species were normalized with phosphatidylcholine (17:0/0:0) for retention times shorter than 300 seconds, phosphatidylcholine (17:0/17:0) for a retention time between 300 and 410 seconds, and TG(17:0/17:0/17:0) for longer retention times.

Quality control of the method showed that the day-to-day repeatability of control serum samples, and the relative standard deviation for values identified was on average less than 25% and 20% for the discovery and validation sets, respectively. The internal standards added to all samples in the study had an average relative standard deviation of 25% and 13% in the discovery and validation sets, respectively.

For further identification of unknown lipids, fractions collected from the UPLC run were infused to a LTQ-Orbitrap (Thermo Fischer Scientific, San Jose, CA) mass spectrometer by a TriVersa Nanomate (Advion Biosciences, Ithaca, NY) using chip-based nanoelectrospray in positive and negative ionization mode. Identifications were based on the exact mass and MS^n spectra. The instrument was calibrated externally according to the manufacturer's instructions. MS^2 and MS^3 were acquired using either low resolution or high resolution up to target mass resolution of R = 60 000 at m/z 400. The normalized collision energies of 30% to 40% were applied in MS^n experiments.

Metabolomic Analysis

Polar metabolites were analyzed using comprehensive 2-dimensional gas chromatography combined with time-of-flight mass spectrometry (GC × GC - TOFMS).[1] A total of 400 μL methanol and 10 μL internal standard mixture (C17:0 [93.3 mg/L], valine-d [18.5 mg/L], and succinic acid-d4 [31.5 mg/L]) were added to 30 μL of plasma samples. The samples were vortex mixed (2 minutes at 20 Hz). After 30 minutes at room temperature the samples were centrifuged for 5 minutes at 10,000 rpm. The supernatant was moved to a GC vial and evaporated to dryness under nitrogen. The samples were trimethylsilylated with 25 μL Methoxyamine (45°C, 60 min) and 25 μL N-Methyl-N-(trimethylsilyl)trifluoroacetamide (45°C, 60 min), and 5 μL of retention index solution in hexane was added to samples (150 mg/L C11, C15, C17, C21, and C25 alkanes).

For metabolomics analysis, a Leco Pegasus 4D GC×GC-TOFMS instrument (Leco Corp, St. Joseph, MI) equipped with a cryogenic modulator was used. The GC part of the instrument was an Agilent 6890 gas chromatograph (Agilent Technologies, Palo Alto, CA), equipped with a split/splitless injector. The first-dimension chromatographic column was a 10-m RTX-5 capillary column with an internal diameter of 0.18 mm and a stationary phase film thickness of 0.20 μm, and the second-dimension chromatographic column was a 1.5 m BPX-50 capillary column with an internal diameter of 100 μm and a film thickness of 0.1 μm. A methyl deactivated retention gap (3 m × 0.53 mm i.d.) was used in the front of the first column. High-purity helium was

used as the carrier gas at a constant pressure mode (39.6 pounds per square inch gauge). A 5-second separation time was used in the second dimension. The MS spectra were measured at 45 to 700 atomic mass unit with 100 spectra/second. For the injection, a pulsed splitless injection (0.5 μL) at 240°C was used, with a pulse pressure of 55 psig for 1 minute. The temperature program was as follows: the first-dimension column oven ramp began at 40°C with a 2-minute hold, after which the temperature was programmed to 295°C at a rate of 7°C/min, and then held at this temperature for 3 minutes. The second-dimension column temperature was maintained at 20°C higher than the corresponding first-dimension column. The programming rate and hold times were the same for the 2 columns.

ChromaTOF vendor software (Leco) was used for within-sample data processing, including quantitation of selected target metabolites, and Guineu software (Espoo, Finland) was used for alignment, normalization, and peak matching across samples.[1] The peaks were first filtered based on the number of detected peaks in the total profile of all sample runs. The normalization was performed by correction for internal standards. Other mass spectra from the GC × GC - TOFMS analysis were searched against The Palisade Complete Mass Spectral Library, 600K Edition (Palisade Mass Spectrometry, Ithaca, NY). Data were processed using the Guineu software.

