MEDS Public Health and Preventive Medicine (2026)
Clausius Scientific Press, Canada

DOI: 10.23977/phpm.2026.060101
ISSN 2616-1915 Vol. 6 Num. 1

# *Constructing a Fatty Liver Prediction Model Based on Physical Examination Data and Machine Learning Algorithm*

**Tao Meiyu, Huang Jiehong, Qin Qingrong, Yao Shunhan, Li Zi, Yao Jianmin, Qi Yue, Zeng Lanqing, Wu Xiangjun***

*Health Management Center of Guangxi Medical University Kaiyuan Langdong Hospital, Nanning, 530000, Guangxi, China*
*895546347@qq.com*
**Corresponding author*

***Keywords:*** Fatty liver; Physical examination data; Machine learning; Logistic regression; Prediction model; Nomogram

***Abstract***: Fatty liver disease is one of the most common chronic liver diseases and is closely associated with metabolic disorders. Early identification of high-risk populations is essential for prevention and intervention. Based on large-scale physical examination data, this study aimed to identify predictors of fatty liver disease and construct a prediction model using machine learning algorithms. A total of 4,700 individuals who underwent physical examinations at the Health Management Center of Guangxi Medical University Kaiyuan Langdong Hospital from 2020 to 2023 were retrospectively included and divided into a fatty liver group (1,372 cases) and a normal group (3,328 cases) according to abdominal ultrasound results. Demographic characteristics and biochemical indicators, including age, sex, body mass index (BMI), alanine aminotransferase (ALT), uric acid (UA), fasting plasma glucose (FPG), total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C), were collected. Univariate analysis and multivariate logistic regression were used to identify independent risk factors for fatty liver disease, and five machine learning models, including logistic regression, support vector machine (SVM), random forest, k-nearest neighbors (KNN), and neural network, were constructed to predict fatty liver risk. The predictive performance of each model was evaluated using accuracy, sensitivity, specificity, F1 score, and area under the receiver operating characteristic curve (AUC). The results showed that age, sex, BMI, ALT, UA, FPG, and TG were independent risk factors for fatty liver disease. Among the five models, logistic regression achieved the best performance with an accuracy of 0.7675 and an AUC of 0.7975. A nomogram based on the logistic regression model was further established and showed good calibration and predictive consistency. These findings suggest that a prediction model based on routine physical examination indicators can effectively assess the risk of fatty liver disease and may provide a convenient tool for early screening and risk assessment in health examination populations.

## 1. Introduction

Fatty liver disease refers to a class of chronic liver diseases characterized by fat accumulation in the liver. It not only affects liver function but is also closely associated with various metabolic diseases, such as obesity, diabetes, and cardiovascular disease [1-3]. In recent years, with the Westernization of lifestyles, changes in dietary structure, and the high prevalence of metabolic-related diseases such as obesity and diabetes, fatty liver disease has become the most common chronic liver disease in China and globally [4, 5]. Although fatty liver disease is often asymptomatic in the early stages, it can progress to steatohepatitis, liver fibrosis, cirrhosis, and even hepatocellular carcinoma, posing a serious threat to public health [6]. Therefore, how to identify high-risk populations at an early stage and implement precise preventive interventions is a major challenge facing current public health and clinical management. However, traditional diagnostic methods for fatty liver disease primarily rely on imaging modalities, such as ultrasound and computed tomography (CT). These methods are limited by high costs, dependence on specialized equipment and personnel, and are not feasible for large-scale early screening [7]. Therefore, developing a risk prediction tool based on readily available routine examination indicators holds significant practical importance.

In recent years, machine learning technology has demonstrated significant advantages in the construction of medical prediction models. Compared to traditional statistical methods, machine learning technology can extract valuable information from massive datasets and construct highly efficient predictive models, which have been widely applied in multiple fields such as disease screening and individualized risk assessment [8]. Currently, some studies have attempted to apply machine learning models to the identification of fatty liver disease. However, most of these studies have small sample sizes, fail to compare the differences in predictive performance among different algorithms, or involve complex clinical application procedures, which have limited the promotion and application of these models in actual physical examination populations [9, 10]. Based on large-scale physical examination data, this study screens out predictive factors related to fatty liver disease, applies multiple machine learning algorithms to construct a fatty liver prediction model, aiming to provide a new method and basis for the early identification of high-risk populations.

