RESEARCH



Machine learning for predicting all-cause mortality of metabolic dysfunction-associated fatty liver disease: a longitudinal study based on NHANES

Xueni Wang¹, Huihui Chen¹, Luqiao Wang¹ and Wenguang Sun^{2*}

Abstract

Background The mortality burden of metabolic dysfunction-associated fatty liver disease (MAFLD) is rising, making it crucial to predict mortality and identify the factors influencing it. While advanced machine learning algorithms are gaining recognition as effective tools for clinical prediction, their ability to predict all-cause mortality of MAFLD individuals remains uncertain. This study aimed to develop different machine learning models to predict all-cause mortality of MAFLD individuals, compare the predictive performance of these models, and identify the risk factors contributing all-cause mortality, which is crucial for management of MAFLD individuals.

Methods We included 3921 MAFLD individuals in NHANES III. After a median follow-up time of 310 months, 1815 (46.3%) deaths were recorded. The data (demographic, behavioral factors and laboratory indicators) were utilized to construct machine learning models (Coxnet, RSF, GBS) after feature selection. Time-dependent AUC, time-dependent brier and C-index were then evaluated the performance of models. We identified the top five factors that contributed significantly to all-cause mortality and further explore the association with all-cause mortality using RCS and Kaplan–Meier survival curves.

Results Coxnet showed the best performance in short-term and long-term predictions with time-dependent AUC of 0.82 at 5 years and 0.88 at 25 years. Age, FORNS, waist circumstance, AAR, FLI were associated positively with all-cause mortality. Compared to the individuals who smoked more than 100 cigarettes, those below 100 had better survival outcome (P < 0.0001).

Conclusions Machine learning has a promising application in predicting all-cause mortality in MAFLD individuals. Combined the results of interpretable machine learning and association analyses, we found risk factors which contributing to the all-cause mortality. These findings provide insights for community health practitioners to intervene in modifiable risk factors, thereby improving the survival and quality of life of MAFLD individuals.

Keywords All-cause mortality, Machine learning, Metabolic dysfunction-associated fatty liver disease, Fatty liver

*Correspondence: Wenguang Sun wgsun@zju.edu.cn Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Metabolic dysfunction-associated fatty liver disease (MAFLD) is the most common cause of chronic liver disease, with a global prevalence reaching up to 25%. [1]. In 2020, an international panel expert consensus panel recommended renaming non-alcoholic fatty liver disease (NAFLD) to MAFLD [2], compared to NAFLD, to better reflect this highly prevalent liver disease affected by metabolic pathophysiology and cardiometabolic implications than the original NAFLD [3], and also acknowledged the heterogeneity of fatty liver disease, facilitating more precise phenotyping and supporting individualized management strategies in clinical practice [4].

MAFLD is often associated with other metabolic disorders such as obesity and type 2 diabetes mellitus (T2DM), and may be accompanied by other metabolic risk factors such as hyperglycemia and hypertension, which are detrimental to the progression of the disease and often lead to an increase in the mortality rate associated with the disease [5]. MAFLD exacerbates the burden of all-cause mortality. Studies conducted in the U.S. have found that individuals with MAFLD have a 17% higher risk of allcause mortality compared to those without fatty liver disease [6, 7]. Additionally, on a global scale, the mortality burden of MAFLD has been rising, with mortality rates showing an upward trend from 1990 to 2021 [8].

Previous studies have explored the relationship between modifiable behavioral factors, metabolic phenotypes, and all-cause mortality in patients with MAFLD, [9-12] as well as other mortality-specific relationships, including cardiovascular disease (CVD) and malignant tumors. These studies identified risk factors contributing to all-cause mortality in MAFLD individuals, including age, sex, married status, alcohol consumption, smoking, body mass index (BMI), FIB-4, and others [13, 14]. However, few studies have explored the predictive value of these demographic information, modifiable behavioral factors and laboratory indicators for all-cause mortality in patients with MAFLD.

Advanced machine learning algorithms have been widely used in medicine, and unlike traditional statistical methods, they can effectively eliminate confounding factors and improve predictive accuracy, assisting health-care professionals in identifying high-risk patients and increasing the accuracy of predictions of diseases and their adverse outcomes [15–17]. And it also shows powerful performance in predicting mortality in clinical scenarios. [18–20] Recent years, studies have used machine learning methods to construct models for identifying populations at high risk of developing MAFLD models, [21, 22] as well as predicting its progression [23, 24].

To the best of our knowledge, few studies have used advanced machine learning models to predict all-cause mortality in individuals with MAFLD. Among the available research, there is a lack of comprehensive demographic information and clinical laboratory biomarkers, as well as limited diversity in predictive model construction approaches [25, 26].

Given the current inability to identify a treatment for MAFLD, it is crucial to predict mortality and identify the factors that influence it. In this study, we utilized large prospective cohort database with potential risk factors to construct a machine learning model for predicting all-cause mortality in MAFLD individuals, comparing multiple machine learning methods to determine the optimal approach. Additionally, we introduced some new indices, such as triglyceride (TG) and obesity-related indices, and the non-invasive test of fatty liver indices as predictive variables. By predicting all-cause mortality risk and identifying risk factors, healthcare providers can intervene timely for MAFLD individuals.

Method

Study population

Data used in this were derived from the third National Health and Nutrition Examination Survey (NHANES III). NHANES-III is a multistage stratified survey conducted from 1988–1994, gathering representative health-related data on repetitive noninstitutionalized US population through household interview and medical examination. Further information about NHANES III database can be found on the website (https://www.cdc.gov/nchs/nhanes/).