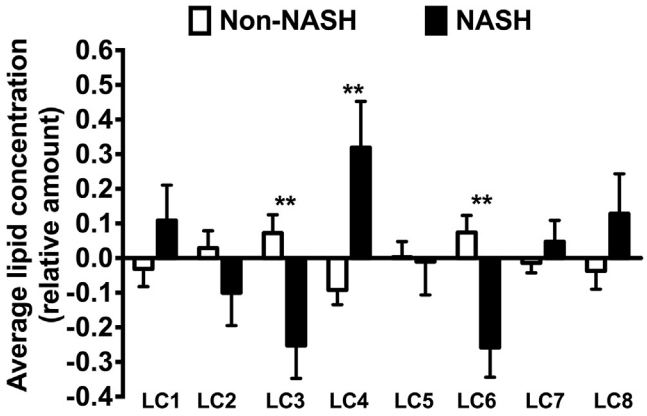
After nontargeted profiling of the discovery set, 13 metabolites were selected for quantitative analysis (amino acids, free fatty acids), which then were quantified in both validation and discovery sets using external calibration curves, after normalization with the labeled group-specific internal standards. Quality control of the method showed that the day-to-day repeatability of control serum samples, and the relative standard deviation for values identified was on average less than 22% and 19% for the discovery and validation sets, respectively. The internal standards added to all the samples in the study had an average relative standard deviation of 20% and 18% in the discovery and validation sets, respectively.

Statistical Analyses

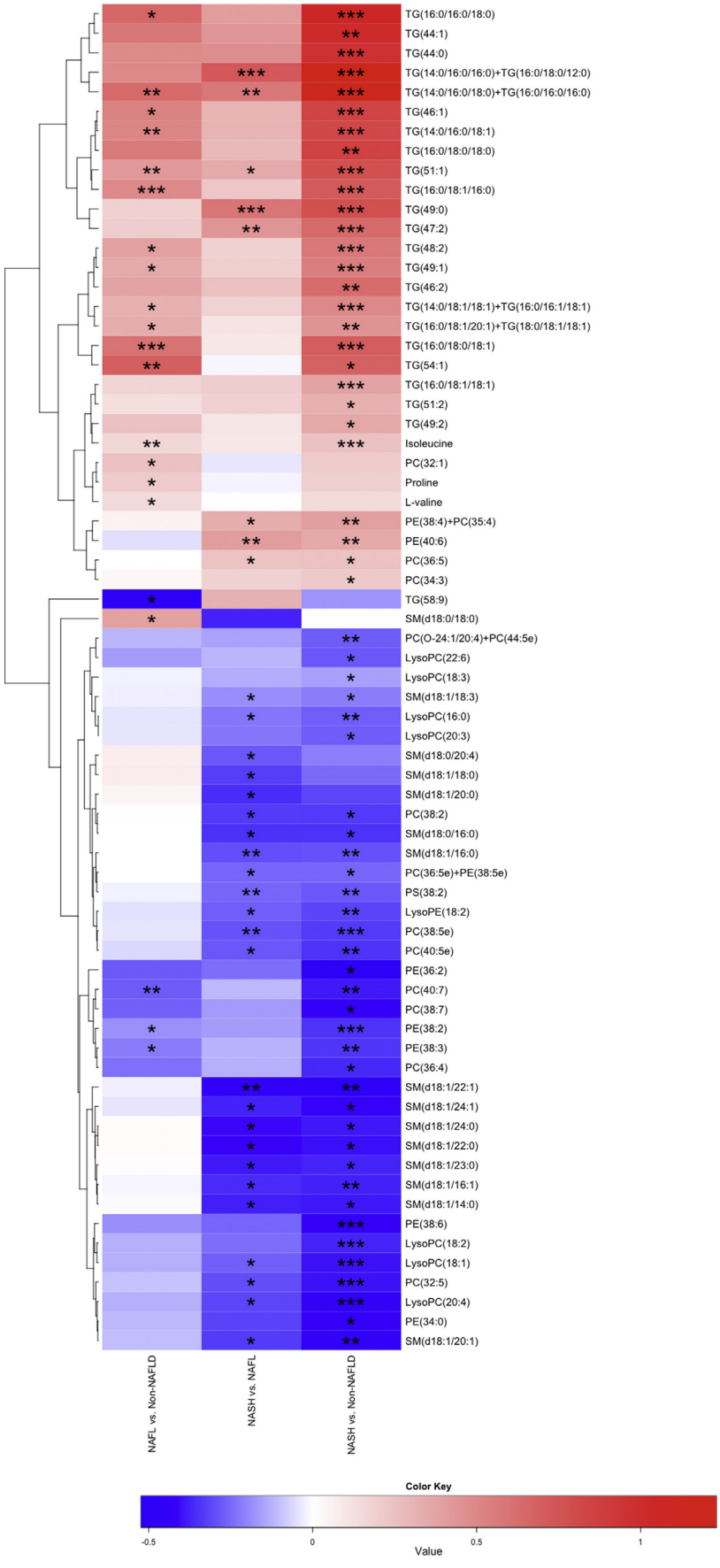
Assessment of abundances of triacylglycerol species. Mean and standard errors of abundances of plasma TG molecular species were calculated. After $\log_2$ transformation, the average abundances of TG molecules were compared between the NASH and non-NASH groups by Student *t* tests. Multiple comparisons were corrected by using the Benjamini–Hochberg method.[2] The comparisons were illustrated by heat maps, which plot chain lengths of fatty acid against the number of double bonds for each TG. R Package, metadar (http://code.google.com/p/metadar) was used for data analysis.

