Received: 1 February 2023 | Accepted: 8 February 2023

DOI: 10.1097/HC9.0000000000000112

EDITORIAL

OPEN

Clustering NAFLD: phenotypes of nonalcoholic fatty liver disease and their differing trajectories

Globally, in the general adult population, the prevalence of NAFLD is estimated at 25%.[1] The prevalence of NAFLD is higher (~40%–60%) in overweight and obese subjects, particularly in the presence of impaired metabolic health,[2] and the highest global prevalence of NAFLD (~55%–70%) is found in patients with diabetes.[3] NAFLD is the main cause of chronic liver disease and HCC.[4] Furthermore, the strong epidemiological relationships of NAFLD with type 2 diabetes and cardiovascular diseases indicate very close pathophysiological relationships between NAFLD and obesity-associated cardiometabolic diseases.[5]

However, there is large heterogeneity among patients with NAFLD with respect to their risk of cardiometabolic diseases.[3] This may result from the fact that different major pathways are involved in the pathogenesis of NAFLD. Among them is NAFLD associated with a stronger hepatic genetic component. For example, genetic variants in *PNPLA3* (148Met allele) and *TM6SF2* (167Lys allele) strongly associate with steatosis and progression to NASH, cirrhosis, and HCC, but also with the absence of insulin resistance, low blood triglycerides, low LDL cholesterol concentration, and protection from coronary artery disease. Furthermore, NAFLD predominantly driven by *de novo* lipogenesis and NAFLD predominantly driven by adipose tissue dysfunction exist, which are both associated with insulin resistance but are also thought to differ in their risk of cardiometabolic diseases.[3]

Yi et al[6] set out to identify clinically important groups among patients with NAFLD and assess the long-term outcomes between different subphenotypes and, most recently, published their findings in *Hepatology Communications*. For this, they analyzed the data from the US Third National Health and Nutrition Examination Survey, where fatty liver was diagnosed in individuals by abdominal ultrasound and used their linked mortality data through December 2019. As a data dimensionality reduction approach, the authors performed a 2-stage cluster analysis (a hierarchical cluster analysis using the Ward method to determine the optimum number of clusters, followed by an allocation of each patient into a particular cluster). Using 21 baseline variables, body mass index (BMI), waist circumference, hemoglobin, glycohemoglobin, waist-to-hip ratio, uric acid, HDL cholesterol, and homeostasis model assessment of insulin resistance were identified as the most important variables in the prediction of the patient clusters. Three distinct clusters were identified. Cluster 1 predominantly comprised younger (mean age 40 y), lean (mean BMI 24 kg/m^2) females (76%) with a low cardiometabolic risk profile and a lower prevalence of comorbidities. Cluster 2 consisted mostly of older (mean age 50 y), obese (mean BMI 34 kg/m^2) females (75%) with a high prevalence of insulin resistance (83%) and diabetes (34%). Cluster 3 was predominantly composed of older (mean age 49 y), overweight/obese (mean BMI 30 kg/m^2) males (75%) with insulin resistance (72%), a moderately elevated prevalence of diabetes (15%), hypertension, and atherogenic dyslipidemia. During a median follow-up period of 312 months compared with patients in cluster 1, patients in cluster 2 and cluster 3 had higher all-cause and cardiovascular mortality, also after the adjustment for age, sex, BMI, and race/ethnicity. No differences in all-cause mortality were observed between patients in clusters 2 and 3.[6]

The authors concluded that patients with NAFLD, who were allocated to cluster 1 and who did not have insulin resistance, hypertension, or dyslipidemia, may have a pathophysiology of NAFLD, which is stronger related to the hepatic genetic component. Unfortunately, the authors could not study incident fibrosis or cirrhosis, as it would be expected that no differences in the incidence of these advanced stages of liver disease would exist between the clusters. The authors further hypothesized that in patients in cluster 2, predominantly being obese and having severe insulin resistance and a high prevalence of diabetes, hepatic steatosis may

Abbreviations: BMI, body mass index; MAFLD, metabolic-associated fatty liver disease.

mainly be driven by adipose dysfunction. The measurements of free fatty acids, which would be expected to be elevated in patients in this cluster and which would have allowed to estimate adipose tissue insulin resistance,[7] would have been important to test this hypothesis. Finally, the authors hypothesized that NAFLD in patients in cluster 3, mostly having atherogenic dyslipidemia, hypertension, and liver and kidney damage, may be predominantly driven by hepatic *de novo* lipogenesis. The determination of hepatic *de novo* lipogenesis, for example, using serum fatty acid ratios in very low-density lipoprotein triglycerides,[8] could have helped to investigate this hypothesis. Altogether, Yi et al[6] provided interesting novel data that the clustering of patients with NAFLD may help to identify the subphenotypes of patients with different main pathomechanisms of NAFLD. Because the NAFLD clusters differed in important cardiometabolic risk factors, such as body fat distribution and insulin resistance, regarding risk stratification or prediction of major outcomes, the proposed cluster approach seems not to be superior to the established cardiometabolic risk prediction models.