Participants

Study was carried out in individuals above 20 years old. We excluded individuals without gall bladder ultrasound video images and hepatic steatosis assessments. We included the MAFLD population based on diagnosis criteria. This analysis included follow-up data collected up to December 31, 2019. Those individuals without followup information were not included in the current study.

Diagnosis criteria and Definitions *MAFLD*

The diagnosis criteria of metabolic associated fatty liver disease (MAFLD) is based on ultrasound images and blood biomarker evidence of hepatic steatosis in addition to one of the following three criteria, including overweight/obesity, T2DM, or presence of at least 2 risk factors of metabolic dysregulation, which was defined in following conditions: (a) Waist circumference ≥ 102 cm in men and 88 cm in women. (b) Blood pressure $\geq 130/85$ mmHg or specific drug treatment. (c) TG ≥ 1.70 mmol/L or specific drug treatment. (d) high density lipoprotein cholesterol (HDL-C) <1.0 mmol/L for male and <1.3

Outcome

The survival data were obtained from the NCHS (https:// www.cdc.gov/nchs/data-linkage/mortality.htm). The all-cause mortality was defined according to the ICD-10 classification system.

Predictors

The covariates analyzed in this study include 11 variables based on questionnaire, including 6 demographic variables (sex, age, ethnicity, education level, married status, family income-poverty ratio level),3 variables related of self-report disease history (heart disease, hypertension, diabetes) and 2 variables of behavioral information (smoking status and alcoholic consumption). Grade of hepatic steatosis, height, weight, waist circumstance, BMI, FBG, glycosylated hemoglobin (GHb), TG, HDL, Glutamyl transpeptidase (GGT), Alanine aminotransferase (ALT), Alanine aspartatetransaminase (AST), platelet counts, total cholesterol, albumin, white blood cell count (WBC), Hematocrit (HCT), total bilirubin were obtained from examination-based information. We further calculated 5 indices refer to the non-invasive test of fatty liver, including FLI, FIB-4, FORNS, APRI, AAR, m_APRI, AARPRI, 8 triglyceride glucose - related and obesity indices, including LAP, ABSI, BRI, VAI, TyG, TyGBMI, TyGWC, TyGWhtR, and HOMA-IR. These additional calculated indices can be found in the Supplementary material (Table S1, see Additional file 1). The distribution of continuous predictors was shown in Figure S1 (see Additional file 1).

Statistical analysis

We checked the missing rate of all features. Variables with more than 30% missing data were excluded, and MICE was applied to fill missing data for the remaining variables. The dataset was divided into training and test sets in an 8:2 ratio, with stratification by mortality status. Continuous variables were described as mean \pm standard deviation and compared using Student's t-test, or described as median (Q25, Q75) and compared using Mann–Whitney tests, according to the distribution. Categorical variables were described as percentages (%) and compared using the chi-square test. *P* < 0.05 was considered statistically significant.

All continuous variables were standardized using z-scores. For feature selection, we conducted Cox

employed to identify the features feeding into modelling. In this study, we employed 3 common machine learning algorithms including Coxnet, Random Survival Forest (RSF), Gradient Boosted Survival (GBS). In the Coxnet model, 10-fold cross-validation was used to determine the optimal parameter based on the C-index. For RSF and GBS models, we used grid search for parameter optimization in the training set. A 1000-time bootstrap was applied in the test set to compute confidence intervals of evaluation indicators at various time points (5, 10, 15, 20, and 25 years), including time-dependent AUC, timedependent Brier score, and the C-index, which we use to evaluate the prediction models.

wise regression Cox proportional hazards regression was

An explainable AI technique, the Sharpley Additive explanation (SHAP) was applied to explain the model, allowing for the ranking of feature importance. Additionally, Restricted Cubic Splines (RCS), and the Kaplan– Meier survival curves and Log-rank test were utilized to analyze the relationship between the top five ranked variables and all-cause mortality in MAFLD patients.

Data processing and model construction were performed using RStudio 2023.12.0. Model interpretation with SHAP was conducted using Python 3.12.4.

Results

Baseline characteristics

The flow chart of inclusion and exclusion of patients and the modeling process was shown in Fig. 1. A total of 3921 patients with MAFLD met inclusion criteria. The mean age was 48.60 years (\pm 15.07), 49.50% were males. During 1045,639 person-months of follow-up (median follow-up, 310.00 months), 1815 all-cause deaths occurred. The characteristics of patients of the training and test set were shown in Table 1. We compared the data from the training and test sets and found no significant differences in their distributions.

Model development

The feature select results for predicting all-cause mortality were shown in Table S2, Figure S2 and Table S3 (see Additional file 1). 22 significant features were then used for model construction, as a result of feature selection. The performance of 3 models at different follow-up times for predicting all-cause mortality in the test set was shown in Table 2. Details of the model parameters are summarized in Table S4 (see Additional file 1).



Fig. 1 The flowchart of this study

For all-cause mortality, the GBS model consistently showed high time-dependent AUC across all followup periods, with 0.86 (95% CI: 0.79–0.92) at 5 years, 0.89 (95% CI: 0.85–0.92) at 10 years, and 0.92 (95% CI: 0.90–0.94) at 25 years. However, despite its strong AUC performance, the GBS model had relatively low C-index values, ranging from 0.62 (95% CI: 0.51–0.73) at 5 years to 0.68 (95% CI: 0.64–0.71) at 25 years, indicating less consistency in its performance. For short-term follow-up, the RSF model demonstrated better consistency with a time-dependent AUC of 0.83 (95% CI: 0.75–0.89) at 5 years and a C-index of 0.82 (95% CI: 0.74–0.89), suggesting good predictive performance in the early stages. Meanwhile, the Coxnet model also performed well in early follow-up, with a time-dependent AUC of 0.82 (95% CI: 0.73–0.88) and a C-index of 0.81 (95% CI: 0.73–0.88) at 5 years.

