Metabolism Clinical and Experimental 101 (2019) 154005

Contents lists available at ScienceDirect

Metabolism Clinical and Experimental

journal homepage: www.metabolismjournal.com

Non-invasive diagnosis of non-alcoholic steatohepatitis and fibrosis with the use of omics and supervised learning: A proof of concept study

Nikolaos Perakakis [a,*,1], Stergios A. Polyzos [b,1], Alireza Yazdani [c,1], Aleix Sala-Vila [d,e], Jannis Kountouras [f], Athanasios D. Anastasilakis [g], Christos S. Mantzoros [a,*]

[a] Division of Endocrinology, Diabetes and Metabolism, Department of Internal Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA
[b] First Department of Pharmacology, Faculty of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece
[c] Division of Applied Mathematics, Brown University, Providence, RI 02906, USA
[d] CIBER de Fisiopatología de la Obesidad y la Nutrición (CIBEROBN), Instituto de Salud Carlos III, Madrid, Spain
[e] Institut d'Investigacions Biomediques August Pi i Sunyer (IDIBAPS), Hospital Clinic of Barcelona, Villarroel 170, Barcelona 08036, Spain
[f] Second Medical Clinic, Faculty of Medicine, Aristotle University of Thessaloniki, Ippokration Hospital, Thessaloniki, Greece
[g] Department of Endocrinology, 424 General Military Hospital, Thessaloniki, Greece

ARTICLE INFO

Article history:
Received 20 August 2019
Accepted 6 November 2019

Keywords:
Non-alcoholic fatty liver disease
Non-alcoholic steatohepatitis
Liver fibrosis
Metabolomics
Lipidomics
Glycomics
Machine learning
Non-invasive

ABSTRACT

Background: Non-alcoholic fatty liver disease (NAFLD) affects 25–30% of the general population and is characterized by the presence of non-alcoholic fatty liver (NAFL) that can progress to non-alcoholic steatohepatitis (NASH), liver fibrosis and cirrhosis leading to hepatocellular carcinoma. To date, liver biopsy is the gold standard for the diagnosis of NASH and for staging liver fibrosis. This study aimed to train models for the non-invasive diagnosis of NASH and liver fibrosis based on measurements of lipids, glycans and biochemical parameters in peripheral blood and with the use of different machine learning methods.

Methods: We performed a lipidomic, glycomic and free fatty acid analysis in serum samples of 49 healthy subjects and 31 patients with biopsy-proven NAFLD (15 with NAFL and 16 with NASH). The data from the above measurements combined with measurements of 4 hormonal parameters were analyzed with two different platforms and five different machine learning tools.

Results: 365 lipids, 61 glycans and 23 fatty acids were identified with mass-spectrometry and liquid chromatography. Robust differences in the concentrations of specific lipid species were observed between healthy, NAFL and NASH subjects. One-vs-Rest (OvR) support vector machine (SVM) models with recursive feature elimination (RFE) including 29 lipids or combining lipids with glycans and/or hormones (20 or 10 variables total) could differentiate with very high accuracy (up to 90%) between the three conditions. In an exploratory analysis, a model consisting of 10 lipid species could robustly discriminate between the presence of liver fibrosis or not (98% accuracy).

Conclusion: We propose novel models utilizing lipids, hormones and glycans that can diagnose with high accuracy the presence of NASH, NAFL or healthy status. Additionally, we report a combination of lipids that can diagnose the presence of liver fibrosis. Both models should be further trained prospectively and validated in large independent cohorts.

1. Introduction

Non-alcoholic fatty liver disease (NAFLD), the hepatic component of the metabolic syndrome, is now recognized as the most prevalent chronic liver disease worldwide; it affects 25–30% of the worldwide population [1]. Initially, an accumulation of fat is observed in the liver (non-alcoholic fatty liver, NAFL), that can progress to an inflammatory state (non-alcoholic steatohepatitis, NASH) and in later stages to liver

Abbreviations: NAFLD, non-alcoholic fatty liver disease; NAFL, non-alcoholic fatty liver; NASH, non-alcoholic steatohepatitis; HCC, hepatocellular carcinoma; IR, insulin resistance; TG, triglycerides; HDL, high-density lipoprotein; LDL, low-density lipoprotein; BMI, body mass index; SAM, significance analysis of microarrays; sPLS-DA, sparse Partial Least Squares – Discriminant Analysis; PCA, principal component analysis; t-SNE, t-distributed stochastic neighbor embedding; ROC, receiver operating characteristic; REF, recursive feature elimination; SVM, support vector machine; AUC, area under curve; OvR, One-vs-Rest; RBF, radial basis function; MCCV, Monte-Carlo cross validation; DG, diglycerides; PG, phosphatidylglycerols; PA, phosphatidic acids; AcCa, acylcarnitines; Che, cholesterol esters; Co, coenzyme Q10; LPC, lysophosphatidylcholines; SM, sphingomyelines; PE, Phosphatidylethanolamines; PC, Phosphatidylcholines; C16:0, palmitic acid; C16:1n7cis, cis-palmitoleic acid; C18:2n6, polyunsaturated linoleic acid; C20:4n6, arachidonic acid; kNN, k-nearest neighbor; Ck-18f, cytokeratin-18 fragment; C18:1n9, oleic acid; C18:3n6, gamma-linoleic acid.

* Corresponding authors at: SL-418, 330 Brookline Avenue, East campus, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA.
E-mail addresses: nperakak@bidmc.harvard.edu (N. Perakakis), cmantzor@bidmc.harvard.edu (C.S. Mantzoros).
[1] Indicates equal co-first authorship.

https://doi.org/10.1016/j.metabol.2019.154005
0026-0495/

fibrosis and cirrhosis that can lead to hepatocellular carcinoma (HCC) [2]. Median time to develop advanced fibrosis when inflammation is absent on the initial biopsy (patients with NAFL) is 13.4 years, while the median time when it is present (patients with NASH) is only 4.2 years [3].

The gold standard for the diagnosis of NAFLD (NAFL or NASH), and for assessing fibrosis is liver biopsy [4]. However, liver biopsy is a costly method that carries certain risks for the patient and the liver biopsy specimen is not always representative of the actual status of the entire liver, thereby often resulting in misclassification due to sampling error and/or inter-observer viability. Regarding the risks related to liver biopsy, the incidence of pain after biopsy is reported to be approximately 20%, whereas the incidence of serious complications ranges between 0.3 and 0.57% and of mortality at 0.01% [5]. Therefore, finding easy to obtain, relatively inexpensive and reliable biomarkers which can be measured with less invasive or even noninvasive techniques is an urgent unmet clinical need. Several studies have tried to identify non-invasive biomarkers or to develop diagnostic scores to distinguish between NAFL and NASH and the stage of liver fibrosis [6]. NAFL and NASH are complex multi factorial disorders and thus no single surrogate marker could probably predict in a reliable manner clinical outcome or benefits of therapeutic interventions. In this context, significant progress has been observed mainly with the use of imaging modalities in diagnosing NAFLD non-invasively [4], the level of hepatic steatosis and the presence or not of advanced liver fibrosis [6]. However, such tools are often not available in the primary care or in smaller community-based gastroenterological or endocrinological departments. Additionally, there are no reliable biomarkers or scores for the non-invasive diagnosis of NASH and for staging liver fibrosis in a similar score as in liver histology [4,6]. This is particularly important, since it will allow us to separate the benign cases (i.e. patients with NAFL) from the patients that require regular follow-up visits and treatment (i.e. patients with NASH and/or liver fibrosis).

Advances in omics in the last years and in the methods for analysis of big data have provided novel insights in the pathophysiology of many diseases including NASH [7]. Specifically, studies focusing on the comprehensive analysis of cellular lipids in serum or liver biopsies of patients with NASH, i.e. lipidomics, have achieved to identify important metabolic pathways involved in the development and progress of the disease [8–11]. Liver is also considered the main site for glycosylation, which is an important post-translational modification of secreted proteins that affects protein stability and folding and consequently their function [12]. The investigation of glycan structures (i.e. glycomics) in other liver diseases have led to the development of diagnostic tests, such as the GlycoCirrhotest [13], the GlycoFibrotest [14] and the GlycoHCC test [15] for the diagnosis of cirrhosis, fibrosis or HCC. Finally, certain hormones (i.e. adiponectin, leptin, follistatins, activins) have been associated both indirectly via regulation of insulin resistance (IR), inflammation [16] and glucose homeostasis [17–19] as well as directly by acting on hepatocytes with the development and progress of NASH and liver fibrosis [18,20–22].

Aim of our study was to investigate whether analyzing combined information from serum lipids, glycans and selected hormones with different machine learning tools can lead to the development of algorithms with high accuracy of predicting simultaneously NASH, NAFL or healthy status as well as the presence of liver fibrosis.

2. Methods

2.1. Study design

The study design has been previously described [20,23]. Inclusion and exclusion criteria are reported in the Supplementary Appendix. Serum was available from 80 subjects: 49 without NAFLD (healthy) and 31 with NAFLD (15 with NAFL and 16 with NASH). Liver biopsy was performed under computed tomography-guidance and histology was evaluated by two pathologists independently. Serum was collected after an overnight fasting and prior to liver biopsy.

2.2. Glycomics

Total N-glycans from serum samples were detected with MS. Details are provided in the Supplementary Appendix.

2.3. Lipidomics

A lipidomics analysis was performed with LC/MS. Details are provided in the Supplementary Appendix.

2.4. Quantification of serum fatty acids

Fatty acids of whole serum were converted to their corresponding fatty acid methyl esters [24]. Comparison of conventional and fast gas chromatography in human plasma fatty acid determination [24] and separated by gas-chromatography using an Agilent HP 7890 Gas Chromatograph equipped with a 30 m × 0.25 μm × 0.25 mm SupraWAX-280 capillary column (Teknokroma, Barcelona, Spain), an autosampler, and a flame ionization detector. The amount of each fatty acid is expressed as a percentage of the total identified fatty acids in the sample.

2.5. Targeted biochemical measurements

We have previously measured adiponectin, leptin, activin A, follistatin and triglycerides in serum samples of this study. Results from these measurements have been previously published [20,23]. These measurements are now included in the analysis to estimate their predictive accuracy in combination with other omics measurements. Details about the previous measurements are presented in the Supplementary Appendix.

2.6. Statistical analysis

2.6.1. Data processing, normalization and scaling

The values from all glycans were normalized initially to the value (ion signal) of the most abandoned glycan (2792). Two glycans (1999 and 2519) were removed from the analysis, as their missing values were >1/3 of the sample size (which was the exclusion threshold in our study). The values from all the lipidomic measurements were normalized to total positive or total negative ion signal. Fatty acid measurements were normalized and expressed as percent of total fatty acid concentration. Targeted biochemical measurements were not normalized. A missing value in mass spectrometric analysis indicated very low signal. Thus, a fraction (1/5th) of the minimum value was added to all variables including the missing values in lipidomics and glycomics. No missing values were present in the fatty acids data set. The missing values in targeted biochemical measurements were very few (2 in Activin A, 1 in cholesterol, triglycerides (TG), high- and low- density lipoproteins (HDL and LDL), and were related to the lack of available serum (and not to low signal) for these subjects and, thus, were replaced with the median value of the group. Since lipidomics, glycomics and fatty acids data are normalized, they carry relative information. For such compositional data, a centered log ratio transformation is recommended [25], where the ratio of measurements by their geometric mean for each variable is log-transformed. Furthermore, a log transformation is performed on targeted biochemical measurements to achieve a balanced distribution. Each dataset is then mean-centered and scaled to have a unit variance [26] (s. Supplementary Fig. A1).

2.6.2. Data analysis and visualization

In Supplementary Table A1 and in Fig. 3E, mean ± SD for normally distributed and median with 25th and 75th percentile are reported.