References

1. Castillo S, Mattila I, Miettinen J, et al. Data analysis tool for comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry. Anal Chem 2011;83:3058–3067.
2. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med 1990;9:811–818.



Supplementary Figure 1. Mean lipid concentrations within each cluster between NASH and non-NASH groups. $^{**}P < .01$ for differences between the groups.



Supplementary Figure 2. Comparison of concentrations of lipids between NASH, NAFL, and non-NAFLD groups. The color code denotes the $\log_2$ of the ratio between the means of the groups for an individual lipid. The y-axes represent the names of lipids, and the x-axes show the groups for comparison (NAFL vs non-NAFLD, NASH vs NAFL, and NASH vs non-NAFLD). The Tukey honestly significant differences post hoc test was used to compare 2 groups after 1-way analysis of variance. Blue represents a decrease and red shows an increase between groups. The brighter the red color, the greater the increase of absolute concentration of the individual lipid between groups. The brighter the blue color, the greater the decrease. The significances for the comparisons are marked as follows: *$P < .05$, **$P < .005$, and ***$P < .0005$.

Supplementary Table 1. Comparison Between the Clinical Characteristics in the Estimation and Validation Groups for Building the NASH Score

Total	Estimation group (n = 223)	Validation group (n = 95)	*P* value
Liver fat, %	15 (5–40)	12.5 (0–47.5)	.9
Grade, 0/1/2/3, n	176/40/7/0	71/19/4/1	.41
Stage, 0/1/2/3/4, n	150/55/12/4/2	65/23/5/1/1	.99
NASH, n (%)	47 (21.1)	24 (25.3)	.5
Age, *y*	46 ± 1	47 ± 1	.79
Sex, n (% women)	140 (62.8)	57 (60)	.73
BMI, kg/m^2	39.9 ± 0.5	39.8 ± 1.0	.61
fP glucose, *mmol/L*	4.9 (4.4–5.8)	5.0 (4.5–5.7)	.55
HbA_{1C}, %	5.7 (5.4–6)	5.7 (5.4–6)	.92
fP triglycerides, *mmol/L*	1.46 (1.09–2.18)	1.36 (0.96–2.03)	.39
fP HDL cholesterol, *mmol/L*	1.17 (0.99–1.46)	1.22 (1.02–1.46)	.37
fP LDL cholesterol, *mmol/L*	2.9 (2.19–3.51)	2.81 (2.24–3.49)	.72
P-AST, *IU/L*	29 (25–42)	32 (24–40)	.67
P-ALT, *IU/L*	40 (31–59)	42 (32–56)	.64
P-GGT, *U/L*	30 (22–54)	31 (23–54)	.63
P-albumin, *g/L*	41.5 ± 0.4	42.5 ± 0.6	.25
B-platelets, $\times 10^9$/L	260 (214–311)	256 (215–303)	.67
PNPLA3, CC/GC/GG, n	108/85/17	50/35/4	.49
Type 2 diabetes, n (%)	43 (19.5)	18 (18.9)	1
Metabolic syndrome, n (%)	145 (65.3)	62 (65.3)	1

NOTE. Data are shown as n (%), means ± SEM, or median (25th–75th percentile), as appropriate.
ALT, alanine aminotransferase; B, blood; BMI, body mass index; GGT, γ-glutamyl transpeptidase; HbA_{1C}, glycosylated hemoglobin 1c; LDL, low-density lipoprotein; P, plasma.