Over the long-term follow-up, the Coxnet model had the best discrimination with highest time-dependent AUC of 0.88 (95% CI: 0.85–0.90) at 25 years and also a

Table 1 Baseline Characteristics and follow-up data in this study cohort

Variables	Total (n = 3921)	Train set (n = 3137)	Test set (n = 784)	<i>P</i> value
Age	48.6 ± 15.07	48.59 ± 15.09	48.68 ± 15.03	0.899
Sex				0.250
Male	1941 (49.50%)	1538 (49.0%)	403 (51.4%)	
Female	1980 (50.50%)	1599 (51.0%)	381 (48.6%)	
Education				0.754
Illiterate	144 (3.70%)	112 (3.57%)	32(4.08%)	
Below high school	1019(26.0%)	807 (25.7%)	212 (27.0%)	
High school	1848 (47.1%)	1488 (47.4%)	360 (45.9%)	
Above high school	910 (23.2%)	730 (23.3%)	180 (23.0%)	
ethnicity				0.586
non-Hispanic White	2812(71.70%)	2259(72.0%)	553 (70.5%)	
non-Hispanic Black	941 (24.00%)	748 (23.8%)	193 (24.6%)	
Other	168 (4.30%)	130 (4.14%)	38 (4.85%)	
Married status				0.407
Married (spouse in household)	2684 (68.50%)	2142 (68.3%)	542 (69.1%)	
Married (spouse not in household)	318 (8.10%)	244 (7.78%)	74 (9.44%)	
Living as married	333 (8.50%)	274 (8.73%)	59 (7.53%)	
Widowed	162 (4.10%)	132 (4.21%)	30 (3.83%)	
Divorced	424 (10.80%)	345 (11.0%)	79 (10.1%)	
Family income-poverty ratio				0.640
< 1.3	1470(37.40%)	1172 (37.36%)	298 (38.01%)	
1.3–3.5	1670(42.59%%)	1347 (42.94%)	323 (41.20%)	
> 3.5	781(19.91%)	618 (19.70%)	163 (20.79%)	
BMI				0.849
>=25	3466 (88.40%)	2775 (88.46%)	691 (88.14%)	
< 25	455 (11.60%)	362 (11.54%)	93 (11.86%)	
Waist	94.00 [102.00,111.00]	102.00 [94.00,111.90]	101.85 [93.78,111.70]	0.875
Weight	82.80 [72.15,95.80]	82.90[72.10,95.80]	82.35[72.79,95.83]	0.986
Height	166.00 [158.90,173.40]	165.90[158.90,173.40]	166.40[158.30,173.53]	0.732
Hypertension				0.387
Yes	1540 (39.30%)	1221 (38.9%)	319 (40.7%)	
No	2381 (60.70%)	1916 (61.1%)	465 (59.3%)	
Number of cigarettes smoked > 100				0.522
Yes	2048 (52.20%)	1630 (52.0%)	418 (53.3%)	
No	1873(47.80%)	1507 (48.0%)	366 (46.7%)	
Alcohol consumption	1.415[1.201,2.110]	1.415 [1.201,2.110]	1.415[1.201,2.110]	0.568
Diabetes status				0.966
Normal	1104(28.20%)	1300 (41.4%)	321 (40.9%)	
Prediabetes	1196(30.50%)	956 (30.5%)	240 (30.6%)	
Diabetes	1621 (41.3%)	881 (28.1%)	223 (28.4%)	
Reported heart disease				0.846
Yes	208 (5.30%)	168 (5.36%)	40 (5.10%)	
No	3713(94.70%)	2969 (94.6%)	744 (94.9%)	
Grade of hepatic steatosis	. ,	. ,	. ,	0.594
Normal/mild	1284(32.75%)	1039 (33.12082%)	245 (31.25000%)	
Moderate	1747(44.55%)	1392 (44.37361%)	355 (45.28061%)	
Sever	890(22.70%)	706 (22.50558%)	184 (23.46939%)	
LAP	71.11[44.79,113.16]	71.21[44.88;113.82]	70.59[44.22;109.48]	0.827
ABSI	0.082[0.079,0.085]	0.082 [0.079,0.085]	0.082[0.079,0.085]	0.661

Table 1 (continued)

