Journal of Gastroenterology and Hepatology

JGHF

doi:10.1111/jgh.16326

ORIGINAL ARTICLE - HEPATOLOGY (CLINICAL)

Identification of clinical phenotypes associated with poor prognosis in patients with nonalcoholic fatty liver disease via unsupervised machine learning

Takanori Ito,[*1] Hikaru Morooka,[†1] Hirokazu Takahashi,[‡] Hideki Fujii,[§] Michihiro Iwaki,[¶] Hideki Hayashi,[**] Hidenori Toyoda,[††] Satoshi Oeda,[‡,‡‡] Hideyuki Hyogo,[§§,¶¶] Miwa Kawanaka,[***] Asahiro Morishita,[†††] Kensuke Munekage,[‡‡‡] Kazuhito Kawata,[§§§] Tsubasa Tsutsumi,[¶¶¶] Koji Sawada,[****] Tatsuji Maeshiro,[††††] Hiroshi Tobita,[‡‡‡‡] Yuichi Yoshida,[§§§§] Masafumi Naito,[§§§§] Asuka Araki,[‡‡‡‡] Shingo Arakaki,[††††] Takumi Kawaguchi,[¶¶¶] Hidenao Noritake,[§§§] Masafumi Ono,[¶¶¶¶] Tsutomu Masaki,[†††] Satoshi Yasuda,[††] Eiichi Tomita,[**] Masato Yoneda,[¶] Akihiro Tokushige,[*****] Masatoshi Ishigami,[*] Yoshihiro Kamada,[†††††] Shinichiro Ueda,[‡‡‡‡‡] Shinichi Aishima,[§§§§§] Yoshio Sumida,[¶¶¶¶¶] Atsushi Nakajima,[¶] Takeshi Okanoue[******] and for the Japan Study Group of Nonalcoholic Fatty Liver Disease (JSG-NAFLD)

Departments of *Gastroenterology and Hepatology, †Emergency and Critical Care Medicine, Nagoya University Graduate School of Medicine, Nagoya, ‡Liver Center, ‡‡Department of Laboratory Medicine, Saga University Hospital, §§§§§Department of Pathology and Microbiology, Faculty of Medicine, Saga University, Saga, §Department of Hepatology, Graduate School of Medicine, Osaka Metropolitan University, §§§§Department of Gastroenterology and Hepatology, Suita Municipal Hospital, †††††Department of Advanced Metabolic Hepatology, Osaka University Graduate School of Medicine, Osaka, ¶Division of Gastroenterology and Hepatology, Yokohama City University Graduate School of Medicine, Yokohama, **Department of Gastroenterology and Hepatology, Gifu Municipal Hospital, Gifu, ††Department of Gastroenterology, Ogaki Municipal Hospital, Ogaki, §§Department of Gastroenterology, JA Hiroshima Kouseiren General Hospital, ¶¶Hyogo Life Care Clinic Hiroshima, Hiroshima, ***Department of General Internal Medicine 2, Kawasaki Medical Center, Kawasaki Medical School, Okayama, †††Department of Gastroenterology and Neurology, Faculty of Medicine, ¶¶¶¶Division of Innovative Medicine for Hepatobiliary and Pancreatology, Faculty of Medicine, Kagawa University, Takamatsu, ‡‡‡Department of Gastroenterology and Hepatology, Kochi Medical School, Kochi, §§§Hepatology Division, Department of Internal Medicine II, Hamamatsu University School of Medicine, Shizuoka, ¶¶¶Division of Gastroenterology, Department of Medicine, Kurume University School of Medicine, Kurume, ****Division of Metabolism and Biosystemic Science, Gastroenterology, and Hematology/Oncology, Asahikawa Medical University, Asahikawa, ††††First Department of Internal Medicine, University of the Ryukyus Hospital, ‡‡‡‡‡Department of Clinical Pharmacology and Therapeutics, Graduate School of Medicine, University of the Ryukyus, Okinawa, ‡‡‡‡Division of Hepatology, Shimane University Hospital, Izumo, *****Department of Cardiovascular Medicine and Hypertension, Kagoshima University Graduate School of Medical and Dental Sciences, Kagoshima, ¶¶¶¶¶Division of Hepatology and Pancreatology, Department of Internal Medicine, Aichi Medical University, Nagakute, ******Hepatology Center, Saiseikai Suita Hospital, Suita, Japan

Key words
BMI, FIB-4 index, NAFLD, NIT, Unsupervised machine learning.

Accepted for publication 1 August 2023.

Correspondence
Hirokazu Takahashi, Liver Center, Saga University Hospital, Saga 8498501, Japan.
Email: takahas2@cc.saga-u.ac.jp

Guarantor of the article: Hirokazu Takahashi
Declaration of conflict of interest: Takanori Ito received speaker's fees from Chugai Pharmaceutical Co., Ltd. Masatoshi Ishigami received grants from Chugai Pharmaceutical Co., Ltd. Hidenori Toyoda received speaker's fees from AbbVie, Bayer, Gilead, Terumo, and Eisai. Masafumi Ono received grants from HITO Medical Center. Takumi Kawaguchi received speaker's fees from Janssen Pharmaceutical K.K., Mitsubishi Tanabe Pharma Corporation, Taisho Pharmaceutical Co., Ltd., Otsuka

Abstract

Background and Aims: Both fibrosis status and body weight are important for assessing prognosis in nonalcoholic fatty liver disease (NAFLD). The aim of this study was to identify population clusters for specific clinical outcomes based on fibrosis-4 (FIB-4) index and body mass index (BMI) using an unsupervised machine learning method.

Methods: We conducted a multicenter study of 1335 biopsy-proven NAFLD patients from Japan. Using the Gaussian mixture model to divide the cohort into clusters based on FIB-4 index and BMI, we investigated prognosis for these clusters.