Shapiro-Wilk test was used to assess the distributions. One-way ANOVA for normally distributed and Kruskal-Wallis test for non-normally distributed features was performed followed (when $p < 0.05$) by post-hoc Tukey's or Dunn's test. ANCOVA was performed subsequently in order to adjust for BMI between groups (non-normally distributed features were logarithmically transformed before the ANCOVA). A multiclass significance analysis of microarrays (SAM) was performed with a false discovery rate adjusted each time so that less than one false significant parameter is expected among all significant features. Additionally, to visualize the differences between the groups, heatmaps were created with hierarchical clustering following the Euclidean distance measure and the ward clustering algorithm. For further visualization and identification of significant features a sparse Partial Least Squares – Discriminant Analysis was performed (sPLS-DA) with maximum 5 components each containing a maximum of 10 variables [27]. Additionally, for two of the predictive models that derived from machine learning further visualization and identification of significant features was performed with a principal component analysis (PCA) and with a t-distributed stochastic neighbor embedding (t-SNE) (s. Fig. 8).

2.6.3. Classification – machine learning

For the classification of the data and development of receiver operating characteristic (ROC) curves different machine learning techniques with different software (Scikit-learn Library in Python and MetaboanalystR) were tested. Specifically, with Scikit-learn [28], a recursive feature elimination (RFE) algorithm along with linear support vector machine (SVM) was performed on all data to rank and identify the important variables [29] (s. Supplementary Fig. A2). We then used the One-vs-Rest (OvR) multiclass classification strategy [30] using different classification techniques including the SVM with both linear and nonlinear (radial basis function or RBF) kernels, k-nearest neighbors and random forest. 2/3 of the data were used to train the model and 1/3 to test it with a 3-fold cross validation scheme, which was repeated 100 times with random slices of data into train and test sets to reduce the variance in the classification accuracy and ROC curves. The main parameters for each classifier (e.g., C and gamma in SVM and number of neighbors in kNN) were tuned using an exhaustive search over specified parameter values also known as "grid search" along with cross validation. Balanced accuracy was used as the metric to identify the best parameter set for each classifier. Optimal parameters were then kept fixed for the rest of the analysis. With Metaboanalyst-R [31], groups were reduced to two (healthy vs NAFL-NASH, NAFL vs healthy-NASH and healthy-NAFL vs NASH) and a linear PLS-DA, a linear SVM and random forest were performed, each time using the respective feature ranking method. ROC curves were generated by Monte-Carlo cross validation (MCCV) using balance sub-sampling. 2/3 of the samples were used to evaluate the feature importance and the top 2, 3, 5…max important features were used to build the model which was tested on the 1/3 of the remaining sample. This procedure was repeated multiple times. The classification analysis and development of models were performed for glycans, lipid classes, individual lipids, hormones and combination of glycans, hormones and individual lipids before and after excluding the lean subjects. The same analysis was also performed between subjects with ($n = 10$) vs without fibrosis ($n = 21$). Data sets are available in [32], R command history related to Metaboanalyst-R is available in [33], and code related to OvR-Python in [34].

3. Results

3.1. Anthropometric and clinical characteristics of study population

The anthropometric and clinical characteristics of study population are described in Supplementary Table A1. The healthy group consisted primarily of two subgroups, i.e. lean/slight overweight (BMI < 27.5 kg/m^2) liver-healthy group and overweight/obese (BMI ≥ 27.5 kg/m^2) liver-healthy group. Thus, we have performed two types of analysis to assess whether BMI is relevant in our measurements. In the first, we have compared all liver-healthy subjects (independent of BMI) vs. NAFL vs. NASH and in the second, we have compared only the overweight/obese (BMI ≥ 27.5 kg/m^2) liver-healthy subjects vs. NAFL vs. NASH.

3.2. Lipidomics - differences in blood concentrations of lipids can be used to discriminate between healthy subjects and patients with NAFL and NASH

A lipidomic analysis was performed which identified 365 lipid species belonging in 17 different lipid classes. Hierarchical clustering based on concentrations of lipid classes (total amount for each class) showed that two main clusters are formed: one consisting of healthy subjects and the other consisting of patients with NAFL or NASH (Fig. 1A). Similarly, sPLS-DA analysis (Fig. 1B) showed that the two main components could mostly differentiate the healthy group (red dots clustering at the left side) from the other two groups (green and blue dots for NAFL and NASH, respectively, clustering together at the right). Specifically, serum concentrations of diglycerides (DG), phosphatidylglycerols (PG) and phosphatidic acids (PA) were significantly increased in NAFL and NASH, whereas the concentrations of acylcarnitines (AcCa), cholesterol esters (Che), coenzyme Q10 (Co), lysophosphatidylcholines (LPC) and sphingomyelines (SM) were reduced (Fig. 1C). Except from AcCa, the observed differences in the lipid classes remained significant after excluding the healthy controls with BMI < 27.5 kg/m^2 (Fig. 1D–F).

Given the limited discriminatory ability of lipid classes, we investigated whether the concentrations of individual lipid species can help differentiate between these groups. Significance Analysis of Microarrays (SAM) with corrected FDR for the number of investigated parameters identified 62 lipid species that significantly differ between healthy, NAFL and NASH (Fig. 2A). Most of these lipid species belong to Phosphatidylethanolamines (PE), Phosphatidylcholines (PC) and SM lipid classes (Fig. 2B). Hierarchical clustering showed that lipid species can specifically best discriminate healthy subjects from subjects with NAFL and NASH (Fig. 2C). sPLS-DA showed that with the use of two main components (Fig. 2D and E), each consisting of 10 lipid species, healthy, NAFL and NAFLD subjects converge into separate clusters (Fig. 2D: red dots healthy subjects located left, green dots NAFL subjects located upper right and blue dots subjects with NASH located low right).

The same analysis was repeated after excluding the healthy subjects with BMI < 27.5 kg/m^2 (Fig. 2F-K) from the healthy group (obese healthy vs. NAFL vs. NASH). The number of significant parameters according to SAM was reduced to 32, with similar representation (mainly PE, PC and SM) as previously. Hierarchical clustering demonstrated that lipid species could best differentiate NASH subjects from healthy and NAFL. Similar to the first analysis, sPLS-DA showed robust discrimination between healthy, NAFL and NASH, suggesting BMI-independent differences in lipidome between healthy subjects and patients with NAFL and NASH.

3.3. Fatty acids - differences in fatty acid composition are observed between healthy, NAFL and NASH subjects

The concentrations of 23 fatty acids as relative percent to the total lipid concentration was assessed chromatographically in the NASH, NAFL and healthy status. Serum from 64 subjects was available for these measurements. sPLS-DA analysis showed only a partial discrimination of healthy subjects (red dots) from NAFL and NASH (green and blue dots, respectively) before (Fig. 3A) and after excluding the controls with BMI < 27.5 kg/m^2 (Fig. 3C). This discrimination was achieved by component 1, where five fatty acids were mainly contributing to its composition (Fig. 3B and D). These fatty acids were significantly different between groups also in the univariate analysis (Fig. 3E). Specifically, C16:0 (palmitic acid) was higher in NAFL and NASH compared to healthy population. Additionally, the monounsaturated fatty acids

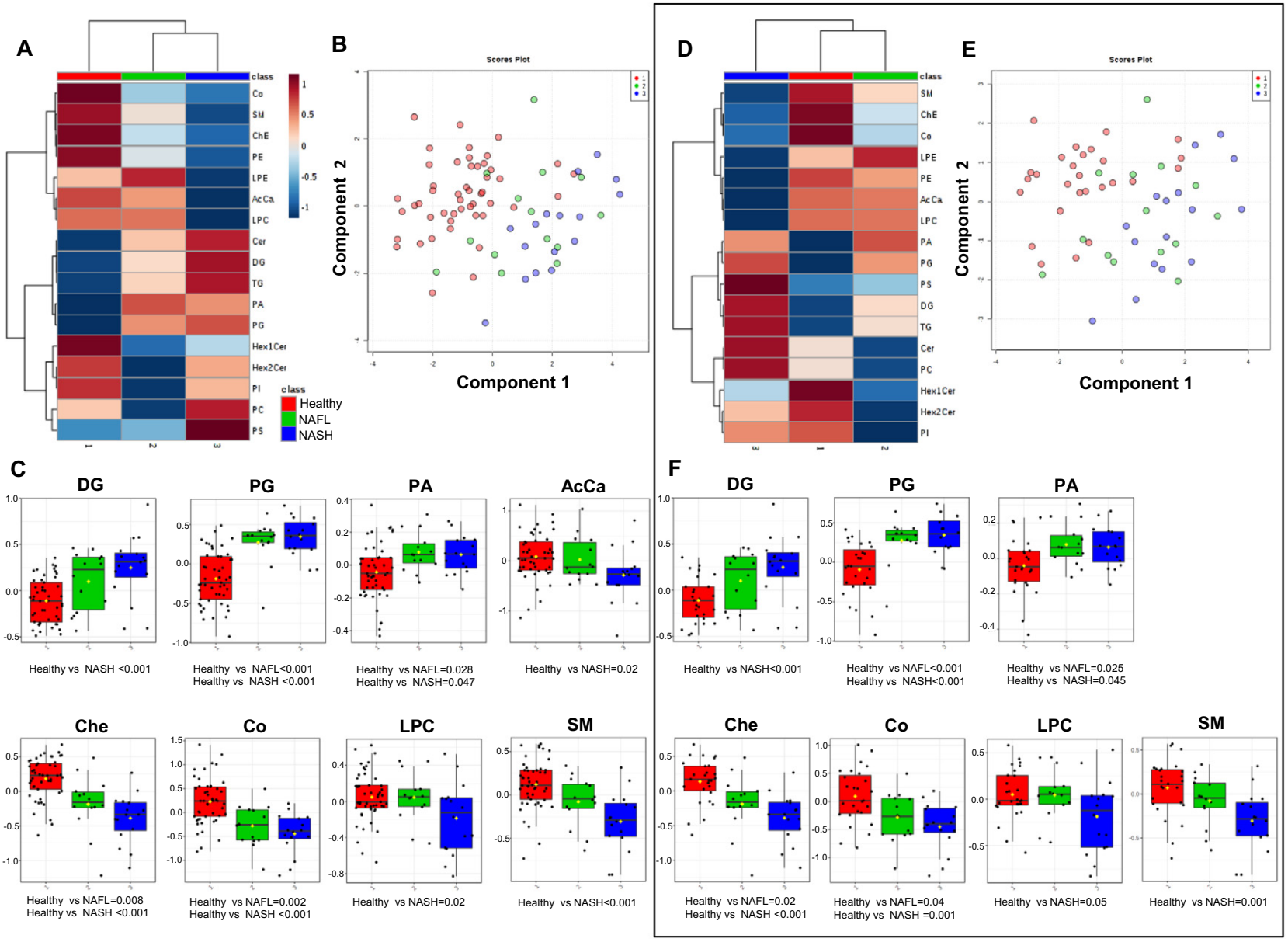


Fig. 1. Profile of circulating lipid classes in healthy vs. NAFL vs. NASH. A. Heatmap of hierarchical clustering displaying the relative concentrations of 17 lipid classes assessed in our study. B. Score plot of component 1 and component 3 of the sPLS-DA analysis (red dots: healthy, green dots: NAFL, blue dots: NASH). C. Boxplots of the significantly different lipid classes between the three groups. Black box (D-F) demonstrates similar analysis when healthy subjects with BMI < 27.5 kg/m^2 were excluded from the control group (overweight/obese control vs NAFL vs NASH): D. Heatmap of hierarchical clustering displaying the relative concentrations of 17 lipid classes. E. Score plot of component 1 and component 3 of the sPLS-DA analysis, F. Boxplots of lipid classes demonstrating significant differences between the three groups. The black dots in boxplots represent the concentrations of the selected feature from each sample. The notch indicates the 95% confidence interval. The yellow diamond indicates the mean concentration of each group. One-way ANOVA followed by Tukey's test was performed for the normally distributed variables and Kruskal-Wallis test followed by Dunn's test was performed for the non-normally distributed ones (1C: DG, PG, Che, 1F: PG).