Variables	Total (n = 3921)	Train set (n = 3137)	Test set (n = 784)	P value
BRI	5.682[4.617,7.075]	5.704[4.628,7.060]	5.622[4.569,7.121]	0.904
CI	0.079[0.075,0.084]	0.079[0.075,0.084]	0.079[0.075,0.084]	0.890
VAI	2.515[1.536,4.173]	2.537[1.543,4.179]	2.450[1.476,4.121]	0.326
FLI	76.49[53.39,91.68]	76.64 [53.27,91.81]	76.15 [53.66,91.44]	0.920
FIB-4	0.874[0.593,1.278]	0.871[0.591,1.269]	0.878[0.598;1.313]	0.521
FORNS	6.29[5.01,6.29]	6.33 [5.07,7.49]	6.28 [5.00,7.36]	0.323
APRI	0.328[0.252,0.441]	0.327[0.252,0.439]	0.334 [0.255,0.449]	0.336
m_APRI	2.323[1.545,3.400]	2.302[1.5384,3.351]	2.393[1.563,3.590]	0.181
AARPRI	0.643[0.472,0.873]	0.639 [0.471;0.873]	0.655 [0.481,0.872]	0.588
AAR	1.17[0.90,1.50]	1.17 [0.91,1.50]	1.17 [0.90,1.47]	0.402
glycosylated hemoglobin (GHb)	5.50[5.10,6.00]	5.50 [5.10,6.00]	5.50 [5.10,6.00]	0.736
Serum triglyceride	1.74[1.19,2.61]	1.75 [1.19,2.61]	1.71 [1.16,2.52]	0.565
Glutamyl transpeptidase (GGT)	29.00[19.00,45.00]	29.00 [19.00,45.00]	29.00 [19.00,46.00]	0.893
Alanine aminotransferase (ALT)	18.00[13.00,27.00]	18.00 [13.00,27.00]	18.00 [13.00,27.00]	0.166
Aspartate transaminase (AST)	21.0[17.00,27.00]	21.00 [17.00,27.00]	21.00 [17.00,27.00]	0.213
High-Density Lipoprotein Cholesterol (HDL-C)	1.14[0.93,1.34]	1.11[0.93,1.34]	1.14[0.96,1.37]	0.370
Platelet counts	271.50[229.50,321.00]	272.00[229.50,322.00]	270.00[228.38,314.63]	0.141
White blood cell	7.20[5.95,9.65]	7.15 [6.00,8.60]	7.20[5.95,8.80]	0.370
Hematocrit	0.422[0.391,0.449]	0.422 [0.391,0.448]	0.422 [0.392,0.451]	0.440
Total Bilirubin	8.55[6.84,11.97]	8.55[6.84,11.97]	8.55[6.84,11.97]	0.345
Total cholesterol	5.46[4.76,6.21]	5.46[4.76,6.21]	5.46 [4.71;6.18]	0.726
Albumin	4.10[3.90,4.30]	4.10 [3.90,4.30]	4.10 [3.90,4.40]	0.975
Time (months)	310[206,336]	310 [207,335]	310 [206,337]	0.881
Death	1815(46.30%)	1452 (46.30%)	363 (46.29%)	1.000

Continuous variables were presented as mean (standard deviation) or median [interquartile range]. Categorical variables were presented as n (%). the *P* value was calculated by the chi-square test and Wilcoxon rank sum test

 Table 2
 Performance of the models for predicting all-cause mortality in the test set

Models	Time-dependent AUC	Time-dependent brier	C-index
Coxnet			
Time = 5 years	0.82(0.73,0.88)	0.04(0.03,0.05)	0.81(0.73,0.88)
Time = 10 years	0.84(0.79,0.88)	0.08(0.07,0.09)	0.82(0.77,0.87)
Time = 15 years	0.86(0.83,0.89)	0.11(0.10,0.12)	0.83(0.80,0.86)
Time = 20 years	0.87(0.84,0.90)	0.12(0.11,0.14)	0.82(0.80,0.85)
Time = 25 years	0.88(0.85,0.90)	0.13(0.12,0.15)	0.82(0.80,0.84)
Random survival forest(RSF)			
Time = 5 years	0.83(0.75,0.89)	0.04(0.03,0.05)	0.82(0.74,0.89)
Time = 10 years	0.85(0.81,0.89)	0.08(0.07,0.10)	0.83(0.79,0.87)
Time = 15 years	0.86(0.83,0.89)	0.11(0.10,0.13)	0.83(0.79,0.86)
Time = 20 years	0.87(0.88,0.90)	0.13(0.12,0.14)	0.83(0.80,0.85)
Time = 25 years	0.87(0.85,0.90)	0.14(0.13,0.15)	0.82(0.79,0.84)
Gradient Boosted Survival (GB	M)		
Time = 5 years	0.86(0.79,0.92)	0.04(0.03,0.05)	0.63(0.52,0.74)
Time = 10 years	0.89(0.85,0.92)	0.07(0.06,0.08)	0.57(0.50,0.64)
Time = 15 years	0.90(0.87,0.93)	0.09(0.08,0.11)	0.61(0.55,0.66)
Time = 20 years	0.91(0.89,0.93)	0.11(0.09,0.12)	0.65(0.60,0.69)
Time = 25 years	0.92(0.90,0.94)	0.11(0.10,0.12)	0.68(0.64,0.71)

high C-index of 0.82 (95% CI: 0.79-0.84). Meanwhile, the time-dependent Brier score of Coxnet remained low throughout the follow-up period, indicating good calibration, as shown in Fig. 2(c).

The time-dependent ROC curves of models for predicting all-cause mortality in test set were shown in Fig. 3.

Model interpretation

We observed that predictions weighted the most important features regarding the decision of Coxnet (Fig. 4). Age, FORNS, waist, number of cigarettes smoked > 100, AAR, FLI were the top 5 important features. Meanwhile, the RCS for the top four important continuous variables were shown in Fig. 5. Notably, we observe that as they increase, they were accompanied by an increase in the all-cause mortality. The risk of all-cause mortality was significantly increased when age >48 years, FORNS >6.16, waist circumstance >96.33 cm, or AAR >1.18. The Kaplan–Meier survival curves (Fig. 6) showed that individuals who smoked fewer than 100 cigarettes had a consistently lower hazard rate throughout the follow-up period, indicating better survival outcomes compared to those who smoked more than 100 cigarettes. The difference between the two groups was significant, with a logrank test result of p < 0.0001.