Recently, another study has identified 5 clusters of metabolic-associated fatty liver disease (MAFLD) in a Chinese cohort and validated the results in the UK Biobank database.[9] That study only used a relatively small number of variables (age, glycohemoglobin, total cholesterol/HDL cholesterol ratio total cholesterol, triglycerides, and lipoprotein (a) levels). Patients in different clusters exhibited different risks of type 2 diabetes, coronary heart disease, and all-causes mortality. Patients in cluster 3, which was referred to as "severe insulin resistance–related MAFLD," had significantly worst survival outcomes and higher cardiometabolic risks than those in other clusters. The other clusters were referred to as "mild obesity and dyslipidemia-related MAFLD" (cluster 1), "age-related MAFLD" (cluster 2), "high lipoprotein (a)-related MAFLD" (cluster 4), and "severe mixed hyperlipidemia-related MAFLD" (cluster 5). Altogether, the results of this study by Ye et al.[9] and the study by Yi et al.[6] highlight that the allocation of the patients with fatty liver to specific clusters strongly depends on the parameters that are used to generate the clusters and that these parameters, and most probably not the liver phenotype, may drive the pathogenesis of the outcomes (Figure 1).

What may be the other approaches to identify major subphenotypes of NAFLD and their mechanisms promoting cardiometabolic diseases? For this purpose, it may be important to focus on hepatokines and adipokines. In this respect, we have most recently proposed how using the clinically important hepatokine fetuin-A and the adipokine adiponectin, together with precisely measured liver fat