## 2. Materials and Methods

### 2.1. Study Population

This study was a retrospective cross-sectional study, enrolling subjects who underwent routine physical examinations at the Health Management Center of Guangxi Medical University's Kaixuan Langdong Hospital between January 2020 and December 2023. Inclusion criteria were: (1) age ≥ 18 years; (2) complete physical examination data, including abdominal color Doppler ultrasound, height and weight measurements, and blood biochemical tests. Exclusion criteria were: (1) history of clear liver disease such as viral hepatitis, drug-induced liver injury, or hereditary metabolic liver disease; (2) presence of severe systemic diseases or tumors; (3) subjects with missing data affecting the analysis. A total of 4,700 subjects were finally included, of which 1,372 were in the fatty liver group and 3,328 were in the normal group.

### 2.2. Data Collection

General information and laboratory test indicators of the subjects were extracted through the Hospital Information System (HIS), including general data such as age, sex, height, weight, and body mass index (BMI), as well as biochemical indicators including alanine aminotransferase

(ALT), uric acid (UA), fasting plasma glucose (FPG), total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C).

### 2.3. Definition of Fatty Liver

The diagnosis of fatty liver was based on judgments made by experienced ultrasound physicians on the day of the physical examination, according to abdominal B-ultrasound imaging results. Diagnostic criteria included imaging manifestations such as enhanced hepatic echogenicity, decreased liver-kidney contrast, blurred liver boundaries, and enhanced far-field attenuation. Individuals meeting the diagnostic criteria established by the "Guidelines for the Prevention and Treatment of Fatty Liver" of the Hepatology Branch of the Chinese Medical Association [11] were diagnosed with fatty liver.

### 2.4. Screening of Independent Risk Factors

First, univariate analysis was employed to compare demographic and biochemical indicators between the normal group and the fatty liver group.For normally distributed continuous variables, the independent samples t-test was used; for non-normally distributed variables, the Mann-Whitney U test was applied; and for categorical variables, the chi^2 test was utilized. Subsequently, variables with statistical significance ($P<0.05$) were incorporated into a multivariate Logistic regression model. Stepwise regression was adopted to screen for independent influencing factors of fatty liver, and the regression coefficients (β), odds ratios (OR), and 95% confidence intervals (CI) were calculated.

### 2.5. Construction of Machine Learning Models

Using the presence or absence of fatty liver as the dependent variable and incorporating the screened key variables, five common machine learning models were constructed using Python (version 3.10): Logistic Regression, Support Vector Machine (SVM), Random Forest, K-Nearest Neighbors (KNN), and Neural Network (MLP). All data were randomly divided into training and testing sets at a 7:3 ratio. All continuous variables were standardized using Z-score normalization prior to model construction. By comparing the predictive performance of each model on the testing set，including Accuracy, Sensitivity, Specificity, F1 Score, and Area Under the ROC Curve (AUC), the effectiveness of different models was evaluated.

### 2.6. Nomogram Construction and Validation

Based on the results of the Logistic regression model, a nomogram was plotted using the "rms" package in R to achieve a visualized prediction of individual fatty liver risk. Simultaneously, calibration curves for the training set and the test set were plotted to assess the consistency between the model's predicted probabilities and the actual incidence rates. A bias-corrected calibration curve was calculated by performing 1000 Bootstrap resamplings to ensure model stability.

### 2.7. Statistical Analysis

Data analysis was performed using SPSS 26.0 and Python 3.10. Normally distributed continuous data are expressed as mean ± standard deviation (bar{x} pm s), and intergroup comparisons were conducted using the independent samples t-test. Non-normally distributed data are presented as median (interquartile range), and intergroup comparisons were performed using the Mann-Whitney

U test. Categorical data are expressed as number (percentage), and intergroup comparisons were carried out using the chi^2 test. A P-value of <0.05 was considered statistically significant.

## 3. Results

### 3.1. Basic Data and Univariate Analysis

A total of 4,700 subjects were enrolled, including 1,372 (29.2%) in the fatty liver group and 3,328 (70.8%) in the normal group. There were significant differences between the two groups in terms of age, sex, body mass index (BMI), alanine aminotransferase (ALT), uric acid (UA), fasting plasma glucose (FPG), total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C). The fatty liver group had a significantly higher mean age, a higher proportion of males, and significantly higher levels of BMI, ALT, UA, FPG, TC, TG, and LDL-C compared to the normal group. Conversely, the level of HDL-C was significantly lower in the fatty liver group. These differences were statistically significant (P<0.05). (See Table 1 for details.)