Discussion

In this study, we developed advanced machine learning models for predicting all-cause mortality in MAFLD patients using data from the third National Health and Nutrition Examination Survey (NHANES III), and the three models trained (Coxnet, GBS, RSF) were evaluated and compared in an internal test set to confirm the predictive ability of the models as well as the reliability of the results. We additionally interpreted the prediction results of the optimal model using the SHAP method to visualize the impact of potential features on all-cause mortality in MAFLD individuals to increase the interpretability of the model at the global level. Overall, the Coxnet model demonstrated excellent predictive power in both predicting short-term and long-term mortality. We also found that in addition to age, FORNS, waist circumstance, AAR, and smoked cigarettes were the largest contributors to all-cause mortality.

To the best of our knowledge, this is the first study to use machine learning to predict all-cause mortality in MAFLD individuals, and most of the past studies aimed to explore the association between clinical characteristics and all-cause mortality in MAFLD, and to explore the value of the corresponding metrics in predicting the prognosis of survival in MAFLD. Although previous studies have used traditional machine learning models (e.g., Cox) to predict adverse outcomes of NAFLD individuals, such as all-cause mortality and hepatocellular carcinoma [27, 28], we believe it is necessary to investigate MAFLD separately due to the differences in clinical definitions between NAFLD and MAFLD. Past studies have identified the value of triglyceride- and glucoserelated [29, 30] (TyG, TyG-WHtR, TyG-BMI, TyG-WC), non-invasive liver test [31] (FLI, FIB-4), and waist circumference [32] in predicting survival in fatty liver individuals, so in this study we also included these potential variables, and additionally we introduced other relevant novel indices with the aim of identifying more risk variables with machine learning and increasing the predictive efficacy of the model.

Previous studies have demonstrated the strong potential of machine learning technology in predicting allcause mortality. Nascimento et al. utilized machine learning techniques to predict mortality caused by respiratory diseases, circulatory diseases, cancer, and other specific diseases [33]. Similarly, Tran et al. employed Bayesian networks to forecast two-year all-cause mortality in patients with chronic kidney diseases [34], Tan et al. developed an all-cause mortality prediction model for disabled and elderly populations using deep learning neural networks [35]. These studies collectively emphasized the capability of machine learning algorithms in accurately predicting both all-cause and disease-specific mortality rates. In the risk management of metabolic dysfunction-associated fatty liver disease (MAFLD), machine learning (ML) techniques also have demonstrated remarkable potential [36]. Deng et al. developed a risk identification model for high-risk MAFLD populations based on large-scale health examination data, which enables community health managers to perform preliminary screening and timely management of MAFLD more efficiently and cost-effectively in large populations [21]. Similarly, Cheung et al. applied ML methods to construct a fibrosis score for MAFLD patients, which significantly outperformed traditional non-invasive index in identifying advanced liver fibrosis [24]. Our study may contribute to identifying individuals with MAFLD at high risk of early adverse outcomes, thereby assisting healthcare providers in initiating timely interventions.

Faced with complex datasets, we employed feature selection prior to model training to reduce model complexity and mitigate the risk of overfitting. We utilized a sequential approach incorporating the filter method (Cox regression), embedded methods with penalized Cox models (Coxnet), and wrapper methods (multivariable Cox stepwise selection), which represent commonly used strategies for feature selection [25]. While the traditional Cox proportional hazards model is widely used for mortality prediction, it may struggle with high-dimensional data. To address these limitations, we incorporated



Fig. 2 Performance of different models predicting all-cause mortality of MAFLD. a Time-dependent AUC (b) C-index (c) Time-dependent brier. Abbreviations: COXNET, net-regularized cox proportional hazards model; GBS, gradient boosted



Fig. 3 Time-dependent AUROC of different models for predicting all-cause mortality (a)Coxnet, b Random Survival Forest, c Gradient Boosted Survival



Fig. 4 Summary plot of optimal model by SHAP

advanced methods into our analysis, including the Coxnet [37, 38], tree-based random survival forest model, and GBS algorithms [39–41]. These machine learning models are better equipped to handle high-dimensional data and capture non-linear relationships in survival analysis compared to traditional Cox models. Additionally, most existing mortality prediction studies consider death as a binary event without accounting for time-to-event data [42, 43]. In our study, we explicitly incorporated survival time, making full use of survival data to provide a more comprehensive analysis.

We found that the Coxnet model performed better than the other two machine learning algorithms in predicting both short-term and long-term mortality, consistent with findings by Duan et al. [44]. This suggests that advanced algorithms do not always outperform traditional ones [45], likely due to insufficient outcome events limiting their performance [46]. We observed that increased waist circumference and FORNS indices were associated with higher all-cause mortality risk in fatty liver patients [47, 48]. Although no studies have linked AAR to mortality in fatty liver, its association with liver fibrosis—a major cause of liver-related deaths—supports our findings. Additionally, consistent with Huang et al. [49], individuals who smoked over 100 cigarettes had higher mortality risks.

In our findings, age emerged as a significant factor influencing all-cause mortality in MAFLD patients.



Fig. 5 Restricted cubic spline for the top four continuous variables of optimal model for predicting all-cause mortality

However, modifiable behavioral factors, such as waist circumference and smoking quantity, also play critical roles in mortality risk. From a clinical perspective, our findings suggest that practitioners should pay greater attention to the liver fibrosis indicators like FORNS and AAR, both of which were widely used in related research and clinical management, which were also aligns with the recommendation of screening liver fibrosis for individuals with hepatic steatosis in clinical practice [50]. To further achieve clinical integration, developing the threshold of these key contributors needed to be collaborated with clinicians.