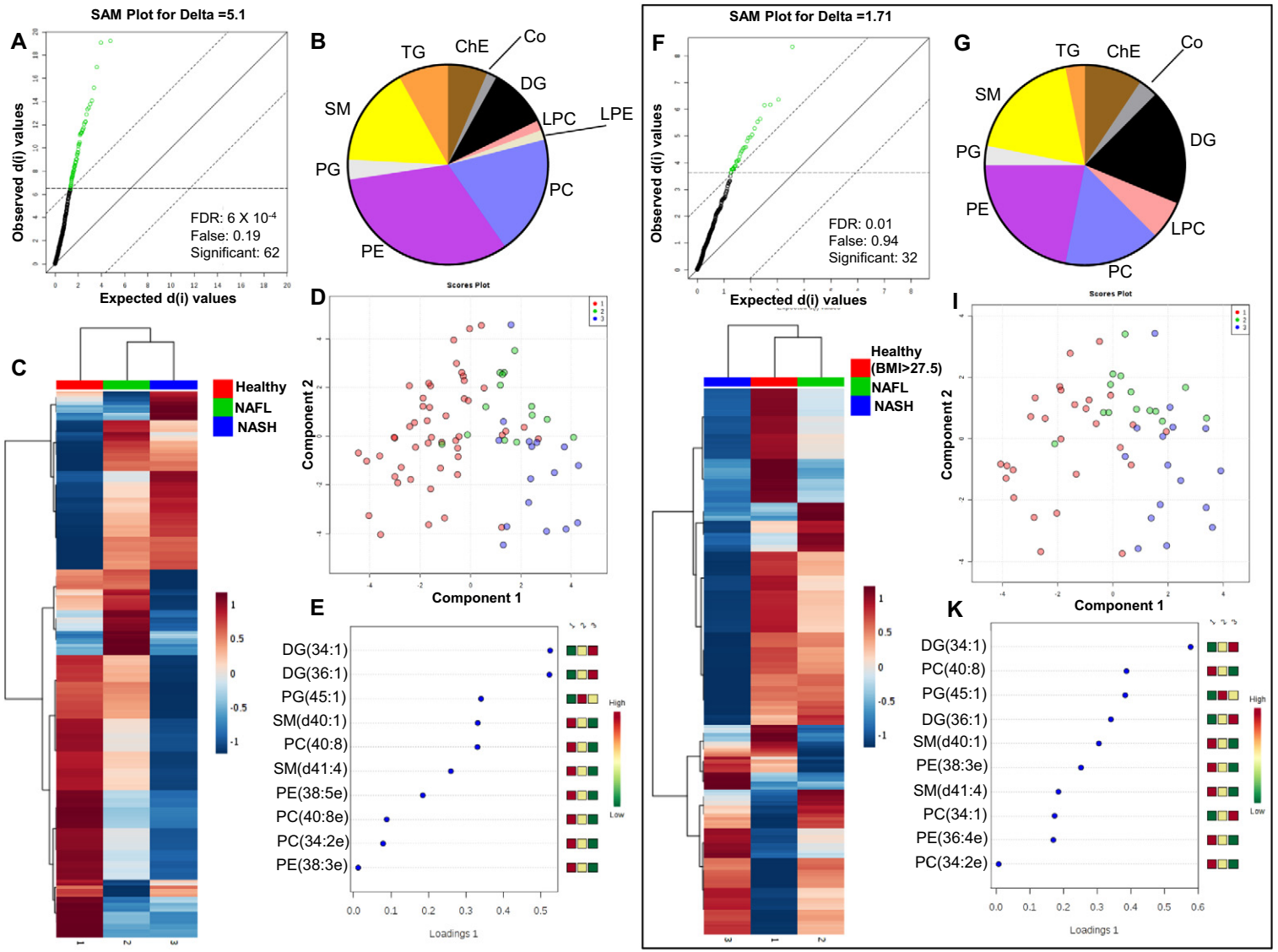
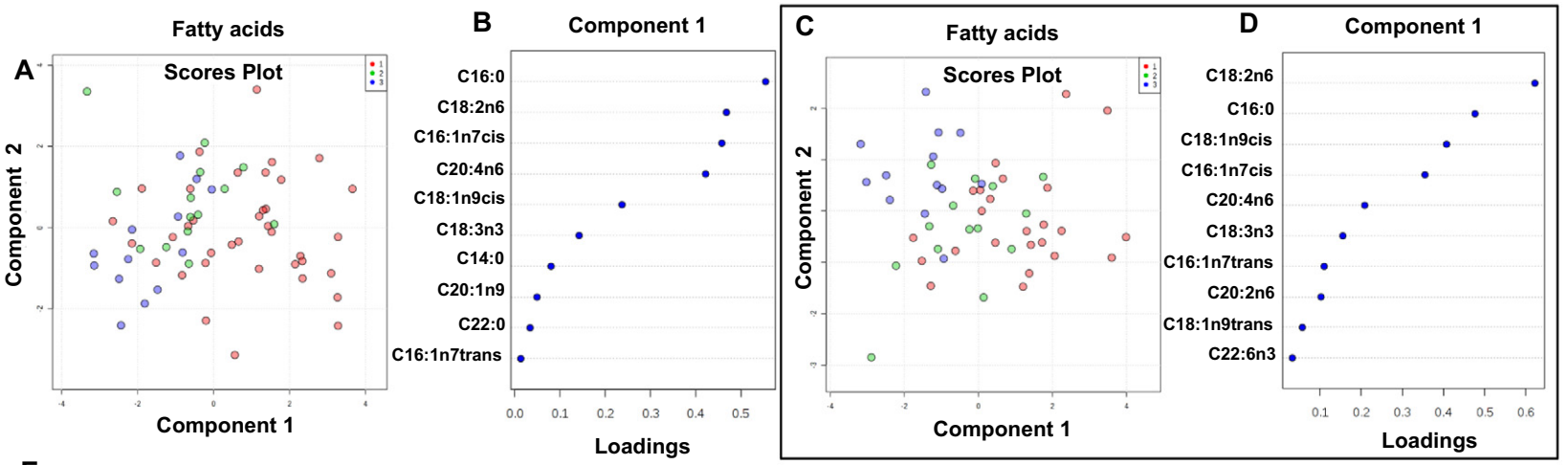


Fig. 2. Profile of circulating lipid species in healthy vs. NAFL vs. NASH. A. SAM identifies 62 lipid species that are significantly different between the three groups with an FDR of 6×10^{-4}. The x-axis represents the expected and the y-axis represents the observed score of the function that depends on a lipid concentration. If the absolute value of (observed–expected) score for a lipid is larger than delta, then this lipid is significantly different between the three groups (healthy, NAFL, NASH), which is marked by a green dot, while the non-significant lipids are marked by black dots. B. Pie chart showing class breakdown of the 62 significant lipid species. C. Heatmap of hierarchical clustering displaying all the lipid species in the three groups. D. Score plot of component 1 and component 2 of the sPLS-DA analysis (red dots: healthy, green dots: NAFL, blue: NASH). E. Loadings of component 1 shown for top 10 ranked lipids. Black box (F-K) demonstrates similar analysis when healthy subjects with BMI < 27.5 kg/m^2 were excluded from the control group (overweight/obese control vs NAFL vs NASH): F. SAM identifies 32 lipid species that are significantly different between the three groups with an FDR of 0.01. G. Pie chart showing class breakdown of the 32 significant lipid species. H. Heatmap of hierarchical clustering displaying all the lipid species. I. Score plot of component 1 and component 2 of the sPLS-DA analysis. K. Loadings of component 1 shown for top 10 ranked lipids.



E

Formula	Name	Class	Healthy (n=37)	NAFL (n=14)	NASH (n=13)	P	P adj.
C14:0	Miristic acid	SFA	0.86 (0.73, 1.24)	1.13 (0.80, 1.62)	1.25 (0.97, 1.46)	.09	.08
C16:0	Palmitic acid	SFA	23.3 (22.1, 25.8)[a,b]	26.0 (25.2, 27.5)[a]	27.2 (25.8, 28.7)[b]	**<.001**	**0.004**
C18:0	Stearic acid	SFA	8.5 (7.8, 9.6)	9.0 (8.0, 9.4)	8.3 (7.9, 9.1)	.89	.78
C20:0	Arachidic acid	SFA	0.07 (0.07,0.09)	0.08 (0.06, 0.08)	0.07 (0.06, 0.08)	.62	.36
C22:0	Behenic acid	SFA	0.06 (0.06,0.07)	0.06 (0.05,0.06)	0.05 (0.05,0.07)	.05	.26
C24:0	Lignoceric acid	SFA	0.06 (0.04,0.10)	0.06 (0.04,0.09)	0.04 (0.04,0.07)	.36	.42
C16:1n7trans	Trans-palmitoleic acid	MUFA	0.62 ± 0.13	0.67 ± 0.11	0.68 ± 0.11	.20	.08
C16:1n7cis	Cis-palmitoleic acid	MUFA	1.54 (1.32, 2.03)[a]	2.02 (1.77, 2.48)	2.51 (1.72, 2.96)[a]	**.002**	**.005**
C18:1n9cis	Oleic acid	MUFA	23.0 (21.5, 26.2)[a]	25.0 (21.9, 26.3)	28.1 (24.9, 29.2)[a]	**.02**	**.003**
C18:1n9-trans	Elaidic acid (trans-oleic acid)	MUFA	1.70 ± 0.19	1.71 ± 0.14	1.80 ± 0.22	.26	.66
C20:1n9		MUFA	0.18 (0.15, 0.28)	0.16 (0.14, 0.17)	0.17 (0.15, 0.20)	.17	.55
C24:1n9	Nervonic acid	MUFA	0.11 (0.07, 0.18)	0.13 (0.08, 0.27)	0.10 (0.06, 0.18)	.63	.54
C18:2n6	Linoleic acid	n-6 PUFA	27.6 ± 4.1[a]	25.4 ± 4.2	22.3 ± 3.3[a]	**<.001**	**<.001**
C18:3n6	gamma-linolenic acid	n-6 PUFA	0.39 (0.30, 0.55)	0.39 (0.29, 0.46)	0.33 (0.27, 0.51)	.53	.90
C20:2n6		n-6 PUFA	0.18 (0.15, 0.21)	0.17 (0.15, 0.20)	0.15 (0.13, 0.19)	.52	.80
C20:3n6		n-6 PUFA	1.54 (1.24, 1.78)	1.45 (1.32, 1.65)	1.57 (1.23, 2.01)	.74	.85
C20:4n6	Arachidonic acid (AA)	n-6 PUFA	5.37 (4.15, 7.08)[a]	4.44 (3.37, 4.94)	4.14 (3.82, 4.50)[a]	**.004**	**.034**
C22:4n6		n-6 PUFA	0.14 (0.11, 0.17)	0.12 (0.10, 0.15)	0.12 (0.10, 0.18)	.27	.46
C22:5n6	n-6 docosapentaenoic acid	n-6 PUFA	0.11 ± 0.03	0.09 ± 0.02	0.11 ± 0.04	.05	.07
C18:3n3	Alpha-linolenic acid	n-3 PUFA	0.32 (0.27, 0.38)	0.26 (0.23, 0.37)	0.27 (0.24, 0.33)	.08	.48
C20:5n3	Eicosapentaenoic acid (EPA)	n-3 PUFA	0.18 (0.15, 0.26)	0.15 (0.10, 0.28)	0.15 (0.12, 0.26)	.34	.25
C22:5n3	n-3 docosapentaenoic acid	n-3 PUFA	0.29 0.25, 0.37)	0.28 (0.24, 0.40)	0.32 (0.25, 0.35)	.92	.87
C22:6n3	Docosahexaenoic acid (DHA)	n-3 PUFA	1.07 (0.87, 1.36)	1.17 (0.71, 1.54)	0.98 (0.81, 1.18)	.57	.70

Fig. 3. Fatty acids concentrations in healthy vs. NAFL vs. NASH. A. Scores Plot of component 1 and component 2 of the sPLS-DA analysis (red = healthy, green = NAFL, blue = NASH). B. Loadings of component 1 showing that mainly the first five fatty acids are contributing to the component. Black box (C-D) demonstrates similar analysis when healthy subjects with BMI < 27.5 kg/m^2 were excluded from the control group (overweight/obese control vs NAFL vs NASH): C. Scores Plot of component 1 and component 2 (red = healthy, green = NAFL, blue = NASH). D. Loadings of component 1 showing that five fatty acids are mainly contributing to the component. E. Table of the 23 fatty acids measured in our study. *p*-values are derived from one-way ANOVA or Kruskal-Wallis test; p adj.-values are derived from one-way ANCOVA after adjusting for BMI (non-normally distributed variables were logarithmically transformed). Same superscript letters indicate significant differences between the two groups based on post-hoc Tukey's or Dunn's test.