Table 1: Baseline Characteristics and Univariate Analysis of the Two Groups

| Variable | Normal group | Fatty Liver Group | Statistics | P Value |
|---|---|---|---|---|
| Age years | 36.37 ±10.78 | 40.12 ±11.11 | 10.610 | <0.001 |
| Sex,n(%) | 812/2516 | 526/846 | 92.011 | <0.001 |
| BMI | 22.21 ±3.15 | 25.53 ±3.85 | 146.329 | <0.001 |
| ALT | 19.09 ±19.53 | 29.65 ±24.50 | 27.461 | <0.001 |
| UA | 318.92 ±86.43 | 378.53 ±96.63 | 104.935 | <0.001 |
| FPG | 4.96 ±0.82 | 5.39 ±1.38 | 53.614 | <0.001 |
| TC | 4.86 ±0.93 | 5.19 ±1.08 | 54.014 | <0.001 |
| TG | 1.23 ±1.35 | 2.96 ±3.25 | 79.421 | <0.001 |
| HDL-C | 1.52 ±0.37 | 1.30 ±0.33 | 105.635 | <0.001 |
| LDL-C | 2.88 ±0.79 | 3.26 ±0.88 | 70.908 | <0.001 |

### 3.2. Multivariate Logistic Regression Analysis

Variables with statistical significance were included in the multivariate Logistic regression model. The results showed that sex (OR=0.481, 95%CI: 0.393–0.587), BMI (OR=1.197, 95%CI: 1.167–1.227), ALT (OR=1.032, 95%CI: 1.023–1.040), UA (OR=1.003, 95%CI: 1.002–1.004), FPG (OR=1.152, 95%CI: 1.069–1.242), and TG (OR=1.203, 95%CI: 1.097–1.320) were independent risk factors for fatty liver. Specific results are shown in Table 2.

Table 2: Multivariate Logistic Regression Analysis

| | Beta Coefficient | Standard Error | Z | OR | 95%CI | P Value |
|---|---|---|---|---|---|---|
| Age years | 0.016 | 0.004 | 4.469 | 1.016 | 1.009–1.023 | <0.001 |
| Sex,n(%) | 0.733 | 0.102 | 7.166 | 0.481 | 0.393–0.587 | <0.001 |
| BMI | 0.179 | 0.013 | 13.86 | 1.197 | 1.167–1.227 | <0.001 |
| ALT | 0.031 | 0.004 | 7.616 | 1.032 | 1.023–1.04 | <0.001 |
| UA | 0.003 | 0.001 | 6.057 | 1.003 | 1.002–1.004 | <0.001 |
| FPG | 0.142 | 0.038 | 3.703 | 1.152 | 1.069–1.242 | <0.001 |
| TC | -0.195 | 0.174 | -1.12 | 0.823 | 0.585–1.157 | 0.263 |
| TG | 0.185 | 0.047 | 3.916 | 1.203 | 1.097–1.32 | <0.001 |
| HDL-C | -0.399 | 0.219 | -1.826 | 0.671 | 0.437–1.03 | 0.068 |
| LDL-C | 0.327 | 0.192 | 1.703 | 1.386 | 0.952–2.019 | 0.088 |

### 3.3. Comparison of Machine Learning Prediction Models and Evaluation Metrics

Using the variables screened by Logistic regression as input features, five machine learning models were constructed: Logistic regression, SVM, Random Forest, K-Nearest Neighbors, and Neural Network. All models demonstrated good predictive performance on the test set, with accuracies above 0.75 and AUCs above 0.74. Among them, the Logistic regression model performed the best, with an AUC of 0.7975, an F1 score of 0.5432, and an accuracy of 0.7675. SVM and Neural Network followed. KNN and Random Forest were slightly superior in sensitivity, but their overall performance was slightly inferior to Logistic regression. (See Table 3 for details.)

Table 3: Evaluation Metric Results of Various Machine Learning Prediction Models on the Test Set

| Machine Learning Model | Accuracy | Sensitivity | Specificity | F1 Score | AUC |
|---|---|---|---|---|---|
| Neural Network | 0.7668 | 0.4029 | 0.9169 | 0.5023 | 0.7877 |
| SVM | 0.7675 | 0.4369 | 0.9039 | 0.5233 | 0.7845 |
| Random Forest | 0.7548 | 0.4903 | 0.8639 | 0.5387 | 0.7675 |
| KNN | 0.7519 | 0.4976 | 0.8569 | 0.5395 | 0.7448 |
| Logistic Regression | 0.7675 | 0.4733 | 0.8889 | 0.5432 | 0.7975 |