This study has several limitations. First, while we using 1000 bootstrap procedure in internal validation strengthens the credibility of our results, the generalizability and real-world applicability of our model for managing MAFLD individuals remain to be established. Therefore, prospective external validation using multicenter and large-scale datasets is essential to further confirm its clinical utility and support its future implementation. Second, the survival data lacked information on cause-specific mortality, preventing us from evaluating the model's performance in predicting specific causes of death in MAFLD. Third, although SHAP method enabled the identification of key factors contributing to the all-cause mortality of MAFLD individuals, and the associations between them were further explored, it is important to acknowledge that these findings were observational. Future experimental studies are warranted to elucidate the underlying biological mechanisms and validate these associations, which may facilitate the development of targeted interventions and precision management strategies for MAFLD individuals. Finally, future research should address gap by focusing on specific mortality causes and exploring the use of advanced deep learning models and



Fig. 6 Kaplan–Meier curves of categorical variable of number of cigarettes smoked

metaheuristic algorithms for model optimization and improved predictive accuracy.

Conclusion

In this study, we developed and compared three machine learning models for predicting all-cause mortality in MAFLD individuals, identifying the Coxnet model as the optimal choice. It is worth noting that while the regularization employed by the Coxnet model helps address challenges associated with high-dimensional data, its ability to capture complex nonlinear relationships may be limited compared to tree-based models such as RSF and GBM, suggesting a potential direction for future research to further enhance model performance. Using SHAP analysis, we highlighted the top five variables contributing most to all-cause mortality: age, FORNS, waist circumference, and AAR, which exhibited significant linear relationships and clear threshold effects, while smoking quantity showed distinct survival patterns. To the best of our knowledge, this is the first study to leverage prospective data and machine learning algorithms to predict all-cause mortality in MAFLD individuals. These findings provide valuable insights for community health practitioners to intervene in modifiable risk factors, thereby improving the survival and quality of life of MAFLD individuals.

Abbreviations

ALT	Alanine aminotransferase
AST	Alanine aspartatetransaminase
APRI	Non-invasive tests
AAR	Non-invasive tests
AAPRI	Non-invasive tests
ABSI	Obesity indices
BMI	Body mass index
BRI	Obesity indices
CRP	C-reactive protein
Coxnet	Net-regularized cox proportional hazards model
FBG	Fasting blood glucose
FIL	Non-invasive tests
FIB-4	Non-invasive tests
FORNS	Non-invasive tests
GGT	Glutamyl transpeptidase
GBS	Gradient boosted survival
GHb	Glycosylated hemoglobin
HDL-C	High density lipoprotein cholesterol
HOMA-IR	Homeostasis model assessment-insulin resistance
LAP	Obesity indices
MAFLD	Metabolic dysfunction-associated fatty liver disease
m_APRI	Non-invasive tests
NAFLD	Non-alcoholic fatty liver disease
NHANES III	The third National Health and Nutrition Examination Survey
RSF	Random Survival Forest
RCS	Restricted Cubic Splines
Shap	The Sharpley Additive explanation
T2DM	Type 2 diabetes mellitus
TG	Triglyceride
TyG	Triglyceride glucose – related indices
TyGBMI	Triglyceride glucose – related indices
TyGWC	Triglyceride glucose – related indices
TyGWhtR	Triglyceride glucose – related indices
VAI	Obesity indices

Acknowledgements

The authors thank the National Health and Nutrition Examination Survey program for providing the data for this study.

Authors' contributions

Study conception and design: XW. Acquisition of data: XW. Statistical analysis: XW, HC, LW. Analysis and interpretation of data: XW, HC. Drafting of manuscript: XW Critical revision: All. Supervision: WS. All authors approved the article's final version.

Funding

None.

Data availability

Data were derived from the following resources available in the public domain: https://www.cdc.gov/nchs/nhanes/index.htm. The survival data were obtained from the NCHS: https://www.cdc.gov/nchs/data-linkage/mortality. htm.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Supplementary Information

Supplementary Material 1.

Author details

¹Center of Clinical Big Data and Analytics of The Second Affiliated Hospital, and Department of Big Data in Health Science School of Public Health, Zhejiang University School of Medicine, Zhejiang University, Hangzhou, Zhejiang 310058, China. ²Center for Data Science, Zhejiang University, 866 Yuhangtang Rd, Room 1411, Hangzhou, Zhejiang 310058, China.

Received: 11 March 2025 Accepted: 28 April 2025 Published online: 15 May 2025

References

- Ma X, Jia J, Cui H, Zhou J, Tian F, Yang J, et al. Association between the triglyceride to high density lipoprotein cholesterol ratio and the incidence of metabolic dysfunction-associated fatty liver disease: a retrospective cohort study. BMC Gastroenterol. 2024;24(1):389.
- Eslam M, Newsome PN, Sarin SK, Anstee QM, Targher G, Romero-Gomez M, et al. A new definition for metabolic dysfunction-associated fatty liver disease: an international expert consensus statement. J Hepatol. 2020;73(1):202–9.
- Pennisi G, Infantino G, Celsa C, Di Maria G, Enea M, Vaccaro M, et al. Clinical outcomes of MAFLD versus NAFLD: a meta-analysis of observational studies. Liver Int. 2024;44(11):2939–49.
- Eslam M, Sanyal AJ, George J. MAFLD: a consensus-driven proposed nomenclature for metabolic associated fatty liver disease. Gastroenterology. 2020;158(7):1999-2014.e1.
- Kapoor N, Kalra S. Metabolic-associated fatty liver disease and diabetes: a double whammy. Endocrinol Metab Clin North Am. 2023;52(3):469–84.
- Kim D, Konyn P, Sandhu KK, Dennis BB, Cheung AC, Ahmed A. Metabolic dysfunction-associated fatty liver disease is associated with increased all-cause mortality in the United States. J Hepatol. 2021;75(6):1284–91.
- Cho SH, Kim S, Oh R, Kim JY, Lee YB, Jin SM, et al. Metabolic dysfunctionassociated fatty liver disease and heavy alcohol consumption increase mortality: a nationwide study. Hepatol Int. 2024;18(4):1168–77.
- Zhang H, Zhou XD, Shapiro MD, Lip GYH, Tilg H, Valenti L, et al. Global burden of metabolic diseases, 1990–2021. Metabolism. 2024;160: 155999.