Results: The cohort consisted of 223 cases (16.0%) with advanced fibrosis (F3–4) as assessed from liver biopsy. Median values of BMI and FIB-4 index were 27.3 kg/m^2 and 1.67. The patients were divided into four clusters by Bayesian information criterion, and all-cause mortality was highest in cluster d, followed by cluster b ($P = 0.001$). Regarding the characteristics of each cluster, clusters d and b presented a high FIB-4 index (median 5.23 and 2.23), cluster a presented the lowest FIB-4 index (median 0.78), and cluster c was associated with moderate FIB-4 level (median 1.30) and highest BMI (median 34.3 kg/m^2). Clusters a and c had lower mortality rates than clusters b and d. However, all-cause of death in clusters a and c was unrelated to liver disease.

Conclusions: Our clustering approach found that the FIB-4 index is an important predictor of mortality in NAFLD patients regardless of BMI. Additionally, non-liver-related diseases were identified as the causes of death in NAFLD patients with low FIB-4 index.

© 2023 Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

Pharmaceutical Co., Ltd., and EA Pharma Co., Ltd. All other authors have no conflict to disclose.
Author contribution: Study design: Takanori Ito, Hikaru Morooka, and Hirokazu Takahashi. Data collection: All authors. Data analysis: Takanori Ito and Hikaru Morooka. Drafting of the article: Takanori Ito, Hikaru Morooka, and Hirokazu Takahashi. Data interpretation, review and/or revision of the manuscript: All authors. Study concept and study supervision: Takanori Ito and Hirokazu Takahashi.
Financial support: This research was financially supported by Japan Strategic Medical Administration Research Center (J-SMARC).
[1]Takanori Ito and Hikaru Morooka contributed equally.

Introduction

Nonalcoholic fatty liver disease (NAFLD) is the leading cause of chronic liver disease, involving 25% of the general population.[1] The prevalence of NAFLD is projected to increase even further in the coming years, mainly because of the rising prevalence of obesity and metabolic syndrome.[2] Multiple factors contribute to the progression of the disease onto its active necro-inflammatory form, termed nonalcoholic steatohepatitis. NAFLD can eventually progress to advanced fibrosis and cirrhosis, and even hepatocellular carcinoma (HCC) in a subset of patients.[3] Liver fibrosis has been identified as the most important factor associated with overall and liver-related mortality in NAFLD.[4–7]

Liver biopsy currently remains the gold standard for evaluating fibrosis stage in NAFLD. However, this procedure is associated with sampling errors in pathological assessment, high cost, and (although rare) risks associated with the procedure, such as hemorrhage.[8,9] To predict liver fibrosis in NAFLD, various noninvasive tests (NITs) have been developed, such as serum markers and imaging for measuring liver stiffness.[10,11] Of these, the fibrosis-4 (FIB-4) index and NAFLD fibrosis score (NFS) are mostly used to exclude and predict advanced fibrosis in NAFLD. These two NITs have been recommended for clinical application by the American Association for the Study of Liver Disease (AASLD) and by the European Association for the Study of the Liver (EASL) to rule out or rule in high-risk NAFLD patients with advanced fibrosis.[3,12] Very recently, a multicenter study showed that the FIB-4 index carries significantly higher predictive power for advanced fibrosis (F3–4) compared with NFS and Hepamet fibrosis score in 1489 biopsy-proven NAFLD patients from Asia.[13] Additionally, Boursier et al. reported that the FIB-4, as well as transient elastography, can stratify the risk of liver-related complications in 1057 patients with NAFLD.[14]

NAFLD is generally associated with obesity and insulin resistance[15]; however, it has also been reported in lean individuals as "lean NAFLD."[16–18] Previous study showed that lean NAFLD, despite the absence of obesity, presented similar cardiovascular disease (CVD) and cancer-related mortality to that observed in obese NAFLD individuals and increased all-cause mortality risk.[19] Thus, when assessing prognosis of NAFLD, it is necessary to take into account not only the stage of fibrosis but also body weight.

Recently, machine learning methods have received increasing attention because of technical advances. Unsupervised machine learning methods, such as Gaussian mixture models (GMMs), can be used to identify potential clusters or phenotypes, allowing objective clustering analysis without the bias introduced by analysts.[20] In this study, we used GMM to reveal potential clusters of the FIB-4 and body mass index (BMI) within a large multicenter cohort dataset of biopsy-proven NAFLD and investigated the outcomes associated with the different clusters.

Methods

Patient selection. In the current study, we used a database from the Japan Study Group of NAFLD (JSG-NAFLD), a multicenter longitudinal cohort study (CLIONE, Clinical Outcome Nonalcoholic Fatty Liver Disease). Data were collected from 15 hepatology centers in Japan and compiled using electronic data capture tools (REDCap). The resulting database collects clinical and pathological findings from 1398 NAFLD patients who were diagnosed by liver biopsy between December 1994 and December 2020. In our study, 63 patients were excluded because of the following reasons: (i) no data for either FIB-4 index or BMI, $n = 4$; (ii) follow-up duration of less than 6 months, $n = 59$. We analyzed data from the remaining 1335 biopsy-proven NAFLD patients. The study protocol was carried out in accordance with the 1975 Declaration of Helsinki. It was approved by the Institutional Review Board at Saga University Hospital (approval no. 2020-04-R-02) and at each participating study center.

NAFLD diagnosis and assessment of pathology. We diagnosed NAFLD based on the AASLD guidelines.[3] Briefly, we confirmed the following: (i) the presence of hepatic steatosis observed via imaging; (ii) the absence of excessive alcohol use (ethanol intake of <30 g per day for men and <20 g per day for women); (iii) no other cause of fatty liver such as drugs; and (iv) no other causes of chronic liver disease, such as viral infection (hepatitis B or C virus), primary biliary cholangitis, or autoimmune hepatitis.

Percutaneous liver biopsy was performed with a 16–17-G needle under ultrasonographic guidance. The liver biopsy specimens were immediately fixed in 10% formalin and embedded in paraffin. An experienced pathologist (S.A.), specialized in NAFLD, evaluated the specimens at Saga University. Pathological evaluations were carried out without knowledge of clinical data. NAFLD was defined as the presence of ≥5% hepatic steatosis based on the criteria set out by Kleiner et al.[21] The severity of hepatic inflammatory grade and fibrosis stage were scored on the basis of the scale proposed by Brunt et al.[22] (Supporting Information).