C16:1n7cis (cis-palmitoleic acid) and C18:1n9cis (oleic acid) were higher in NASH compared to healthy group, whereas the polyunsaturated linoleic fatty acid (C18:2n6) and arachidonic fatty acid (C20:4n6) were lower (Fig. 3E). These differences between groups remained significant after adjusting for BMI.

3.4. Glycomics – differences in N-glycans are observed between NAFLD vs healthy subjects

GC–MS identified 61 N-glycans in the serum of healthy subjects and patients with NAFL and NASH. Concentrations of seventeen glycans were significantly different between groups (Fig. 4A), while five glycans concentrations were significantly different after excluding the lean controls (Fig. 4C). Main structural characteristics of the glycans were the presence of fucose (10 out of 17) and of sialic acids (13 out of 17). sPLS-DA analysis showed that only two from the five components could partially discriminate the healthy group from the other two (both before and after excluding the healthy controls BMI < 27.5 kg/m^2) but could not discriminate patients with NAFL from those with NASH (Fig. 4B, D).

3.5. Evaluation of predictive accuracy of NASH, NAFL or healthy status with the use of five machine learning methods

We have used two different platforms (Scikit-learn Library in Python and MetaboanalystR) to evaluate the performance of five different machine learning techniques (linear and non-linear SVM, k-nearest neighbor (kNN), linear PLS-DA and random forest) in predicting whether a subject belongs to NASH, NAFL or healthy groups (Table 1). A binary classification method was used to reduce the groups from three to two following the OvR (One-vs-Rest) strategy in Scikit-learn. In MetaboanalystR, three individual classifications were performed, and subjects were grouped as healthy vs. NAFL-NASH, NAFL vs. healthy-NASH and NASH vs. healthy-NAFL. Each omics data set (lipidomics, glycomics, biochemical parameters including four hormones and triglycerides) was analyzed individually as well as in combination aiming to achieve the highest accuracy with a few number of variables. Fatty acids were analyzed only individually since the available fatty acids data were acquired for 64 subjects (due to limited serum availability), whereas the data from other measurements (i.e. glycomics, lipidomics, hormones) were collected for 80 subjects. The highest accuracy was observed for the SVM (with nonlinear radial basis function kernel) method for lipid species both alone and in different combinations with the other omics data (Table 1).

3.6. Development of predictive models for the differentiation between NASH, NAFL and healthy status

Based on the above analysis, nonlinear SVM with RFE was selected as a method to develop ROC curves by using a limited number of lipids, hormonal or glycan variables (Table 2). With the use of 29 different lipid species, a very high predictive accuracy was achieved (Fig. 5A) for all the three groups. Fatty acids and hormonal levels demonstrated high accuracy for discriminating healthy and NASH status but low for predicting NAFL (Fig. 5B and D). Glycans performed better for predicting NAFL compared to fatty acids and hormones, but worse for predicting healthy and NASH (Fig. 5C). Combining 19 lipids and adiponectin (Fig. 5E) or 19 lipids and one glycan (Fig. 5F) led also to very high accuracies and high sensitivities and specificities. Similarly, combinations of nine lipids and adiponectin (Fig. 5G) or nine lipids and one glycan (Fig. 5H) could discriminate with high sensitivity and specificity between the groups, whereas combination of lipids with glycans and hormones led to no further improvement (data not shown).

The same analysis was repeated after excluding subjects with BMI < 27.5 kg/m^2 from the healthy group. This had a small impact on the observed predictive accuracy especially for NAFL group (Fig. 6A) in the model consisting of 29 lipids. However, combinations of lipids and glycans (Fig. 6F) or lipids, hormones and glycans (Fig. 6H) could discriminate very efficiently between groups.

3.7. Exploratory analysis of the discriminatory and predictive potentials of lipidomics, hormones, glycomics and fatty acids in liver fibrosis

Among the 31 subjects with NAFLD, 21 had fibrosis in the histology according to Kleiner fibrosis score and 10 did not. We performed an exploratory analysis of these two groups (fibrosis “yes” or “no”), which showed that similar to what we observed in the whole study population, lipidomics primarily were able to discriminate robustly (with 95% sensitivity and 99% specificity) between the two groups whereas glycans and fatty acids had less predictive potentials (Fig. 7 and Table 3).

4. Discussion

We report novel combinations of glycans, lipids and hormonal variables that can diagnose simultaneously and with high accuracy the presence of NASH, NAFL or healthy status.

Additionally, we report a combination of lipids that can diagnose the presence of liver fibrosis. We obtained excellent classification performance, which we believe can be improved further. Both diagnostic models should be further trained and validated in large independent cohorts both cross-sectionally and, more importantly, prospectively.

Several studies have aimed to develop non-invasive diagnostic scores for advanced fibrosis, whereas fewer efforts have focused on the diagnosis of NASH [6,35]. Among them, circulating cytokeratin-18 fragment (Ck-18f), has been the most investigated biomarker so far, with two meta-analyses suggesting an AUROC of 0.82 with sensitivity ranging from 66%–78% and specificity of 82%–97% for NASH diagnosis [36,37]. The low sensitivity, observed especially in studies including multiethnic populations [38], the variations in proposed diagnostic cut-off levels and the lack of a widely available assay for clinical purposes have resulted in the limited use of CK-18f in the clinical setting. Several combinations of serum biomarkers and/or clinical variables have also been proposed (reviewed in [11]), but none of them has reported AUROCs as high as in our study. Furthermore, many of these studies were performed in morbidly obese individuals thus, may have limited applicability to the general population.

Our study, using a stepwise approach and multiple machine learning methods, has focused on variables that are key players in pathophysiological mechanisms related to the development of steatosis and NASH such as changes in lipidome representing the abnormal lipid metabolism [39], in glycans reflecting the ability of the liver to synthesize proteoglycans [12] and in hormones associated with the increased IR and abnormal glucose homeostasis commonly observed in these patients [40]. Glycomics is a novel subspecialty in omics systems sciences that offers substantial promise for next-generation biomarkers on disease susceptibility, drug target discovery, and precision medicine. Glycosylation disorders have been reported in NAFLD and loss of glycosylation of crucial uptake and efflux transporters in patients with NASH could influence transporter function and contribute to changed drug disposition [41].

Most importantly, in our study, we have followed a novel approach that has several major advantages: First, we have comparatively tested several supervised learning methods to identify the most accurate to predict among the groups. Previous studies primarily used one classification method (mostly based on logistic regression analysis) and features were selected in many cases through univariate tests (*t*-tests, Welch's tests, individual AUC scores) [42,43]. Such methods, however, tend to be overly optimistic and are highly susceptible to overfitting, thus, it is difficult to be reproduced. Among the methods we have tested, SVM with RFE has demonstrated the highest accuracy. RFE method removes the redundant features and builds a predictive model only with the highly scored variables, which helps avoid overfitting. Since

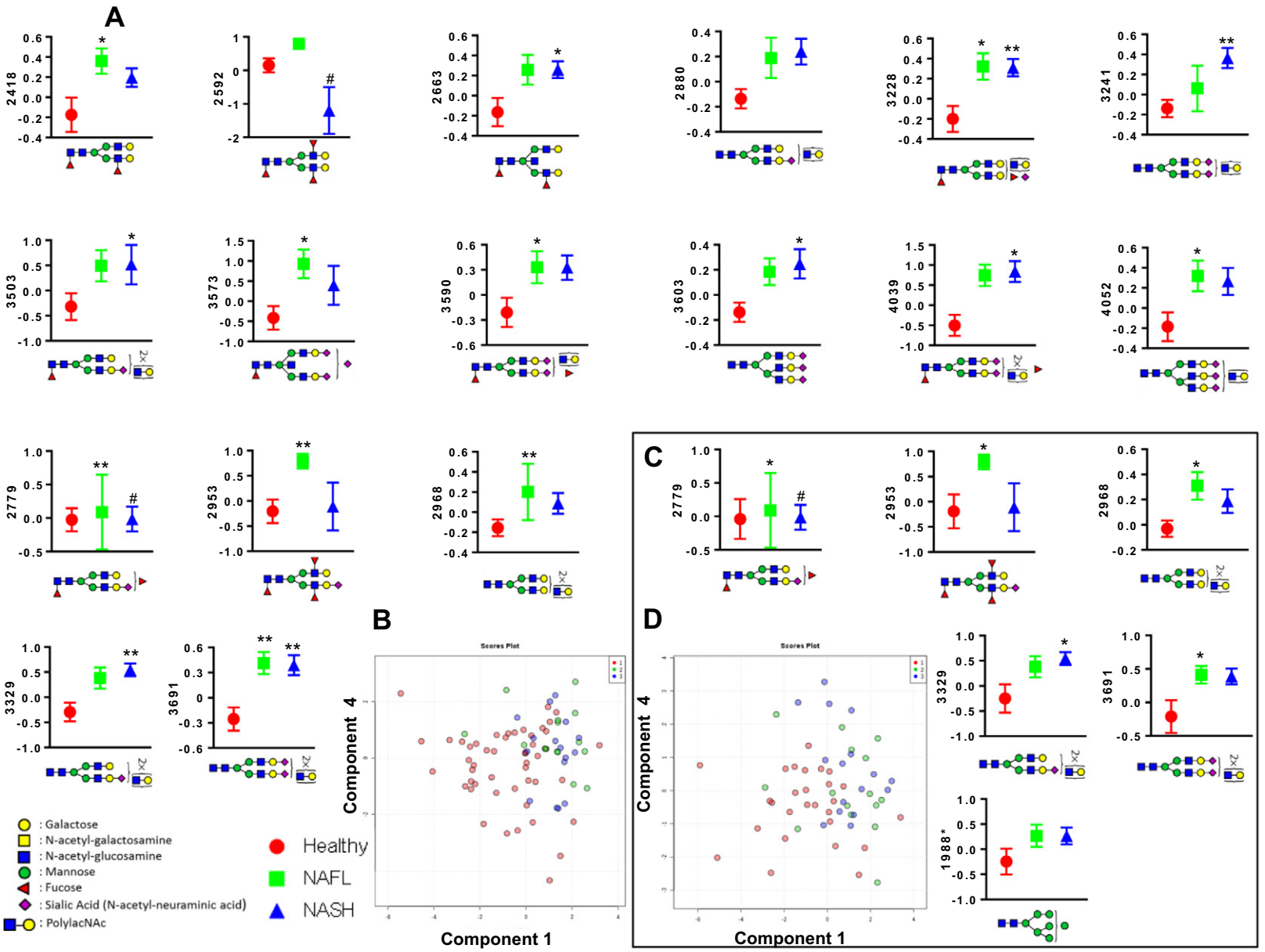


Fig. 4. N-glycan profile in healthy vs. NAFL vs. NASH. A. Glycans with significantly different concentrations between the three groups. B. Score plot of component 1 and component 4 from the sPLS-DA analysis (red = healthy, green = NAFL, blue = NASH). Black box (C-D) demonstrates similar analysis when healthy subjects with BMI < 27.5 kg/m^2 were excluded from the control group (overweight/obese control vs NAFL vs NASH): C. Glycans with significantly different concentrations between the three groups. D. Score plot of component 1 and component 4 from the sPLS-DA analysis (red = healthy, green = NAFL, blue = NASH). x-axis shows the most likely structure of the glycan and y-axis its mass (m/z) according to mass spectrometry. *,**,***, $p < .05$, <0.01 and < 0.001 in the post hoc t-test between the group with the star (NAFL or NASH) vs. healthy. #, $p < .05$ in the post-hoc t-test between NASH vs. NAFL. The 2953 and the 1988 glycans were participating regularly in the predictive models.

Table 1
Predictive accuracies (%) of five machine learning techniques for multiclass (OvR) and binary classification using two different software.