- Åberg F, Puukka P, Salomaa V, Männistö S, Lundqvist A, Valsta L, et al. Risks of light and moderate alcohol use in fatty liver disease: follow-up of population cohorts. Hepatology. 2020;71(3):835–48.
- Åberg F, Helenius-Hietala J, Puukka P, Jula A. Binge drinking and the risk of liver events: a population-based cohort study. Liver Int. 2017;37(9):1373–81.
- 11. Younossi ZM, Stepanova M, Ong J, Yilmaz Y, Duseja A, Eguchi Y, et al. Effects of alcohol consumption and metabolic syndrome on mortality in patients with nonalcoholic and alcohol-related fatty liver disease. Clin Gastroenterol Hepatol. 2019;17(8):1625-33.e1.
- Charatcharoenwitthaya P, Karaketklang K, Aekplakorn W. Impact of metabolic phenotype and alcohol consumption on mortality risk in metabolic dysfunction-associated fatty liver disease: a population-based cohort study. Sci Rep. 2024;14(1):12663.
- Cheng WC, Chen HF, Cheng HC, Li CY. Comparison of all-cause mortality associated with non-alcoholic fatty liver disease and metabolic dysfunction-associated fatty liver disease in Taiwan MJ cohort. Epidemiol Health. 2024;46: e2024024.
- Zhu Y, Xu X, Fan Z, Ma X, Rui F, Ni W, et al. Different minimal alcohol consumption in male and female individuals with metabolic dysfunctionassociated fatty liver disease. Liver Int. 2024;44(3):865–75.
- Luo L, Gao P, Yang C, Yu S. Predictive modeling of COVID-19 mortality risk in chronic kidney disease patients using multiple machine learning algorithms. Sci Rep. 2024;14(1):26979.
- Fei J, Yong J, Hui Z, Yi D, Hao L, Sufeng M, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2(4):230.
- Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med. 2019;380(14):1347–58.
- Guo X, Ma M, Zhao L, Wu J, Lin Y, Fei F, et al. The association of lifestyle with cardiovascular and all-cause mortality based on machine learning: a prospective study from the NHANES. BMC Public Health. 2025;25(1):319.
- Wang K, Tian J, Zheng C, Yang H, Ren J, Liu Y, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. Comput Biol Med. 2021;137: 104813.
- Zhou H, Liu L, Zhao Q, Jin X, Peng Z, Wang W, et al. Machine learning for the prediction of all-cause mortality in patients with sepsis-associated acute kidney injury during hospitalization. Front Immunol. 2023;14: 1140755.
- Deng J, Ji W, Liu H, Li L, Wang Z, Hu Y, et al. Development and validation of a machine learning-based framework for assessing metabolic-associated fatty liver disease risk. BMC Public Health. 2024;24(1):2545.
- Njei B, Osta E, Njei N, Al-Ajlouni YA, Lim JK. An explainable machine learning model for prediction of high-risk nonalcoholic steatohepatitis. Sci Rep. 2024;14(1):8589.
- Fan R, Yu N, Li G, Arshad T, Liu WY, Wong GL, et al. Machine-learning model comprising five clinical indices and liver stiffness measurement can accurately identify MASLD-related liver fibrosis. Liver Int. 2024;44(3):749–59.
- Cheung JTK, Zhang X, Wong GL, Yip TC, Lin H, Li G, et al. MAFLD fibrosis score: using routine measures to identify advanced fibrosis in metabolic-associated fatty liver disease. Aliment Pharmacol Ther. 2023;58(11–12):1194–204.
- Bonfiglio C, Campanella A, Donghia R, Bianco A, Franco I, Curci R, et al. Development and Internal validation of a model for predicting overall survival in subjects with MAFLD: a cohort study. J Clin Med. 2024;13(4):1181.
- Drozdov I, Szubert B, Rowe IA, Kendall TJ, Fallowfield JA. Accurate prediction of all-cause mortality in patients with metabolic dysfunction-associated steatotic liver disease using electronic health records. Ann Hepatol. 2024;29(5): 101528.
- Carrillo-Larco RM, Guzman-Vilca WC, Castillo-Cara M, Alvizuri-Gómez C, Alqahtani S, Garcia-Larsen V. Phenotypes of non-alcoholic fatty liver disease (NAFLD) and all-cause mortality: unsupervised machine learning analysis of NHANES III. BMJ Open. 2022;12(11): e067203.
- Suárez M, Gil-Rojas S, Martínez-Blanco P, Torres AM, Ramón A, Blasco-Segura P, et al. Machine learning-based assessment of survival and risk factors in non-alcoholic fatty liver disease-related hepatocellular carcinoma for optimized patient management. Cancers (Basel). 2024;16(6):1114.
- Zhang Y, Wang F, Tang J, Shen L, He J, Chen Y. Association of triglyceride glucose-related parameters with all-cause mortality and cardiovascular

disease in NAFLD patients: NHANES 1999–2018. Cardiovasc Diabetol. 2024;23(1):262.