Follow-up and outcomes. The start of follow-up was defined as the date on which liver biopsy was carried out. Monitoring of clinical events was performed every 3–12 months until the last visit, death, or liver transplantation using imaging and/or blood tests. Specialists at each hospital reviewed causes of death from medical records. If a patient dies after being transferred to another hospital, the outcome was informed to the original hospital as much as possible. The diagnosis of HCC was based on imaging characteristics specified by the guidelines of the AASLD.[23] Death

© 2023 Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

14401746, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jgh.16326 by Int. Uni Of Health & Welfare, Wiley Online Library on [22/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

due to liver failure other than HCC was classified as jaundice, hepatic encephalopathy, esophageal or gastric varices, and ascites accompanied by liver cirrhosis. The definition of CVD in this study involved myocardial infarction, angina pectoris, stroke, abdominal aortic aneurysm, and congestive heart failure.

Clustering by the Gaussian mixture model. The aim of our clustering method was to divide patients into subgroups using GMM. We used the FIB-4 index and BMI as clustering inputs for the GMM. First, we used the Bayesian information criterion (BIC) to assess the best clustering number. With the most appropriate model and cluster number, we calculated the GMM (model selection: VVV, which is ellipsoidal, varying volume, shape, and orientation).

Statistical analysis. Continuous variables were expressed as medians (interquartile range). Categorical variables were expressed as numbers (percentages). We used the chi-square test to compare categorical variables and the Kruskal–Wallis test to compare continuous variables between groups. An actuarial analysis of overall survival (OS) was performed using the Kaplan–Meier method. Differences across groups were compared using a log-rank test. The incidence of liver-related versus non-liver-related death was estimated using the cumulative incidence method and compared using Gray's test, because these events were competing risk factors. Cox proportional hazard models were used to estimate the hazard ratio (HR) and associated 95% confidence interval (CI). The FIB-4 index and NFS were calculated based on the previous reports[24,25] (Supporting Information). We relied on previous reports to set FIB-4 and BMI cutoff values for Kaplan–Meier curves applied to our Asian sample.[26,27] The R package "mclust" was used for GMM analysis.[20] For all tests, statistical significance was set at $P < 0.05$. We used R (version 4.1.2, R foundation for Statistical Computing, Vienna, Austria; http://www.R-project.org/) for all statistical analyses.

Results

Baseline characteristics across entire sample of patients. Table 1 lists baseline characteristics across the 1335 patients included in this study. Median age was 57.0 years with a predominance of women (57.2%). Of these cases, 36.5%, 42.1%, and 57.8% had type2 diabetes mellitus, hypertension, and dyslipidemia, respectively. Median values of BMI and FIB-4 index were 27.3 kg/m^2 and 1.67. Regarding the degree of liver fibrosis, 595 cases (44.6%) presented significant fibrosis (F2–4), and 223 cases (16.0%) had advanced fibrosis (F3–4) on liver biopsy. The median follow-up period was 4.9 years (range 0.5–21.6).

Overall survival rate by FIB-4 index and BMI. During the follow-up period, 46 patients (3.4%) died because of liver-related (n = 17, 37.0%) and non-liver-related diseases (n = 29, 63.0%). Detailed information on the causes of death is listed in Supplementary Table. In cases of liver-related death, eight were caused by HCC and nine by liver failure. Deaths caused by non-liver-related diseases involved 15 malignancies (mainly gastrointestinal cancers other than HCC), three infections, four CVD, and seven others. As a first step in the analysis, we examined whether FIB-4 index and BMI could predict the prognosis of NAFLD patients using previously established cutoff values. Figures Figure S1 and S2 show cumulative OS for FIB-4 and BMI, respectively. When patients were categorized based on low and high cutoff points of FIB-4 index as <1.3 (low, n = 550), 1.3–2.67(intermediate, n = 459), and ≥2.67 (high, n = 326), higher FIB-4 index values were associated with significantly lower OS rates ($P < 0.0001$) (Fig. S1). Conversely, when subdividing the cohort by BMI based on World Health Organization (WHO) criteria for the Asian population[27] where BMI < 23 is classified as lean (n = 167), 23–27.5 as overweight (n = 523), and >27.5 as obese (n = 645), we found no connection with OS (P = 0.55) (Fig. S2).

Table 1 Patient characteristics of the overall cohort

Factors		n = 1335
Age	(years)	57 [45, 65]
Sex	Female	764 (57.2)
	Male	571 (42.8)
Type2 diabetes	Presence (%)	487 (36.5)
Hypertension	Presence (%)	561 (42.1)
Dyslipidemia	Presence (%)	772 (57.8)
BMI	(kg/m^2)	27.3 [24.8, 30.6]
AST	(U/L)	52 [36, 75]
ALT	(U/L)	73 [47, 110]
GGT	(U/L)	60 [40, 99]
Platelet counts	(10^9/L)	213 [173, 262]
FIB-4 index		1.67 [0.90, 2.60]
Pathological findings		
Steatosis	0	8 (0.6)
	1	946 (70.9)
	2	252 (18.9)
	3	129 (9.7)
Activity	0	66 (4.9)
	1	841 (63.0)
	2	347 (26.0)
	3	81 (6.1)
Fibrosis	0	225 (16.9)
	1	515 (38.6)
	2	382 (28.6)
	3	189 (14.2)
	4	24 (1.8)
Observation time*	(years)	4.9 [0.5–21.6]
Liver-related death	Presence (%)	17 (1.3)
Non-liver-related death	Presence (%)	29 (2.2)

Values are expressed as median (first to third interquartiles), *median (range), and number (%).
ALT, alanine transaminase; AST, aspartate aminotransferase; BMI, body mass index; FIB-4, fibrosis-4; GGT, gamma glutamyl transferase; NFS, NAFLD fibrosis score.