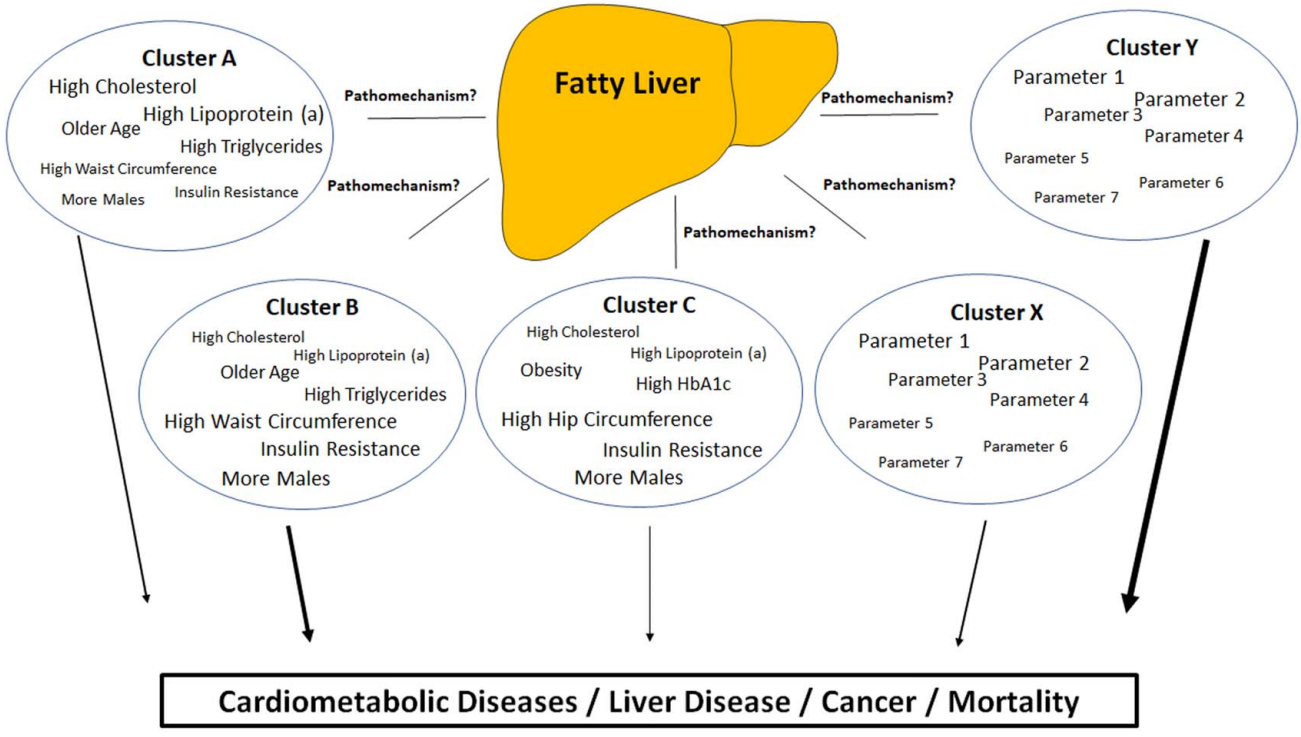

FIGURE 1 Hypothetical depiction of risk clusters in patients with fatty liver and their relationships with the incidence of diseases and mortality. Several risk clusters can be identified in patients with NAFLD. For the hypothetical clusters, A–C parameters that are often used to generate the clusters are depicted. The font size used for these parameters indicates that the parameters have a large, moderate, or small impact on the assignment of subjects to the individual clusters. The hypothetical clusters X and Y represent additional clusters that can be identified in case additional parameters (1-7) are being used to generate clusters. In most cases, the pathomechanisms linking fatty liver and these risk clusters are unknown. Furthermore, it is unclear whether the risk clusters mediate the fatty liver–associated incidence of diseases and mortality.

content and visceral fat mass, can identify the clusters of patients with NAFLD in whom dysfunctional hepatic or dysfunctional adipose tissue are the major pathomechanisms of cardiometabolic diseases.[10]

In summary, data dimensionality reduction approaches, such as cluster, principal component, and factor analyses, became very popular in clinical research. They are also being used to identify subphenotypes of NAFLD and study their impact on diseases and mortality. This field of research is very important, considering the large heterogeneity of the pathogenesis of NAFLD.[3] However, for such analytical approaches, a careful selection of the most important parameters related to the different pathomechanisms of NAFLD is necessary to advance this field of clinical research.

CONFLICTS OF INTEREST

The authors have no conflicts to report.

Konstantinos Kantartzis[1,2,3]
Norbert Stefan[1,2,3]

[1]*Department of Internal Medicine IV, University Hospital Tübingen, Tübingen, Germany*
[2]*Institute of Diabetes Research and Metabolic Diseases (IDM) of the Helmholtz Centre Munich, Tübingen, Germany*
[3]*German Center for Diabetes Research (DZD), Neuherberg, Germany*

Correspondence
Norbert Stefan, Department of Internal Medicine IV, University Hospital Tübingen, Otfried-Müller-Str. 10, 72076 Tübingen, Germany.
Email: Norbert.stefan@med.uni-tuebingen.de

ORCID

Konstantinos Kantartzis https://orcid.org/0000–0002–0584–1138
Norbert Stefan https://orcid.org/0000–0002–2186–9595

REFERENCES

1. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-metanalytic assessment of prevalence, incidence, and outcomes. Hepatology. 2016;64:73–84.
2. Stefan N, Schick F, Häring HU. Causes, characteristics, and consequences of metabolically unhealthy normal weight in humans. Cell Metab. 2017;26:292–300.
3. Stefan N, Cusi K. A global view of the interplay between non-alcoholic fatty liver disease and diabetes. Lancet Diabetes Endocrinol. 2022;10:284–96.
4. Powell EE, Wong VW, Rinella M. Non-alcoholic fatty liver disease. Lancet. 2021;397:2212–4.
5. Targher G, Tilg H, Byrne CD. Non-alcoholic fatty liver disease: a multisystem disease requiring a multidisciplinary and holistic approach. Lancet Gastroenterol Hepatol. 2021;6:578–88.
6. Yi J, Wang L, Guo J, Ren X. Novel metabolic phenotypes for extrahepatic complication of nonalcoholic fatty liver disease. Hepatol Commun. 2023;7:e0016.
7. Gastaldelli A, Cusi K. From NASH to diabetes and from diabetes to NASH: mechanisms and treatment options. JHEP Rep. 2019;1:312–28.
8. Silbernagel G, Kovarova M, Cegan A, Machann J, Schick F, Lehmann R, et al. High hepatic SCD1 activity is associated with low liver fat content in healthy subjects under a lipogenic diet. J Clin Endocrinol Metab. 2012;97:E2288–92.
9. Ye J, Zhuang X, Li X, Gong X, Sun Y, Wang W, et al. Novel metabolic classification for extrahepatic complication of metabolic associated fatty liver disease: a data-driven cluster analysis with international validation. Metabolism. 2022;136:155294.
10. Stefan N, Schick F, Birkenfeld AL, Häring HU, White MF. The role of hepatokines in NAFLD. Cell Metab. 2023;35:236–52.