- Min Y, Wei X, Wei Z, Song G, Zhao X, Lei Y. Prognostic effect of triglyceride glucose-related parameters on all-cause and cardiovascular mortality in the United States adults with metabolic dysfunction-associated steatotic liver disease. Cardiovasc Diabetol. 2024;23(1):188.
- Decraecker M, Dutartre D, Hiriart JB, Irles-Depé M, Chermak F, Foucher J, de Lédinghen V. Long-term prognosis of patients with metabolic (dysfunction)-associated fatty liver disease by non-invasive methods. Aliment Pharmacol Ther. 2022;55(5):580–92.
- Liu W, Yang X, Zhan T, Huang M, Tian X, Tian X, Huang X. Weight-adjusted waist index is positively and linearly associated with all-cause and cardiovascular mortality in metabolic dysfunction-associated steatotic liver disease: findings from NHANES 1999–2018. Front Endocrinol (Lausanne). 2024;15:1457869.
- 33. do Nascimento CF, Dos Santos HG, de Moraes Batista AF, Roman Lay AA, Duarte YAO, Chiavegatto Filho ADP. Cause-specific mortality prediction in older residents of São Paulo, Brazil: a machine learning approach. Age Ageing. 2021;50(5):1692–8.
- Tran NTD, Balezeaux M, Granal M, Fouque D, Ducher M, Fauvel J-P. Prediction of all-cause mortality for chronic kidney disease patients using four models of machine learning. Nephrol Dial Transplant. 2022;38(7):1691–9.
- Tan HC, Zeng LJ, Yang SJ, Hou LS, Wu JH, Cai XH, et al. Deep learning model for the prediction of all-cause mortality among long term care people in China: a prospective cohort study. Sci Rep. 2024;14(1):14639.
- Anwar A, Rana S, Pathak P. Artificial intelligence in the management of metabolic disorders: a comprehensive review. J Endocrinol Invest. Published online February 19, 2025. https://doi.org/10.1007/ s40618-025-02548-x.
- Magga L, Maturana S, Olivares M, Valdevenito M, Cabezas J, Chapochnick J, et al. Identifying factors predicting kidney graft survival in Chile using elastic-net-regularized Cox's regression. Medicina (Kaunas). 2022;58(10):1348.
- Bortz J, Guariglia A, Klaric L, Tang D, Ward P, Geer M, et al. Biological age estimation using circulating blood biomarkers. Commun Biol. 2023;6(1):1089.
- Ning C, Ouyang H, Shen D, Sun Z, Liu B, Hong X, et al. Prediction of survival in patients with infected pancreatic necrosis: a prospective cohort study. Int J Surg. 2024;110(2):777–87.
- Rahman SA, Maynard N, Trudgill N, Crosby T, Park M, Wahedally H, et al. Prediction of long-term survival after gastrectomy using random survival forests. Br J Surg. 2021;108(11):1341–50.
- Yang X, Qiu H, Wang L, Wang X. Predicting colorectal cancer survival using time-to-event machine learning: retrospective cohort study. J Med Internet Res. 2023;25: e44417.
- Liu X, Xie Z, Zhang Y, Huang J, Kuang L, Li X, et al. Machine learning for predicting in-hospital mortality in elderly patients with heart failure combined with hypertension: a multicenter retrospective study. Cardiovasc Diabetol. 2024;23(1):407.
- Lin J, Gu C, Sun Z, Zhang S, Nie S. Machine learning-based model for predicting the occurrence and mortality of nonpulmonary sepsis-associated ARDS. Sci Rep. 2024;14(1):28240.
- 44. Duan S, Wu Y, Zhu J, Wang X, Zhang Y, Gu C, Fang Y. Development of interpretable machine learning models associated with environmental chemicals to predict all-cause and specific-cause mortality: a longitudinal study based on NHANES. Ecotoxicol Environ Saf. 2024;270: 115864.
- Shamsutdinova D, Stamate D, Stahl D. Balancing accuracy and Interpretability: an R package assessing complex relationships beyond the Cox model and applications to clinical prediction. Int J Med Inform. 2024;194: 105700.
- Fansler SD, Bakulski KM, Park SK, Walker E, Wang X. Use of biomarkers of metals to improve prediction performance of cardiovascular disease mortality. Environ Health. 2024;23(1):96.
- Golabi P, Paik JM, Arshad T, Younossi Y, Mishra A, Younossi ZM. Mortality of NAFLD according to the body composition and presence of metabolic abnormalities. Hepatol Commun. 2020;4(8):1136–48.
- Rasmussen DN, Thiele M, Johansen S, Kjærgaard M, Lindvig KP, Israelsen M, et al. Prognostic performance of 7 biomarkers compared to liver biopsy in early alcohol-related liver disease. J Hepatol. 2021;75(5):1017–25.

- Huang Y, Xu J, Yang Y, Wan T, Wang H, Li X. Association between lifestyle modification and all-cause, cardiovascular, and premature mortality in individuals with non-alcoholic fatty liver disease. Nutrients. 2024;16(13):2063.
- 50. Budd J, Cusi K. Nonalcoholic fatty liver disease: what does the primary care physician need to know? Am J Med. 2020;133(5):536–43.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.