Clustering by GMM and patient characteristics in each cluster. Next, we performed clustering analysis with GMM, an unsupervised machine learning method, fed with data for FIB-4 index and BMI. First, we determined the appropriate number of clusters via a BIC plot (Fig. S3) to be four. On the basis

© 2023 Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

14401746, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jgh.16326 by Int. Uni Of Health & Welfare, Wiley Online Library on [22/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of this result, we classified our cohort into four clusters: "a" (n = 465), "b" (n = 631), "c" (n = 126), and "d" (n = 113). Cluster distributions are shown in Figure 1.

Table 2 details patient characteristics for each cluster. The FIB-4 index of cluster d was the highest (median 5.23), followed by cluster b (median 2.23). While almost all cases in cluster d (99.1%) had a high FIB-4 index (>2.67), the FIB-4 index for cluster a was <1.3 (Table 2 and Fig. S4). Consistent with the distribution of FIB-4, the rate of advanced fibrosis (F3–4) was highest in group d (54.9%), followed by group b (20.1%). For BMI, obese patients (BMI > 27.5 kg/m^2) were most prevalent in cluster c (95.2%), while cluster b showed the lowest BMI (median 26.1 kg/m^2).

Outcomes by clusters. Figure 2 shows that OS differed significantly across clusters (P < 0.0001). When each cluster was assessed separately, cluster d had the poorest prognosis, followed by cluster b. A forest plot for survival with cox hazard models adjusted by sex showed the highest HR in cluster d (HR 12.4, 95% CI: 4.85–31.6), followed by cluster b (HR 3.2, 95% CI: 1.43–7.2) against the reference cluster a. However, there was no significant difference between clusters a and c with respect to HR (Fig. 3). Interestingly, all causes of death in clusters a and c with lower FIB-4 index were non-liver-related events (Table 2). Conversely, the main cause of death in cluster d with the highest FIB-4 was liver-related death (72.7%). Liver-related and non-liver-related death in cluster b, which presented the second highest FIB-4 and the lowest BMI, were 36.0% (n = 9) and 64.0% (n = 16), respectively. Finally, we investigated liver- versus non-liver-related mortality by clusters with Gray's test (Fig. S5). We found significant differences for the occurrence of liver-related mortality (P < 0.0001), but mortality rates due to non-liver-related diseases did not differ across clusters (P = 0.48).

Discussion

In our multicenter cohort of 1335 biopsy-proven NAFLD patients in Japan, we aimed to reveal the potential clusters combining FIB-4 index and BMI using machine learning models and to assess the characteristics and prognosis of each cluster. Our results demonstrated that the cluster with highest FIB-4 index (median 5.23) and moderate BMI (median 28.9 kg/m^2), which we named d, showed the poorest prognosis, followed by cluster b with the second highest FIB-4 (median 2.23) and lowest BMI (median 26.1 kg/m^2). In contrast, clusters a with lowest FIB-4 (median 0.78), and cluster c with moderate FIB-4 levels (median 1.30) and highest BMI (median 34.3 kg/m^2), presented lower mortality rates than clusters b and d. Furthermore, all deaths in clusters a and c were caused by non-liver-related diseases. Using a clustering approach, we therefore revealed that the FIB-4 index, an established fibrosis marker, is an important factor for predicting the occurrence of liver-related events in NAFLD, regardless of BMI. Our analysis further indicates that, in the case of NAFLD patients with low FIB-4 index, preventive screening should focus on non-liver-related diseases, rather than liver-related ones.

In the present study, liver fibrosis progression with high FIB-4 levels was more strongly associated with liver-related mortality as opposed to non-liver-related mortality. Younes et al. reported that, while various NITs (such as FIB-4 index and NFS) can predict long-term outcomes caused by liver-related complications and death, all NITs showed limited performance for extrahepatic events.[28] Additionally, a previous study in biopsy-proven NAFLD patients from Japan reported that the FIB-4 index carried limited potential for predicting the development of extrahepatic malignancies or CVD.[29] Using magnetic resonance elastography, Tamaki et al. reported that CVD events were more common in the minimal or moderate-to-advanced fibrosis stage than in the cirrhosis stage.[30] Considering these reports and our own results, we conclude that screening for liver-related diseases should be carried out in populations with a higher FIB-4 index, such as clusters b and d, while extrahepatic malignancies and CVD development should be monitored in populations with a low FIB-4 index, such as clusters a and c.

In addition to fibrosis state, BMI is also known to be an important component in examining the prognosis of NAFLD. In particular, the incidence of lean NAFLD with normal weight in Asia is higher than in Europe and in the United States.[17,18,31] A recent meta-analysis in Japan using data from 14 887 individuals

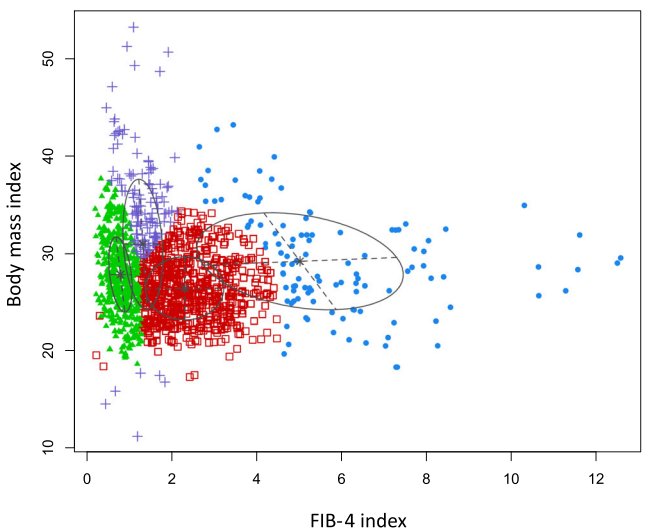

Figure 1 Scatter plot of estimated mixture densities for each cluster. FIB-4, fibrosis-4. Cluster a, green; cluster b, red; cluster c, purple; cluster d blue.