Models (A)	All subjects								
	OvR (Scikit-learn)					Binary (MetaboanalystR)			
	Variables (n)	SVM (linear)	kNN	SVM (RBF)	Random forest	Variables (n)	PLS-DA (linear)	Random forest	SVM linear
1. Lipid classes	4	58 ± 8	50 ± 8	63 ± 10	56 ± 10	3	84/58/75	80/60/73	76/51/72
2. Lipid species	29	86 ± 9	65 ± 10	88 ± 7	65 ± 9	50	77/55/70	83/60/80	81/58/76
3. Hormones	4	54 ± 8	49 ± 8	55 ± 9	54 ± 9	4	86/60/80	80/57/76	85/54/77
4. Glycans	5	57 ± 9	45 ± 7	54 ± 10	49 ± 9	5	54/44/58	60/60/51	59/49/53
5. Fatty acids	5	50 ± 9	43 ± 8	54 ± 11	45 ± 10	5	61/48/57	71/49/63	66/48/62
6. Lipid species + Glycans	10	80 ± 8	67 ± 10	78 ± 9	66 ± 9				
7. Lipid species + Glycans	20	86 ± 8	67 ± 10	82 ± 9	66 ± 10				
8. Lipid species + Hormones	10	74 ± 10	59 ± 9	76 ± 9	67 ± 10				
9. Lipid species + Hormones	20	80 ± 10	64 ± 10	87 ± 8	65 ± 9				
10. Lipid species + Hormones + Glycans	10	77 ± 9	71 ± 11	77 ± 9	68 ± 10				
11. Lipid species + Hormones + Glycans	20	83 ± 8	63 ± 9	86 ± 7	65 ± 9				
Models (B)	All subjects (excluded controls with BMI < 27.5 kg/m^2)								
	OvR (Scikit-learn)					Binary (MetaboanalystR)			
	Variables (n)	SVM (linear)	kNN	SVM (RBF)	Random forest	Variables (n)	PLS-DA (linear)	Random forest	SVM (linear)
1. Lipid classes	4	59 ± 9	49 ± 9	55 ± 9	56 ± 10	5	77/50/67	77/53/68	77/52/69
2. Lipid species	29	84 ± 9	70 ± 11	88 ± 9	64 ± 9	50	70/52/68	79/49/77	75/53/72
3. Hormones	4	52 ± 8	48 ± 9	51 ± 9	49 ± 9	4	79/51/72	74/47/68	78/50/69
4. Glycans	5	62 ± 10	54 ± 9	62 ± 9	49 ± 10	5	58/41/63	52/53/48	55/49/53
5. Fatty acids	5	51 ± 10	42 ± 10	51 ± 12	50 ± 9	5	57/43/57	74/43/62	61/41/59
6. Lipid species + Glycans	10	71 ± 10	67 ± 10	71 ± 10	65 ± 11				
7. Lipid species + Glycans	20	88 ± 7	67 ± 10	90 ± 9	64 ± 11				
8. Lipid species + Hormones	10	78 ± 9	73 ± 10	79 ± 10	67 ± 11				
9. Lipid species + Hormones	20	86 ± 7	72 ± 10	89 ± 7	68 ± 10				
10. Lipid species + Hormones + Glycans	10	82 ± 9	75 ± 9	83 ± 9	68 ± 10				
11. Lipid species + Hormones + Glycans	20	83 ± 8	74 ± 10	85 ± 9	69 ± 9				

For the OvR (Scikit-learn) % of predictive accuracy ± SD of the cross validations is reported. For the Binary (MetaboanalystR) predictive accuracy of healthy vs NAFL-NASH/NAFL vs healthy-NASH/NASH vs healthy-NAFL is reported. Number of variables included each time in the model is reported. For the Binary (MetaboanalystR) analysis the number of variables selected was the one that achieved the highest accuracy.

our study included a small number of subjects, we have additionally used the biggest part of the population to train the model but also a small part (through balanced subsampling) to validate it. Second, our predictive models are able to diagnose between three (and not two) conditions i.e., healthy status, NAFL and NASH. Thus, they allow the evaluation of a subject with a single blood draw without the need of an imaging modality for diagnosing NAFL. This concept is especially attractive for routine screening of the population at high risk for the development of the disease such as obese patients or patients with type 2 diabetes in a primary care setting and without the need of specially trained personnel. Additionally, our approach can be used to build a similar model for differentiating between "none vs. non-advanced vs. advanced fibrosis". Importantly, we have additionally developed models that can differentiate healthy status, NAFL and NASH among overweight and obese people, taking into consideration that this population is not only more probable to have NASH, but is also more likely to be tested by doctors. We recognize, though, that despite significant progress, mass-spectrometric methods are still not widely available, cost-effective and optimally standardized [7]. Thus, we are reporting different models combining a variety of detection methods (mass-spectrometry, chromatography, ELISA) and with variable number of parameters for which accurate cost-effectiveness has to be assessed in the future. Despite the above issues, these models seem to be more cost-effective than liver biopsy. Specifically, the costs of liver biopsy have been reported to be at £956 (i.e. $1153) in 2012–2013 [44], whereas in the US currently the costs of liver biopsy vary between $2000–8000 (personal communication). The costs of the model combining lipids, glycans and adiponectin in our study was $605/subject, thus significantly lower compared to liver biopsy. We are also expecting the costs to be further reduced as mass-spectrometric methods are becoming widely available and if targeted (instead of untargeted) measurements of lipids and glycans are performed. Finally, we believe that these models may be further improved or simplified by including targeted biochemical, genetic and clinical parameters. The best candidates are liver transaminases, platelets, age, BMI and PNPLA3 genotype [11]. These variables were not included in the current study since some of them were used as inclusion criteria to organize the study groups (e.g. liver transaminases and BMI), whereas no material was available for others (i.e. DNA for PNPLA3 genotype).

Regarding the variables that were selected by our models, the most robust differences and consequently the most profound contribution to the predictive algorithms were observed in lipid species. Previous lipidomic descriptive studies in NAFLD reported higher DG and TG and lower PC levels in the liver and in serum in NAFLD as well as significant changes in fatty acid composition [9,10,45]. A recent study suggested also that changes in specific triglyceride groups can differentiate between NAFL and NASH [46]. In line with these findings, we observe in our study higher total DG, a trend to higher total TG and lower levels in many of the PC species in NAFL and NASH. Several of these variables are also selected by our predictive models. Additionally, similar to a previous lipidomic analysis [10], we observed a progressive increase of palmitoleic (16:1n7) and oleic acid (18:1n9), as well as a progressive decrease of linoleic acid (18:2n6) from healthy to NAFL and NASH. Unfortunately, quantification of the total amount of these fatty acids was only possible to a subset of the study population, which affected their performance as diagnostic predictors in our models. Furthermore, we identified changes in TG species that can particularly contribute to the diagnosis of liver fibrosis (fibrotic model). Recent data indicate that inhibition of diacylglycerol acyltransferase-2, one of the two enzymes that catalyze the final step of TG synthesis, is considered to treat patients

Table 2
Top-ranked variables considered in different predictive models of NASH vs. NAFL vs. healthy.

All subjects					
Lipids (29 variables)			Fatty acids (5 variables)	Glycans (5 variables)	Hormones (4 variables)
AcCa(10:0)	PC(34:2e)	PE(38:6)	C18:2n6	1661	Leptin
Cer(d34:2)	PC(35:3)	PI(36:1)	C18:3n6	2592	Adiponectin
DG(34:1)	PC(36:4)	SM(d32:0)	C20:0	2764	Activin A/Follistatin
DG(36:4)	PC(36:5e)	SM(d32:2)	C20:4n6	2968	Triglycerides
LPC(20:0e)	PC(37:2)	SM(d40:1)	C22:4n6	4039	
LPC(22:5)	PC(40:6e)	TG(38:0)			
LPE(16:0)	PC(40:7)	TG(38:2)			
PC(32:0)	PC(40:8)	TG(43:1)			
PC(32:1e)	PC(42:6)	TG(53:5)			
PC(34:0)	PE(38:1)				
Lipids + Hormones (20 variables)		Lipids + Glycans (20 variables)		Lipids + Hormones (10 variables)	Lipids + Glycans (10 variables)
AcCa(10:1)	PC(36:5e)	Cer(d18:2_25:1)	PC(40:5e)	DG(34:1)	LPC(22:5)
Cer(d43:0)	PC(37:2)	LPC(20:0e)	PC(40:7)	LPC(20:0e)	LPE(16:0)
DG(34:1)	PC(37:3)	LPC(22:5)	PC(42:6)	LPE(18:0)	PC(34:1)
DG(36:4)	PC(40:8)	LPE(16:0)	PE(38:1)	PC(34:2e)	PC(34:1e)
LPC(20:0e)	PE(34:1)	PC (32: 1e) ≥ *PC(16:0e_16:1)*	PE(38:6)	PC(36:4)	PC(35:3)
LPC(22:5)	PE(38:1)	PC (32:1e) ≥ *PC(16:1e_16:0)*	SM(d35:1)	PC(36:5e)	PC(42:6)
LPE(18:0)	SM(d40:1)	PC(34:0)	SM(d38:1)	PC(37:3)	PE(38:1)
PC(34:2e)	TG(38:0)	PC(34:1)	TG(40:0)	PC(40:8)	SM(d35:1)
PC(36:3)	TG(38:2)	PC(34:1e)	TG(52:4)	PE(38:1)	SM(d40:1)
PC(36:4)	Adiponectin	PC(35:3)	2592	Adiponectin	2592

All subjects (excluded controls with BMI < 27.5 kg/m^2)					
Lipids (29 variables)			Fatty acids (5 variables)	Glycans (5 variables)	Hormones (4 variables)
AcCa(10:0)	PC(34:1)	PE(40:4e)	C16:0	1988	Leptin
Cer(d34:2)	PC(35:3)	PI(36:1)	C18:1n9cis	2592	Adiponectin
DG(34:1)	PC(36:5e)	SM(d32:0)	C18:2n6	2764	Activin A/Follistatin
DG(36:4)	PC(37:2)	SM(d32:2)	C18:3n6	2968	Triglycerides
Hex1Cer(d34:2)	PC(40:7)	SM(d36:0)	C22:4n6	3953	
LPC(20:0e)	PC(40:8)	SM(d36:4)			
LPC(22:5)	PC(42:6)	SM(d37:1)			
LPE(18:0)	PE(34:1)	SM(d40:1)			
LPE(22:5)	PE(34:3e)	TG40:0)			
PA(44:4)	PE(38:1)				
Lipids + Hormones (20 variables)		Lipids + Glycans (20 variables)		Lipids + Hormones (10 variables)	Lipids + Glycans (10 variables)
Cer(d34:0)	PC(36:5e)	Cer(d34:0)	PC(36:2)	LPC(20:0e)	Cer(d43:3)
DG(36:4)	PC(37:2)	Cer(d43:3)	PC(40:7)	LPC(20:3)	DG(34:1)
LPC(20:0e)	PC(40:7)	DG(34:1)	PC(42:6)	LPE(16:0)	DG(36:2)
LPC(20:3)	PC(40:8)	DG(36:4)	PE(34:1)	PC(34:1)	LPC(20:0e)
LPE(16:0)	PE(38:1)	LPC(20:0e)	PE(34:3e)	PC(34:2e)	LPC(22:5)
LPE(18:0)	PE(40:4e)	LPC(22:5)	CO(36:1)	PC(36:5e)	LPE(16:0)
LPE(22:5)	SM(d32:0)	LPE(16:0)	TG(38:0)	PC(40:8)	PC(35:2)
PC(34:0)	SM(d36:0)	PC(34:0)	TG(38:2)	PE(38:1)	PC(40:7)
PC(34:1)	SM(d40:1)	PC(34:1)	1988	SM(d32:0)	PE(34:1)
PC(34:2e)	Adiponectin	PC(35:2)	2592	Adiponectin	PE(34:3e)
Lipids + Hormones + Glycans (20 variables)			Lipids + Hormones + Glycans (10 variables)		
DG(34:2)	PC(36:4)		LPC(22:5)		
DG(36:4)	PC(40:8)		LPE(16:0)		
LPC(20:0e)	PE(34:1)		PC(34:1)		
LPC(22:5)	PE(40:4e)		PC(40:8)		
LPE(16:0)	PI(36:1)		PE(34:1)		
LPE(18:0)	SM(d35:1)		PI(36:1)		
PA(44:4)	SM(d38:0)		SM(d38:0)		
PC(34:1)	SM(d40:1)		SM(d40:1)		
PC(34:2e)	Adiponectin		Adiponectin		
PC(36:3)	2592		2592		

with NAFLD/NASH [47,48]. We subsequently tested whether including adiponectin, leptin, follistatin and activin A (measured all through validated immunoassays) could improve the accuracy of the models or maintain high accuracy with lower number of variables. We have selected these hormones based on their involvement in glucose regulation, IR, lipid metabolism and their associations with obesity, type 2 diabetes [17–19] and NAFLD [18,20–22]. Among them, only adiponectin is contributing to models combining lipids with hormones or lipids with hormones and glycans. Of note, adiponectin improves the accuracy of our predictive models even after excluding the subjects with BMI < 27.5 kg/m^2, suggesting that its relation to NAFL and NASH is not explained just by the higher weight and fat mass of these patients. Adiponectin, beyond its weight- and glucose-regulatory effects exerts anti-steatotic, anti-inflammatory and anti-fibrotic actions in the liver, suggesting a protective role against development and progress of NAFLD [18,20–22]. In line with the experimental results, in a meta-analysis of 27 studies adiponectin levels progressively decreased from controls to subjects with NAFL and to subjects with NASH independently of BMI, age, sex and the presence of T2D [22]. Adiponectin together with leptin and ghrelin have been previously used to develop a