Supplementary Table 2. Clinical Characteristics of the NASH and the Non-NASH Groups in the Estimation Group

Estimation group	Non-NASH (n = 176)	NASH (n = 47)	*P* value
Liver fat, %	10 (0, 25)	40 (26.3, 67.5)	<.001
Grade, 0/1/2/3, n	176/0/0/0	0/40/7/0	<.001
Stage, 0/1/2/3/4, n	145/24/5/0/2	5/31/7/4/0	<.001
Age, *y*	46 ± 1	48 ± 2	.21
Sex, n (% women)	115 (65.3)	25 (53.2)	.17
BMI, kg/m^2	39.5 ± 0.6	41.4 ± 1.1	.15
fP glucose, *mmol/L*	4.8 (4.4–5.7)	5.7 (5–6.5)	<.001
HbA_{1C}, %	5.6 (5.4–6)	6 (5.8–6.7)	<.001
fP triglycerides, *mmol/L*	1.38 (1.07–1.95)	1.84 (1.36–2.54)	.005
fP HDL cholesterol, *mmol/L*	1.22 (1.01–1.51)	1.07 (0.95–1.25)	.012
fP LDL cholesterol, *mmol/L*	2.89 (2.17–3.52)	3 (2.32–3.51)	.46
P-AST, *IU/L*	28 (24–38)	43 (28–71)	<.001
P-ALT, *IU/L*	37 (30–52)	53 (40–103)	<.001
P-GGT, *U/L*	28 (21–48)	48 (29–74)	<.001
P-albumin, *g/L*	41.6 ± 0.5	41.3 ± 0.9	.89
B-platelets, $\times 10^9$/L	264 (219–314)	243 (199–296)	.11
PNPLA3, CC/CG/GG, n	90/66/10	18/19/7	.064
Type 2 diabetes, n (%)	24 (13.6)	19 (42.2)	<.001
Metabolic syndrome, n (%)	106 (60.2)	39 (84.9)	.0033
Hyperglycemic medication, n (%)	18 (10.2)	15 (31.9)	<.001

NOTE. Data are shown as n (%), means ± SEM, or median (25th–75th percentile), as appropriate.
ALT, alanine aminotransferase; B, blood; BMI, body mass index; GGT, γ-glutamyl transpeptidase; HbA_{1C}, glycosylated hemoglobin 1c; LDL, low-density lipoprotein; P, plasma.

Supplementary Table 3. Clinical Characteristics of the NASH and the Non-NASH Groups in the Validation Group

Validation group	Non-NASH (n = 71)	NASH (n = 24)	*P* value
Liver fat, %	5 (0–31.3)	42.5 (30–75)	<.001
Grade, 0/1/2/3, n	71/0/0/0	0/19/4/1	<.001
Stage, 0/1/2/3/4, n	63/6/2/0/0	2/17/3/1/1	<.001
Necroinflammation, 0/1/2, n	70/1/0	0/15/9	<.001
Age, *y*	45 ± 2	52 ± 2	<.001
Sex, n (% women)	43 (60.6)	14 (58.3)	1
BMI, kg/m^2	39.4 ± 1.2	40.8 ± 1.7	.36
fP glucose, *mmol/L*	5.0 (4.6–5.6)	5.5 (4.5–6.3)	.2
HbA_{1C}, %	5.6 (5.4–6.0)	5.9 (5.6–6.4)	.042
fP triglycerides, *mmol/L*	1.27 (0.94–1.82)	1.86 (1.28–2.77)	.021
fP HDL cholesterol, *mmol/L*	1.25 (1.07–1.51)	1.05 (0.91–1.29)	.028
fP LDL cholesterol, *mmol/L*	2.81 (2.19–3.5)	2.7 (2.4–3.48)	.71
P-AST, *IU/L*	31 (24–38)	40 (30–48)	.0047
P-ALT, *IU/L*	39 (31–51)	54 (40–60)	.012
P-GGT, *U/L*	29 (22–56)	41 (29–52)	.18
P-albumin, *g/L*	42.9 ± 0.7	41.5 ± 1.2	.46
B-platelets, $\times 10^9/L$	256 (219–296)	248 (195–325)	.65
PNPLA3, CC/CG/GG, n	42/24/0	8/11/4	<.001
Type 2 diabetes, n (%)	8 (11.3)	10 (41.7)	.0028
Metabolic syndrome, n (%)	41 (57.7)	21 (87.5)	.016
Hyperglycemic medication, n (%)	6 (8.4)	7 (30.4)	.021

NOTE. Data are shown as n (%), means ± SEM, or median (25th–75th percentile), as appropriate.

ALT, alanine aminotransferase; B, blood; BMI, body mass index; GGT, γ-glutamyl transpeptidase; HbA_{1C}, glycosylated hemoglobin 1c; LDL, low-density lipoprotein; P, plasma.