© 2023 Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

14401746, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jgh.16326 by Int. Uni Of Health & Welfare, Wiley Online Library on [22/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2 Patient characteristics by each cluster

Cluster	a (*n* = 465)	b (*n* = 631)	c (*n* = 126)	d (*n* = 113)	*P*-value
Age (years)	44 [35, 52]	63 [57, 69]	49 [40, 57]	66 [62, 72]	<0.001
Male	267 (57.4)	212 (33.6)	57 (45.2)	35 (31.0)	<0.001
BMI (kg/m^2)	27.6 [25.2, 30.3]	26.1 [24.1, 28.5]	34.3 [31.9, 37.7]	28.9 [26.1, 32.5]	<0.001
Lean	49 (10.5)	100 (15.8)	6 (4.8)	12 (10.6)	<0.001
Overweight	181 (38.9)	308 (48.8)	0 (0.0)	34 (30.1)	
Obese	235 (50.5)	223 (35.3)	120 (95.2)	67 (59.3)	
FIB-4 index	0.78 [0.57, 0.98]	2.23 [1.75, 2.95]	1.30 [1.12, 1.56]	5.23 [4.62, 7.09]	<0.001
<1.3	465 (100.0)	23 (3.7)	61 (48.8)	0 (0.0)	<0.001
1.3–2.67	0 (0.0)	393 (62.4)	64 (51.2)	1 (0.9)	
>2.67	0 (0.0)	214 (34.0)	0 (0.0)	112 (99.1)	
Fibrosis stage					<0.001
0	159 (34.2)	54 (8.6)	10 (7.9)	2 (1.8)	
1	230 (49.5)	228 (36.1)	41 (32.5)	16 (14.2)	
2	70 (15.1)	222 (35.2)	57 (45.2)	33 (29.2)	
3	6 (1.3)	118 (18.7)	16 (12.7)	49 (43.4)	
4	0 (0.0)	9 (1.4)	2 (1.6)	13 (11.5)	
Advanced fibrosis (≥F3)	6 (1.3)	127 (20.1)	18 (14.3)	62 (54.9)	<0.001
All death	8 (1.7)	25 (4.0)	2 (1.6)	11 (9.7)	
Liver-related death	0 (0.0)	9 (36.0)	0 (0.0)	8 (72.7)	
Non-liver-related death	8 (100.0)	16 (64.0)	2 (100.0)	3 (27.3)	
Observation time (years)*	6.2 [0.5–21.6]	6.6 [0.5–21.1]	4.1 [0.5–16.2]	3.9 [0.5–17.1]	<0.001

Values are expressed as median (first to third interquartiles), *median (range), and number (%).
BMI, body mass index; FIB-4, fibrosis-4.

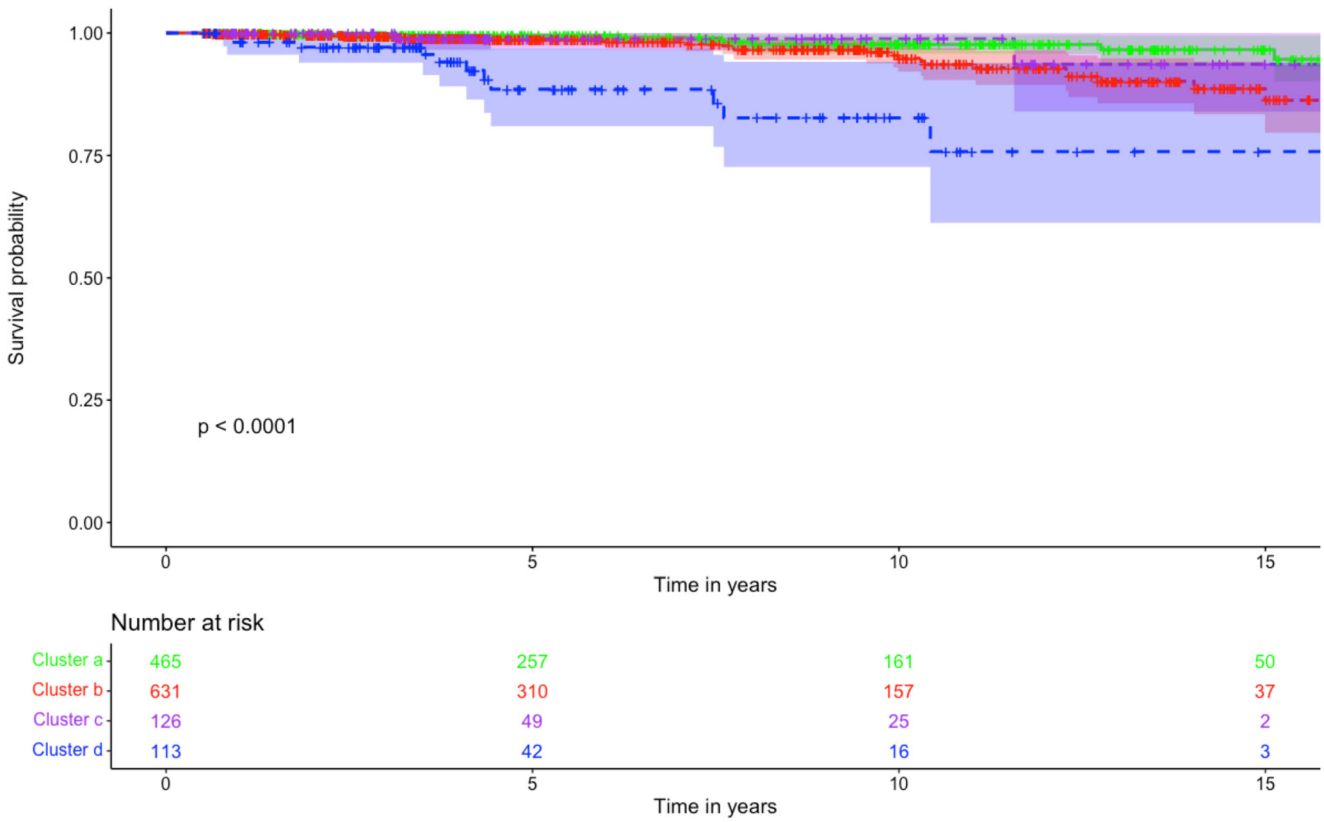

Figure 2 Kaplan–Meier plots of overall survival for different clusters. Among these clusters, there was a significant difference on survival ($P < 0.0001$). Cluster d had the poorest prognosis, followed by cluster b. ⊣, Cluster a; ⊣, Cluster b; ⊣, Cluster c; ⊣, Cluster d.