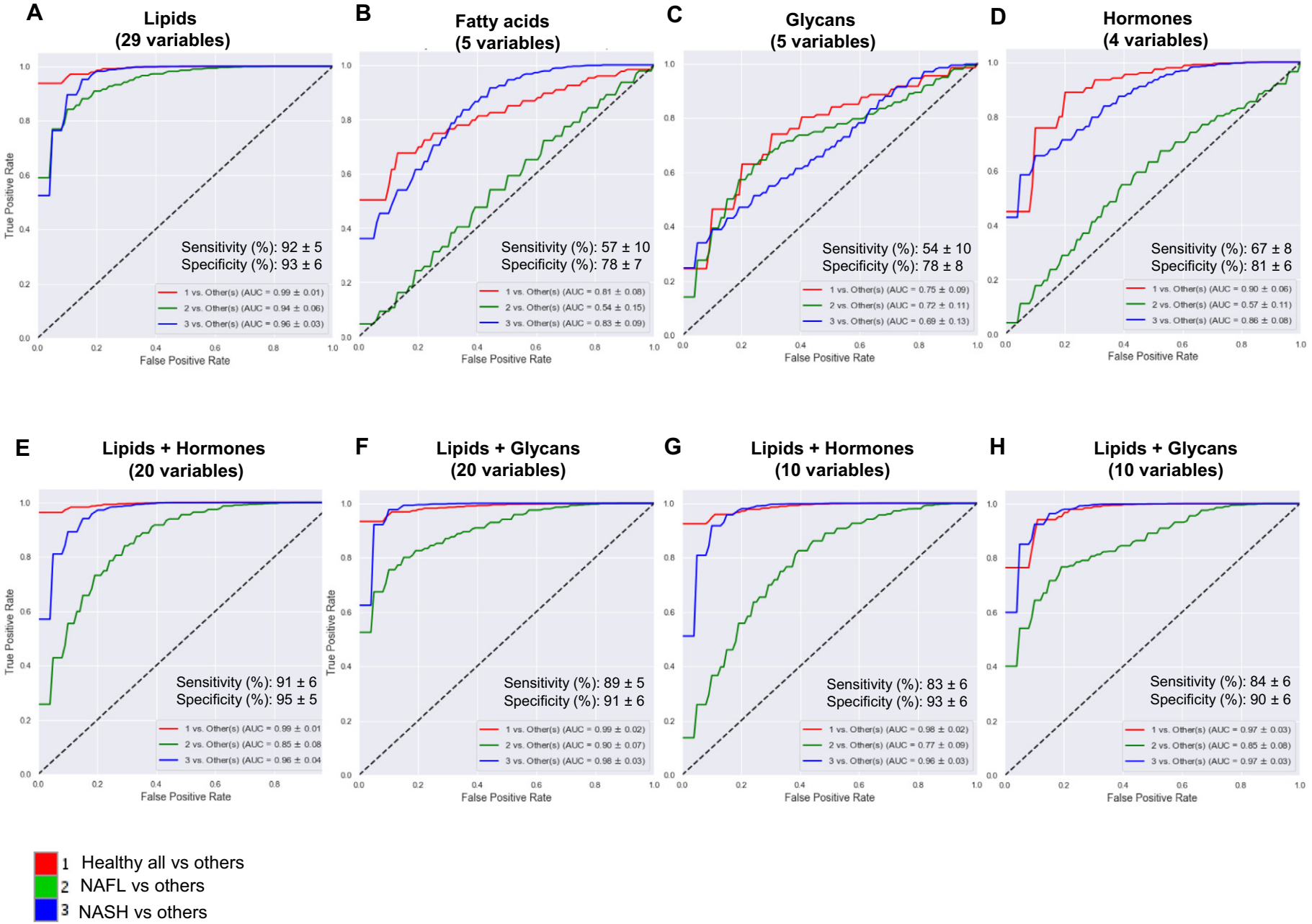


Fig. 5. ROC Curves for the classification of healthy vs. NAFL vs. NASH. A-D. ROC curves for single lipids, fatty acids, glycans and hormones data set computed using the nonlinear SVM classifier. E-H. ROC curves for the integrated lipids with hormones and lipids with glycans data sets computed using the nonlinear SVM classifier. The number of variables used to produce each curve is indicated in parenthesis.

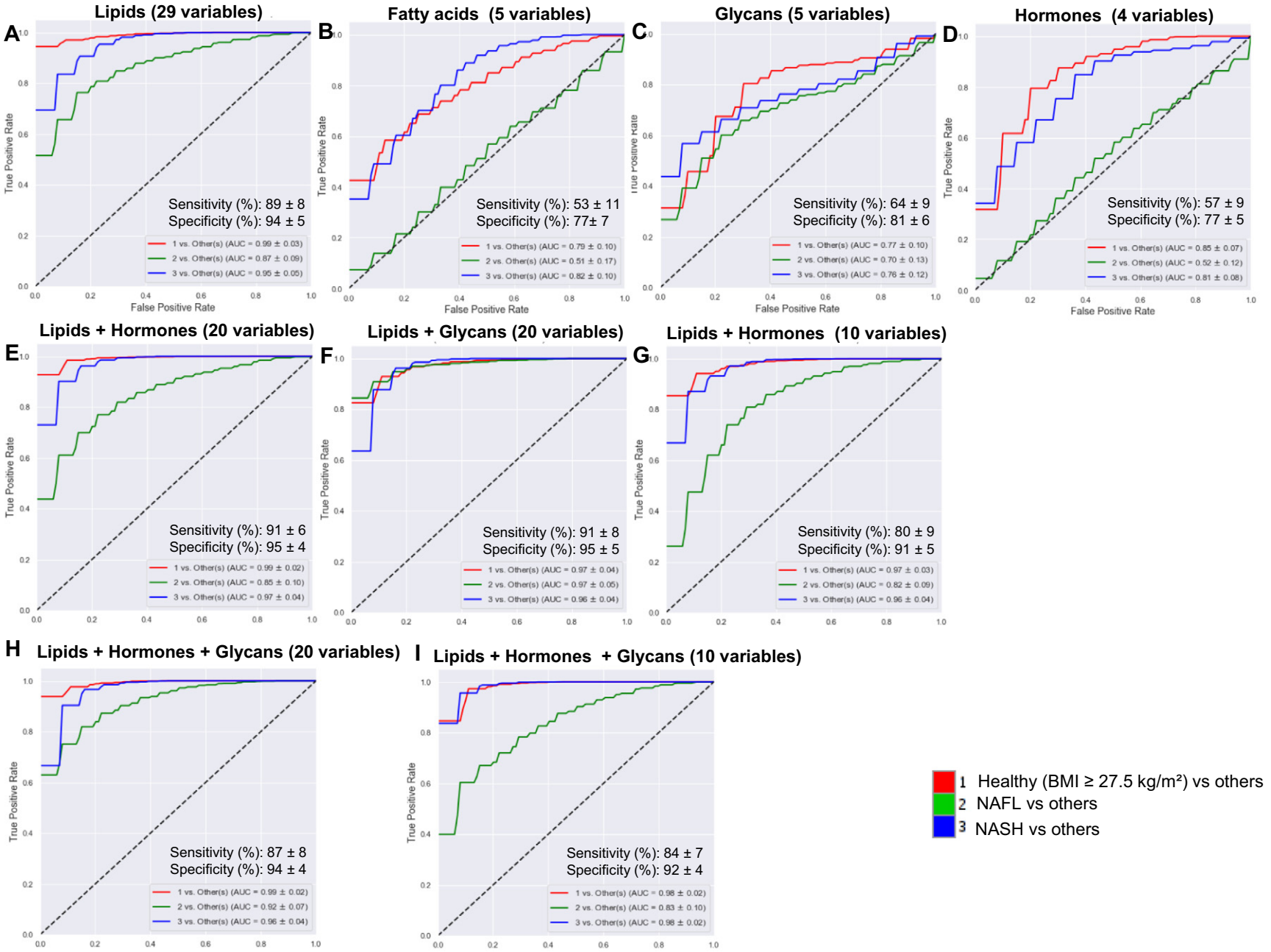


Fig. 6. ROC Curves for the classification of healthy (obese) vs. NAFL vs. NASH. A-D. ROC curves for single lipids, fatty acids, glycans and hormones data set computed using the nonlinear SVM classifier. E-I. ROC curves for the integrated lipids with hormones, lipids with glycans and lipids with hormones and glycans data sets computed using the nonlinear SVM classifier. The number of variables used to produce each curve is indicated in parenthesis.

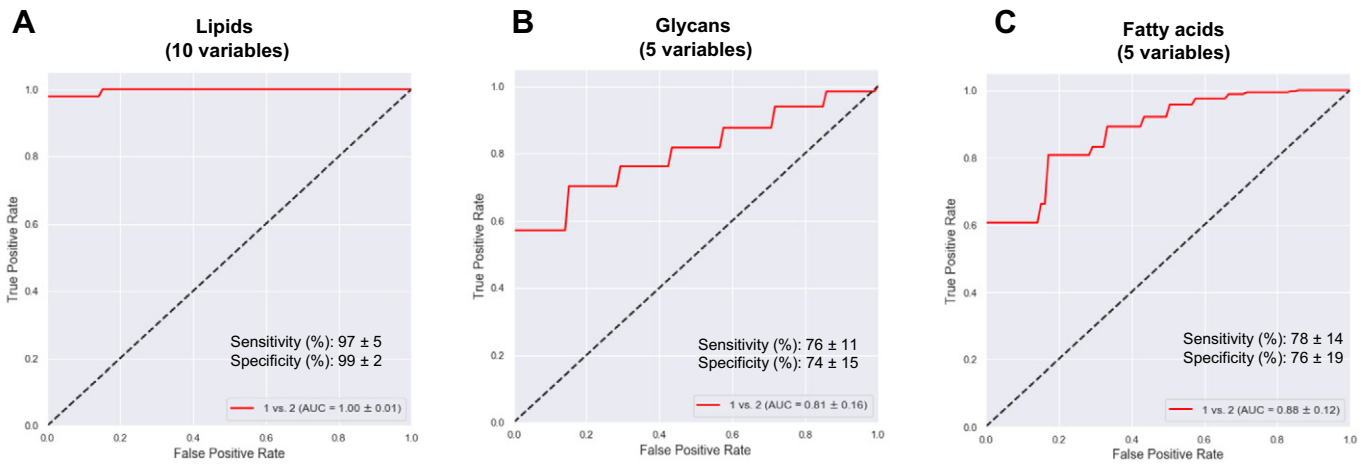


Fig. 7. ROC Curves for the classification of fibrosis status. A-C. ROC curves for single lipids, glycans and fatty acids data sets in subjects with liver fibrosis ($n = 21$) vs. subjects without liver fibrosis ($n = 10$).

diagnostic model for NASH, which demonstrated lower sensitivity and specificity (AUROC of 0.789) compared to the present models [49].