© 2023 Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

14401746, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jgh.16326 by Int. Uni Of Health & Welfare, Wiley Online Library on [22/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

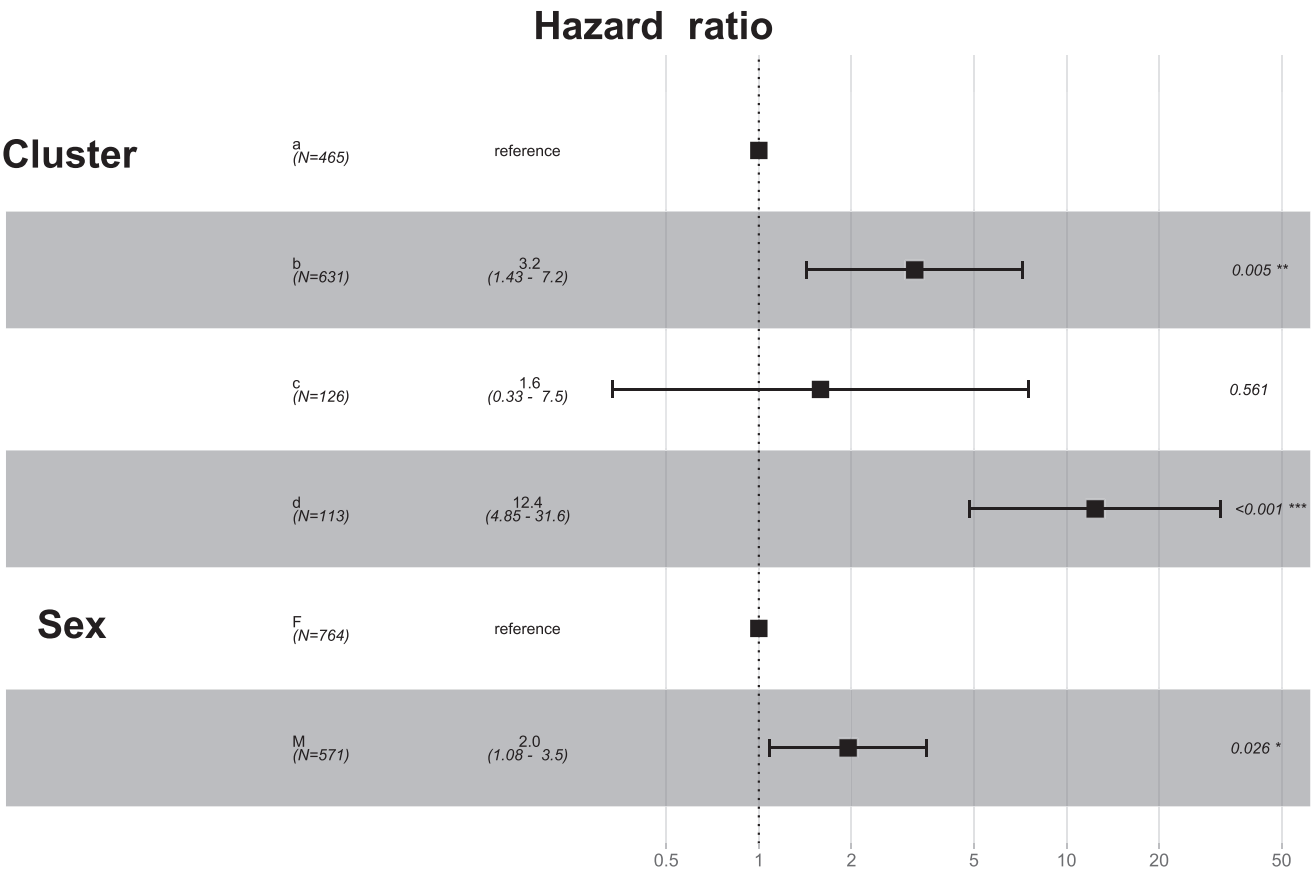

Figure 3 Forest plot of multivariate Cox hazard models for overall survival. Cluster d showed the highest hazard ratio (HR), followed by cluster b against the reference cluster a. However, there was no significant difference on HR between clusters a and c.

demonstrated that lean NAFLD (BMI < 23 kg/m^2) accounts for 20.7% of all NAFLD cases, with higher all-cause mortality compared with non-lean NAFLD. In relation to specific causes of death, this study revealed high rates of liver-related deaths within the non-lean group, while mortality within the lean NAFLD group was mostly caused by non-liver-related diseases.[16] Consistent with these results, Ahmed et al. previously reported that, in comparison with obese NAFLD, NAFLD with normal BMI (<25 kg/m^2) was associated with a trend toward milder liver disease progression but similar risk of CVD and malignancies and worse survival. Their study was based on the analysis of 4834 NAFLD patients from Minnesota (USA) and used a medical record linkage system.[32] Our results do not show that the lean state leads to significantly poorer prognosis; however, this apparent discrepancy may be due to differences in diagnostic methods (such as ultrasound *vs* biopsy). Furthermore, lean NAFLD is related to aging, genetic factors such as non-synonymous single nucleotide polymorphisms of the patatin-like phospholipase 3 rs738409,[33] the presence of sarcopenic visceral obesity,[34] and changes in intestinal microbiota.[35] Future studies will need to take these factors into account for the subclassification of lean NAFLD by risk for mortality.

Although both the FIB-4 index and BMI are continuous variables, most previous studies have used reported cutoff values for examining prognosis.[14,27,29] However, the optimal cutoff values of the NITs for the diagnosis of advanced fibrosis in NAFLD are still being debated. Indeed, the previously published cutoff values show different results for predictive abilities, which brings into question their generalizability to clinical applications in primary care.[10] In particular, the FIB-4 index is strongly influenced by age and metabolic factors,[13,36] and its diagnostic ability and suitable cutoff values may require individualized adjustment. For the lean or obese state, the definition based on the WHO criteria for BMI is generally used, but these cutoff values also vary between populations in Asia and the West.[27] Clustering analysis using GMM is more objective because it supports classification of FIB-4 index and BMI as continuous variables without potentially biased intervention.[20] Unsupervised machine learning methods can detect potential phenotypes by including multiple inputs, such as FIB-4 and BMI in our study. Therefore, our study identified a possible clinical picture that aligns with the picture engaged by clinicians when treating patients. GMM is a powerful method that was also able to expose potential phenotypes within an observational cohort from a previous report.[37,38] GMM is a powerful tool for phenotyping patients into interpretable groups on a study-by-study basis with different analyses and factors.

Our study presents several strengths. To the best of our knowledge, this is the first report investigating the prognosis of NAFLD

© 2023 Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

14401746, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jgh.16326 by Int. Uni Of Health & Welfare, Wiley Online Library on [22/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

with the combined parameters of FIB-4 index and BMI using machine learning methods. Machine learning methods, especially unsupervised clustering analyses, can discover more homogeneous phenotypes from heterogeneous datasets.[39] Furthermore, our study was based on a large NAFLD dataset with pathological findings diagnosed by an experienced liver pathologist.

Our study also presents some limitations. First, it only includes patients from Japan, so our results may not generalize to patients in other regions and/or belonging to different racial/ethnic groups. Second, the recommendation of a liver biopsy for NAFLD patients was at the discretion of their treating physicians, which can lead to patient selection bias in our study. Actually, our study included 213 cases (16.0%) with advanced fibrosis (F3–4), and the prevalence of cases with fibrosis may be higher than that in the general population. Tada et al. previously reported that most deaths of NAFLD patients from Japan diagnosed with ultrasonography are caused by non-liver-related events.[40] Therefore, our results on the cause of death differ from those in the general population. However, we believe that our results have a clinical impact to follow-up NAFLD patients in the secondary and tertiary referral hospitals. Lastly, while our study includes a large cohort of over 1300 patients, the follow-up period was relatively short (median 4.9 years).

Conclusion

We used an unsupervised machine learning method to cluster patients using the FIB-4 index and BMI. This approach was able to predict prognosis in biopsy-proven Asian NAFLD patients. In cases with advanced fibrosis like clusters b and d, surveillance for liver-related diseases is needed. On the other hand, non-liver-related diseases consisted the causes of death in patients with low FIB-4 index, such as clusters a and c, though the number of deaths was small (under $n = 10$). For the subgroup with low FIB-4, it may be unnecessary to follow up for liver-related diseases strictly. In addition, the present study found that the prognosis for NAFLD is similar between lean and obese patients. Further large-scale studies with longer follow-up periods that include both Asian and Western populations are needed to validate our results, possibly using the same or similar machine learning methods.

Acknowledgments

We wish to thank all those who contributed to the collection of data for this study. We also express gratitude for the English language support provided Edanz (https://jp.edanz.com/ac).

Ethical approval

This study was conducted in accordance with the tenets of the Declaration of Helsinki and approved by the Institutional Review Board of Saga University Hospital (approval no. 2020-04-R-02) and at each participating study center. Informed consent for this study was obtained from the website of each participating study center.

Data availability statement. The data that support the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy or ethical restrictions.

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**: 73–84.
2 Estes C, Anstee QM, Arias-Loste MT *et al*. Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016-2030. *J. Hepatol.* 2018; **69**: 896–904.
3 Chalasani N, Younossi Z, Lavine JE *et al*. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of Liver Diseases. *Hepatology* 2018; **67**: 328–57.
4 Angulo P, Kleiner DE, Dam-Larsen S *et al*. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology* 2015; **149**: 389–97.
5 Ekstedt M, Hagstrom H, Nasr P *et al*. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology* 2015; **61**: 1547–54.
6 Dulai PS, Singh S, Patel J *et al*. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: systematic review and meta-analysis. *Hepatology* 2017; **65**: 1557–65.
7 Fujii H, Iwaki M, Hayashi H *et al*. Clinical outcomes in biopsy-proven nonalcoholic fatty liver disease patients: a multicenter registry-based cohort study. *Clin. Gastroenterol. Hepatol.* 2023; **21**: 370–9.
8 Bruzzi JF, O'Connell M, Murray JG. Liver biopsy. *N. Engl. J. Med.* 2001; **344**: 2030.
9 Myers RP, Fong A, Shaheen AA. Utilization rates, complications and costs of percutaneous liver biopsy: a population-based study including 4275 biopsies. *Liver Int.* 2008; **28**: 705–12.
10 Xiao G, Zhu S, Xiao X, Yan L, Yang J, Wu G. Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: a meta-analysis. *Hepatology* 2017; **66**: 1486–501.
11 Selvaraj EA, Mozes FE, Jayaswal ANA *et al*. Diagnostic accuracy of elastography and magnetic resonance imaging in patients with NAFLD: a systematic review and meta-analysis. *J. Hepatol.* 2021; **75**: 770–85.
12 European Association for the Study of the Liver. EASL clinical practice guidelines on non-invasive tests for evaluation of liver disease severity and prognosis - 2021 update. *J. Hepatol.* 2021; 659–89.
13 Ito T, Nguyen VH, Tanaka T *et al*. Poor diagnostic efficacy of noninvasive tests for advanced fibrosis in obese or younger than 60 diabetic NAFLD patients. *Clin. Gastroenterol. Hepatol.* 2023; **21**: 1013–22.
14 Boursier J, Hagstrom H, Ekstedt M *et al*. Non-invasive tests accurately stratify patients with NAFLD based on their risk of liver-related events. *J. Hepatol.* 2022; **76**: 1013–20.
15 Younossi ZM, Golabi P, de Avila L *et al*. The global epidemiology of NAFLD and NASH in patients with type 2 diabetes: A systematic review and meta-analysis. *J. Hepatol.* 2019.
16 Ito T, Ishigami M, Zou B *et al*. The epidemiology of NAFLD and lean NAFLD in Japan: a meta-analysis with individual and forecasting analysis, 1995-2040. *Hepatol. Int.* 2021; **15**: 366–79.
17 Ye Q, Zou B, Yeo YH *et al*. Global prevalence, incidence, and outcomes of non-obese or lean non-alcoholic fatty liver disease: a systematic review and meta-analysis. *Lancet Gastroenterol. Hepatol.* 2020; **5**: 739–52.