Lastly, the N-glycome profile has been assessed in very few studies and with less advanced methods [50–52]. In one of them, eight out of the 41 detected glycans by glycoblotting have been reported to be significantly different in NAFL vs. NASH in a Japanese population [52]. A maximum AUROC of 0.866 was achieved by analyzing each glycan individually [52]. In another study, twelve glycans were detected by glycome mapping on DNA sequencing equipment in serum and only one was significantly different between NAFL and NASH [50]. In a more recent work, with the use of immunoprecipitation followed by MALDI-TOF MS higher serum levels of a specific alpha-1 antitrypsin glycosylation were reported in NASH, a result that was derived by the comparison of 5 NAFL vs. 6 NASH subjects [51]. In our study, with the use of standardized GC–MS methods, 61 glycans were detected among which 17 were significantly different between the groups. However, the differences were observed mainly between healthy and NAFLD patients and in many cases diminished after excluding the subjects with BMI < 27.5 kg/m^2 from the control group, suggesting that they are primarily weight-related. Most of the significant glycans were core- or multi-fucosylated and many of them contained sialic acids. The levels of sialic acids have been associated with metabolic syndrome [53] as well as NAFLD [54,55]. Increased core-fucosylation of proteins has been associated with liver fibrosis as well as HCC with a core-fucosylated form of alpha-fetoprotein being approved by the FDA as a biomarker of HCC [56,57]. Using a recent algorithm for diagnosing advanced liver fibrosis in NASH, fucose levels were one of the ten metabolites included in the diagnostic panel [35]. In our diagnostic models of healthy, NAFL and NASH, which combine lipids, glycans and/or hormones, a multi- and core-fucosylated glycan (2592) was often selected, along with a second agalactosylated non-fucosylated glycan (1988) in some cases. In the exploratory analysis for the presence of liver fibrosis, 4 out of 5 selected glycans were core-fucosylated. Nevertheless, N-glycome profile is strongly related to body weight, which can help improve the diagnostic accuracy of NAFL subjects in our models, but may be especially useful for the non-invasive discrimination between different fibrosis stages. Of note, serum fucosylated haptoglobin appears to be a useful glyco-biomarker of liver fibrosis and a predictor of HCC in patients with chronic hepatitis C, a prominent feature of which is NAFLD [58,59].

Our study has both strengths and limitations. First, the sample size was not large, albeit adequate to prove our hypotheses. A validation cohort was not available, but as described above, we have used balanced-subsampling and cross-validation. With this approach, 2/3 of the study population is used for training the model and 1/3 of the population is used to validate the model. This procedure is repeated multiple times with different slices of the data and serves as a control against random results. Another limitation is that the population under study was generally homogenous (white people from northern Greece) and consequently, our results should be further tested in larger and more heterogeneous cohorts and in different populations. Additionally, we have not accounted for variables that may be useful for enhancing the predictive capabilities of our models such as PNPLA3 genotype, age, BMI, decompensated diabetes, liver function tests, and ethnicity. On the other hand, our study integrated measurements from different omics procedures, which were evaluated with several supervised learning methods. Given the exploratory character of our study, we are also reporting several models that include a variety of variables. These models should be further trained by larger cohorts from different populations.

In conclusion, we propose several non-invasive models consisting of lipids, hormones and glycans that can diagnose with very high accuracy (>90%) and simultaneously the presence of NASH, NAFL or healthy status. We additionally propose a combination of lipids that can diagnose with very high accuracy (97%) the presence of liver fibrosis. These models may serve as a low-risk cost-effective alternative method to liver biopsy for diagnosing and staging NAFLD. Future studies should aim to further improve these models and validate them in large independent cohorts both cross-sectionally as well as prospectively.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.metabol.2019.154005.

Table 3
Top-ranked variables considered in the predictive models of fibrosis.

Lipids	Glycans	Fatty acids
DG(36:3)	2081	C14:0
LPC(18:0)	2315	C18:1n9trans
PC(36:2)	2592	C18:3n3
PC(37:2)	2605	C20:3n6
PC(40:5)	4587	C22:6n3
TG(38:0)		
TG(50:0)		
TG(51:1)		
TG(57:1)		
TG(60:2)		

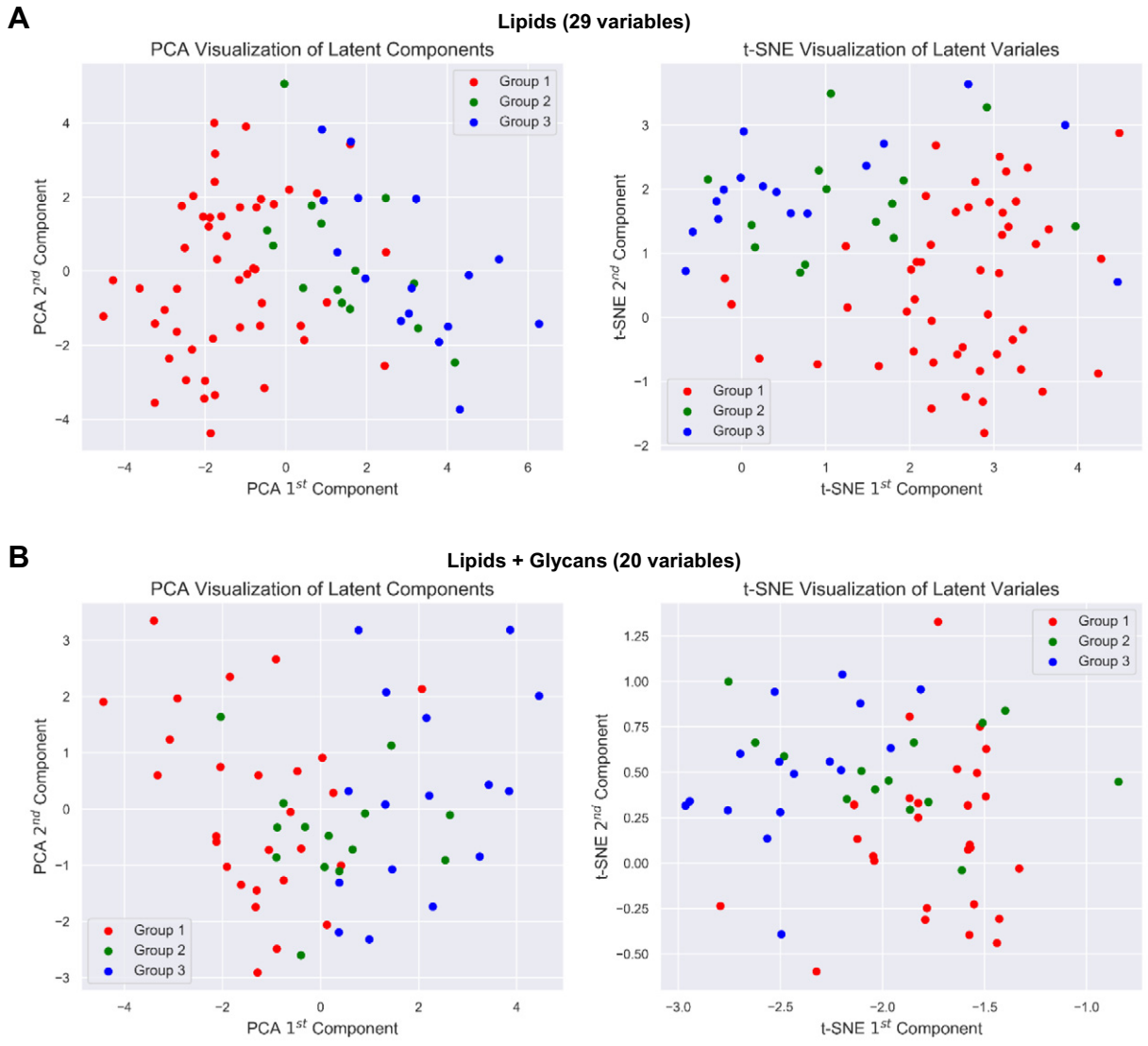


Fig. 8. PCA and tSNE visualizations of the samples in the projected space. A. PCA projection in the subspace spanned by the first two components as well as two-component tSNE visualizations for the reduced lipids data that contain 29 variables. The best ROC curves and classification of healthy vs. NAFL vs. NASH were achieved for this data set with 80 subjects. B. PCA projection in the subspace spanned by the first two components as well as two-component tSNE visualizations for the reduced lipids + glycans data that contain 20 variables (18 lipids + 2 glycans). The best ROC curves and classification of healthy (obese) vs. NAFL vs. NASH were achieved for this data set with 57 subjects. Group 1: Healthy, Group 2: NAFL, Group 3: NASH.

Acknowledgments

Author contributions

C.S.M designed the experiment. All authors contributed to the performance of the experiments and acquisition of the data. NP and AY analyzed the data. NP and AY wrote the manuscript with input from all the authors.

Funding

The current study was funded by NIH K24DK081913. NP was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) –389891681 (PE 2431/2-1). AS-V is recipient of the Instituto de Salud Carlos III Miguel Servet fellowship (grant CP II 17/00029).

Declaration of competing interest

CSM is consultant to Intarcia and is grant recipient through BIDMC and consultant to Novo Nordisk and Ansh Labs LLC.

References

[1] Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 2016;64(1):73–84 [Epub 2015/12/29].
[2] Polyzos SA, Kountouras J, Mantzoros CS. Obesity and nonalcoholic fatty liver disease: from pathophysiology to therapeutics. Metabolism: clinical and experimental 2018;92:82–97 [Epub 2018/12/07].
[3] Argo CK, Northup PG, Al-Osaimi AM, Caldwell SH. Systematic review of risk factors for fibrosis progression in non-alcoholic steatohepatitis. J Hepatol 2009;51(2):371–9 [Epub 2009/06/09].
[4] 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(1):328–57. Epub 2017/07/18.
[5] Sumida Y, Nakajima A, Itoh Y. Limitations of liver biopsy and non-invasive diagnostic tests for the diagnosis of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. World J Gastroenterol 2014;20(2):475–85 [Epub 2014/02/28].
[6] Castera L, Friedrich-Rust M, Loomba R. Noninvasive assessment of liver disease in patients with nonalcoholic fatty liver disease. Gastroenterology 2019;156(5):1264–81 e4. [Epub 2019/01/21].
[7] Perakakis N, Yazdani A, Karniadakis GE, Mantzoros C. Omics, big data and machine learning as tools to propel understanding of biological mechanisms and to discover novel diagnostics and therapeutics. Metabolism: clinical and experimental 2018;87:A1–9 [Epub 2018/08/12].

[8] Musso G, Cassader M, Paschetta E, Gambino R. Bioactive lipid species and metabolic pathways in progression and resolution of nonalcoholic Steatohepatitis. Gastroenterology 2018;155(2):282–302 e8. [Epub 2018/06/16].

[9] Puri P, Baillie RA, Wiest MM, Mirshahi F, Choudhury J, Cheung O, et al. A lipidomic analysis of nonalcoholic fatty liver disease. Hepatology. 2007;46(4):1081–90. Epub 2007/07/27.

[10] Puri P, Wiest MM, Cheung O, Mirshahi F, Sargeant C, Min HK, et al. The plasma lipidomic signature of nonalcoholic steatohepatitis. Hepatology. 2009;50(6):1827–38. Epub 2009/11/26.

[11] Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: clinical prediction rules and blood-based biomarkers. J Hepatol 2018;68(2):305–15 [Epub 2017/11/21].

[12] Blomme B, Van Steenkiste C, Callewaert N, Van Vlierberghe H. Alteration of protein glycosylation in liver diseases. J Hepatol 2009;50(3):592–603 [Epub 2009/01/23].