© 2023 Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

14401746, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jgh.16326 by Int. Uni Of Health & Welfare, Wiley Online Library on [22/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

18 Li J, Zou B, Yeo YH *et al*. Prevalence, incidence, and outcome of non-alcoholic fatty liver disease in Asia, 1999-2019: a systematic review and meta-analysis. *Lancet Gastroenterol. Hepatol.* 2019; **4**: 389–98.
19 Sookoian S, Pirola CJ. Systematic review with meta-analysis: risk factors for non-alcoholic fatty liver disease suggest a shared altered metabolic and cardiovascular profile between lean and obese patients. *Aliment. Pharmacol. Ther.* 2017; **46**: 85–95.
20 Scrucca L, Fop M, Murphy TB, Raftery AE. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *R J.* 2016; **8**: 289–317.
21 Kleiner DE, Brunt EM, Van Natta M *et al*. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005; **41**: 1313–21.
22 Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am. J. Gastroenterol.* 1999; **94**: 2467–74.
23 Heimbach JK, Kulik LM, Finn RS *et al*. AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology* 2018; **67**: 358–80.
24 Sterling RK, Lissen E, Clumeck N *et al*. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 2006; **43**: 1317–25.
25 Angulo P, Hui JM, Marchesini G *et al*. The NAFLD fibrosis score: A noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology* 2007; **45**: 846–54.
26 Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ, Nash Clinical Research Network. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clin. Gastroenterol. Hepatol.* 2009; **7**: 1104–12.
27 Consultation WHOE. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* 2004; **363**: 157–63.
28 Younes R, Caviglia GP, Govaere O *et al*. Long-term outcomes and predictive ability of non-invasive scoring systems in patients with non-alcoholic fatty liver disease. *J. Hepatol.* 2021; **75**: 786–94.
29 Ito T, Ishigami M, Ishizu Y *et al*. Utility and limitations of noninvasive fibrosis markers for predicting prognosis in biopsy-proven Japanese non-alcoholic fatty liver disease patients. *J. Gastroenterol. Hepatol.* 2019; **34**: 207–14.
30 Tamaki N, Higuchi M, Kurosaki M, Loomba R, Izumi N, Group MLS. Risk difference of liver-related and cardiovascular events by liver fibrosis status in nonalcoholic fatty liver disease. *Clin. Gastroenterol. Hepatol.* 2022; **20**: 1171–3.e2.
31 Zou B, Yeo YH, Nguyen VH, Cheung R, Ingelsson E, Nguyen MH. Prevalence, characteristics and mortality outcomes of obese, nonobese and lean NAFLD in the United States, 1999-2016. *J. Intern. Med.* 2020; **288**: 139–51.
32 Ahmed OT, Gidener T, Mara KC, Larson JJ, Therneau TM, Allen AM. Natural history of nonalcoholic fatty liver disease with normal body mass index: a population-based study. *Clin. Gastroenterol. Hepatol.* 2022; **20**: 1374–81 e6.
33 Kawaguchi T, Sumida Y, Umemura A *et al*. Genetic polymorphisms of the human PNPLA3 gene are strongly associated with severity of non-alcoholic fatty liver disease in Japanese. *PLoS ONE* 2012; **7**: e38322.
34 Cho Y, Chang Y, Ryu S *et al*. Skeletal muscle mass to visceral fat area ratio as a predictor of NAFLD in lean and overweight men and women with effect modification by sex. *Hepatol. Commun.* 2022.
35 Chen F, Esmaili S, Rogers GB *et al*. Lean NAFLD: a distinct entity shaped by differential metabolic adaptation. *Hepatology* 2020; **71**: 1213–27.
36 McPherson S, Hardy T, Dufour JF *et al*. Age as a confounding factor for the accurate non-invasive diagnosis of advanced NAFLD fibrosis. *Am. J. Gastroenterol.* 2017; **112**: 740–51.
37 Morooka H, Tanaka A, Inaguma D, Maruyama S. Clustering phosphate and iron-related markers and prognosis in dialysis patients. *Clin. Kidney. J.* 2022; **15**: 328–37.
38 Yamamoto T, Morooka H, Ito T *et al*. Clustering using unsupervised machine learning to stratify the risk of immune-related liver injury. *J. Gastroenterol. Hepatol.* 2023; **38**: 251–8.
39 Wang M, Abrams ZB, Kornblau SM, Coombes KR. Thresher: determining the number of clusters while removing outliers. *BMC Bioinf.* 2018; **19**: 9.
40 Tada T, Kumada T, Toyoda H *et al*. Progression of liver fibrosis is associated with non-liver-related mortality in patients with nonalcoholic fatty liver disease. *Hepatol. Commun.* 2017; **1**: 899–910.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Kaplan–Meier plots of overall survival by fibrosis-4 (FIB-4) index. FIB-4 index <1.3 was labelled as low (n = 550), 1.3–2.67 as intermediate (n = 459), and ≥2.67 as high (n = 326).
Figure S2. Kaplan–Meier plots of overall survival by body mass index (BMI). BMI < 23 was defined as lean (n = 167), 23–27.5 as overweight (n = 523), and >27.5 as obese (n = 645) based on World Health Organization (WHO) criteria for the Asian population.
Figure S3. Plot of the Bayesian information criterion (BIC). VVV, ellipsoidal, varying volume, shape and orientation.
Figure S4. Tukey box plot. Baseline fibrosis-4 (FIB-4) index and body mass index (BMI) of each cluster.
Figure S5. Cumulative incidence rates of liver-related death and non-liver-related death for different clusters.

© 2023 Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

14401746, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jgh.16326 by Int. Uni Of Health & Welfare, Wiley Online Library on [22/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License