[13] Callewaert N, Van Vlierberghe H, Van Hecke A, Laroy W, Delanghe J, Contreras R. Noninvasive diagnosis of liver cirrhosis using DNA sequencer-based total serum protein glycomics. Nat Med 2004;10(4):429–34 [Epub 2004/05/22].

[14] Vanderschaeghe D, Laroy W, Sablon E, Halfon P, Van Hecke A, Delanghe J, et al. GlycoFibroTest is a highly performant liver fibrosis biomarker derived from DNA sequencer-based serum protein glycomics. Molecular & cellular proteomics : MCP. 2009;8(5):986–94. Epub 2009/02/03.

[15] Liu XE, Desmyter L, Gao CF, Laroy W, Dewaele S, Vanhooren V, et al. N-glycomic changes in hepatocellular carcinoma patients with liver cirrhosis induced by hepatitis B virus. Hepatology. 2007;46(5):1426–35. Epub 2007/08/09.

[16] Liu X, Perakakis N, Gong H, Chamberland JP, Brinkoetter MT, Hamnvik OR, et al. Adiponectin administration prevents weight gain and glycemic profile changes in diet-induced obese immune deficient Rag1−/− mice lacking mature lymphocytes. Metabolism: clinical and experimental. 2016;65(12):1720–30. Epub 2016/11/12.

[17] Perakakis N, Kokkinos A, Peradze N, Tentolouris N, Ghaly W, Tsilingiris D, et al. Follistatins in glucose regulation in healthy and obese individuals. Diabetes, obesity & metabolism. 2019;21(3):683–90. Epub 2018/11/06.

[18] Boutari C, Perakakis N, Mantzoros CS. Association of Adipokines with development and progression of nonalcoholic fatty liver disease. Endocrinol Metab (Seoul) 2018;33(1):33–43 [Epub 2018/03/29].

[19] Perakakis N, Upadhyay J, Ghaly W, Chen J, Chrysafi P, Anastasilakis AD, et al. Regulation of the activins-follistatins-inhibins axis by energy status: Impact on reproductive function. Metabolism: clinical and experimental. 2018;85:240–9. Epub 2018/05/13.

[20] Polyzos SA, Kountouras J, Anastasilakis AD, Triantafyllou G, Mantzoros CS. Activin A and follistatin in patients with nonalcoholic fatty liver disease. Metabolism: clinical and experimental 2016;65(10):1550–8 [Epub 2016/09/14].

[21] Polyzos SA, Perakakis N, Mantzoros CS. Fatty liver in lipodystrophy: a review with a focus on therapeutic perspectives of adiponectin and/or leptin replacement. Metabolism: clinical and experimental 2019;96:66–82 [Epub 2019/05/10].

[22] Polyzos SA, Toulis KA, Goulis DG, Zavos C, Kountouras J. Serum total adiponectin in nonalcoholic fatty liver disease: a systematic review and meta-analysis. Metabolism: clinical and experimental 2011;60(3):313–26 [Epub 2010/11/03].

[23] Polyzos SA, Kountouras J, Anastasilakis AD, Geladari EV, Mantzoros CS. Irisin in patients with nonalcoholic fatty liver disease. Metabolism: clinical and experimental 2014;63(2):207–17 [Epub 2013/10/22].

[24] Bondia-Pons I, Castellote AI, Lopez-Sabater MC. Comparison of conventional and fast gas chromatography in human plasma fatty acid determination. J Chromatogr B Analyt Technol Biomed Life Sci 2004;809(2):339–44 [Epub 2004/08/19].

[25] Kalivodová A, Hron K, Filzmoser P, Najdekr L, Janečková H, Adam T. PLS-DA for compositional data with application to metabolomics. J Chemometr 2015;1(29):21–8.

[26] van den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, van der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics 2006;1(7):142.

[27] Lê Cao KA, Boitard S, Besse P. Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. BMC bioinformatics 2011;1(12):253.

[28] Pedregosa FaV, G. and Gramfort, A. and Michel, V. and Thirion, B. and Grisel, O. and Blondel, M. and Prettenhofer, P. and Weiss, R. and Dubourg, V. and Vanderplas, J. and Passos, A. and Cournapeau, D. and Brucher, M. and Perrot, M. and Duche. Scikit-learn: machine learning in python. Journal of Machine Learning Research 2011;12:2825–30.

[29] Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. Machine learning 2002;1-3(46):389–422.

[30] Bishop CM. Pattern Recognition and Machine Learning. Springer; 2006.

[31] Chong J, Xia J. MetaboAnalystR: an R package for flexible and reproducible analysis of metabolomics data. Bioinformatics 2018;34(24):4313–4 [Epub 2018/06/30].

[32] Perakakis N. Data sets - NAFLD. figshare. Dataset 2019;https://doi.org/10.6084/m9.figshare.9774809.

[33] Perakakis N. R command history in MetaboanalystR - NAFLD. figshare. Dataset 2019; https://doi.org/10.6084/m9.figshare.9773843.

[34] Yazdani A. Python code - NAFLD. 2019;https://github.com/alirezayazdani1/NAFLD.

[35] Caussy C, Ajmera VH, Puri P, Hsu CL, Bassirian S, Mgdsyan M, et al. Serum metabolites detect the presence of advanced fibrosis in derivation and validation cohorts of patients with non-alcoholic fatty liver disease. Gut. 2018. [Epub 2018/12/21].

[36] Kwok R, Tse YK, Wong GL, Ha Y, Lee AU, Ngu MC, et al. Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease–the role of transient elastography and plasma cytokeratin-18 fragments. Alimentary pharmacology & therapeutics. 2014;39(3):254–69. Epub 2013/12/07.

[37] Musso G, Gambino R, Cassader M, Pagano G. Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. Ann Med 2011;43(8):617–49 [Epub 2010/11/03].

[38] Cusi K, Chang Z, Harrison S, Lomonaco R, Bril F, Orsak B, et al. Limited value of plasma cytokeratin-18 as a biomarker for NASH and fibrosis in patients with non-alcoholic fatty liver disease. Journal of hepatology. 2014;60(1):167–74. Epub 2013/08/27.

[39] Buzzetti E, Pinzani M, Tsochatzis EA. The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NAFLD). Metabolism: clinical and experimental 2016;65(8):1038–48 [Epub 2016/01/30].

[40] Reccia I, Kumar J, Akladios C, Virdis F, Pai M, Habib N, et al. Non-alcoholic fatty liver disease: a sign of systemic disease. Metabolism: clinical and experimental. 2017;72:94–108. Epub 2017/06/24.

[41] Clarke JD, Novak P, Lake AD, Hardwick RN, Cherrington NJ. Impaired N-linked glycosylation of uptake and efflux transporters in human non-alcoholic fatty liver disease. Liver international : official journal of the International Association for the Study of the Liver 2017;37(7):1074–81 [Epub 2017/01/18].

[42] Poynard T, Ratziu V, Charlotte F, Messous D, Munteanu M, Imbert-Bismut F, et al. Diagnostic value of biochemical markers (NashTest) for the prediction of non alcoholo steato hepatitis in patients with non-alcoholic fatty liver disease. BMC gastroenterology. 2006;6:34. Epub 2006/11/14.

[43] Younossi ZM, Jarrar M, Nugent C, Randhawa M, Afendy M, Stepanova M, et al. A novel diagnostic biomarker panel for obesity-related nonalcoholic steatohepatitis (NASH). Obesity surgery. 2008;18(11):1430–7. Epub 2008/05/27.

[44] Crossan C, Tsochatzis EA, Longworth L, Gurusamy K, Davidson B, Rodriguez-Peralvarez M, et al. Cost-effectiveness of non-invasive methods for assessment and monitoring of liver fibrosis and cirrhosis in patients with chronic liver disease: systematic review and economic evaluation. Health Technol Assess. 2015;19(9):1–409, v-vi. Epub 2015/01/31.

[45] Mato JM, Alonso C, Noureddin M, Lu SC. Biomarkers and subtypes of deranged lipid metabolism in non-alcoholic fatty liver disease. World J Gastroenterol 2019;25(24):3009–20 [Epub 2019/07/12].

[46] Mayo R, Crespo J, Martinez-Arranz I, Banales JM, Arias M, Minchole I, et al. Metabolomic-based noninvasive serum test to diagnose nonalcoholic steatohepatitis: results from discovery and validation cohorts. Hepatology communications. 2018;2(7):807–20. Epub 2018/07/22.

[47] Pabst B, Futatsugi K, Li Q, Ahn K. Mechanistic characterization of long residence time inhibitors of Diacylglycerol Acyltransferase 2 (DGAT2). Biochemistry 2018;57(51):6997–7010 [Epub 2018/11/14].

[48] Vitali C, Khetarpal SA, Rader DJ. HDL cholesterol metabolism and the risk of CHD: new insights from human genetics. Curr Cardiol Rep 2017;19(12):132 [Epub 2017/11/06].

[49] Machado MV, Coutinho J, Carepa F, Costa A, Proenca H, Cortez-Pinto H. How adiponectin, leptin, and ghrelin orchestrate together and correlate with the severity of nonalcoholic fatty liver disease. Eur J Gastroenterol Hepatol 2012;24(10):1166–72 [Epub 2012/06/28].

[50] Blomme B, Francque S, Trepo E, Libbrecht L, Vanderschaeghe D, Verrijken A, et al. N-glycan based biomarker distinguishing non-alcoholic steatohepatitis from steatosis independently of fibrosis. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2012;44(4):315–22. Epub 2011/11/29.

[51] Kobayashi T, Ogawa K, Furukawa JI, Hanamatsu H, Hato M, Yoshinaga T, et al. Quantifying Protein-Specific N-Glycome Profiles by Focused Protein and Immunoprecipitation Glycomics. Journal of proteome research. 2019;18(8):3133–41. Epub 2019/07/04.

[52] Yamasaki Y, Nouso K, Miyahara K, Wada N, Dohi C, Morimoto Y, et al. Use of non-invasive serum glycan markers to distinguish non-alcoholic steatohepatitis from simple steatosis. Journal of gastroenterology and hepatology. 2015;30(3):528–34. Epub 2014/08/30.

[53] Sriharan M, Reichelt AJ, Opperman ML, Duncan BB, Mengue SS, Crook MA, et al. Total sialic acid and associated elements of the metabolic syndrome in women with and without previous gestational diabetes. Diabetes care. 2002;25(8):1331–5. Epub 2002/07/30.

[54] He J, Mao W, Zhang J, Jin X. Association between serum Sialic acid levels and nonalcoholic fatty liver disease: a cross-sectional study. Ann Nutr Metab 2015;67(2):69–75 [Epub 2015/09/01].

[55] Lu Z, Ma H, Xu C, Shao Z, Cen C, Li Y. Serum Sialic acid level is significantly associated with nonalcoholic fatty liver disease in a nonobese Chinese population: a cross-sectional study. Biomed Res Int 2016;2016:5921589 [Epub 2016/04/05].

[56] Ma J, Sanda M, Wei R, Zhang L, Goldman R. Quantitative analysis of core fucosylation of serum proteins in liver diseases by LC-MS-MRM. J Proteomics 2018;189:67–74 [Epub 2018/02/11].

[57] Shimizu K, Katoh H, Yamashita F, Tanaka M, Tanikawa K, Taketa K, et al. Comparison of carbohydrate structures of serum alpha-fetoprotein by sequential glycosidase digestion and lectin affinity electrophoresis. Clinica chimica acta; international journal of clinical chemistry. 1996;254(1):23–40. Epub 1996/10/15.

[58] Tawara S, Tatsumi T, Iio S, Kobayashi I, Shigekawa M, Hikita H, et al. Evaluation of Fucosylated Haptoglobin and Mac-2 Binding Protein as Serum Biomarkers to Estimate Liver Fibrosis in Patients with Chronic Hepatitis C. PloS one. 2016;11(3):e0151828. Epub 2016/03/24.

[59] Adinolfi LE, Rinaldi L, Guerrera B, Restivo L, Marrone A, Giordano M, et al. NAFLD and NASH in HCV Infection: Prevalence and Significance in Hepatic and Extrahepatic Manifestations. International journal of molecular sciences. 2016;17(6). Epub 2016/05/28.