### 3.4. Construction of the Nomogram

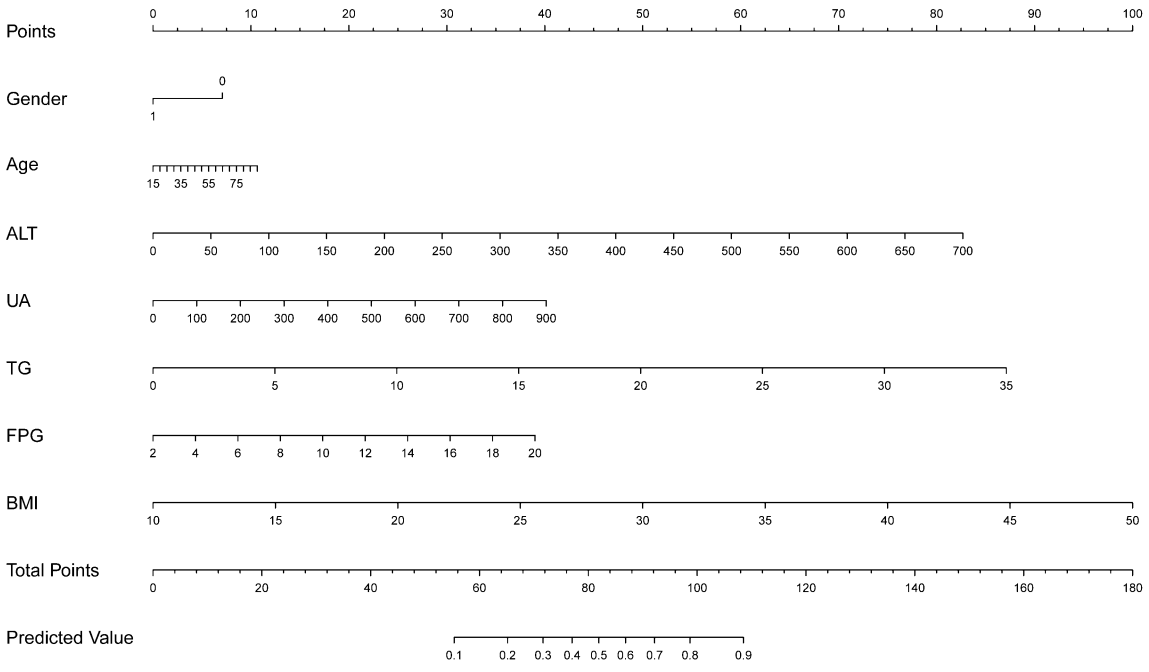


Figure 1: A nomogram for predicting fatty liver.

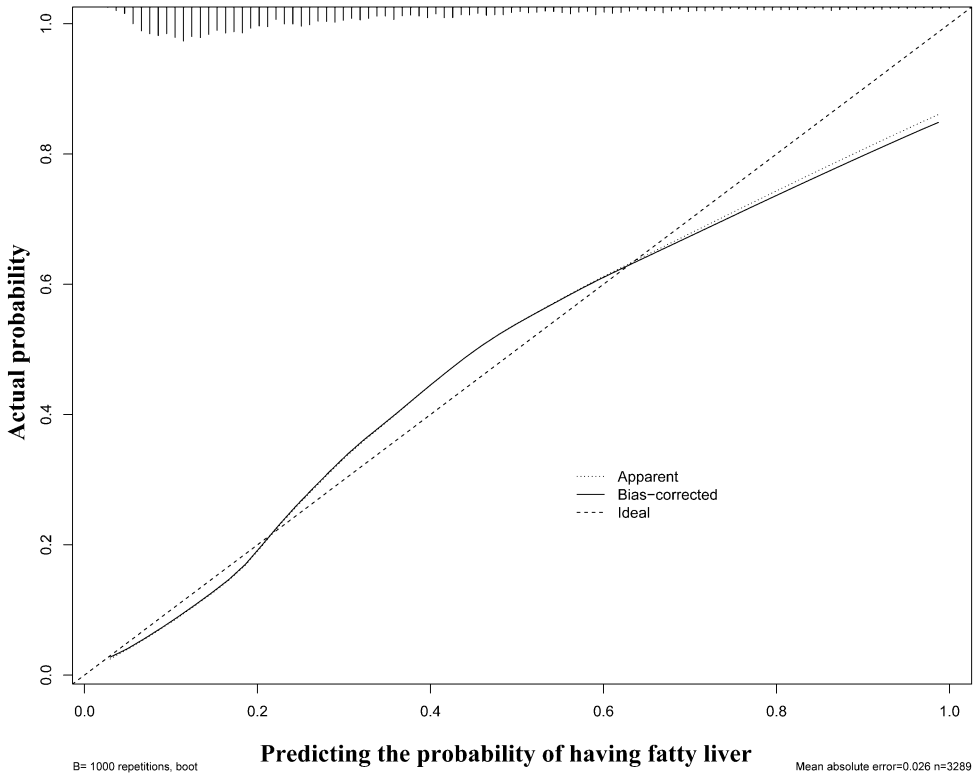


Figure 2: Calibration curve of nomogram on training set.

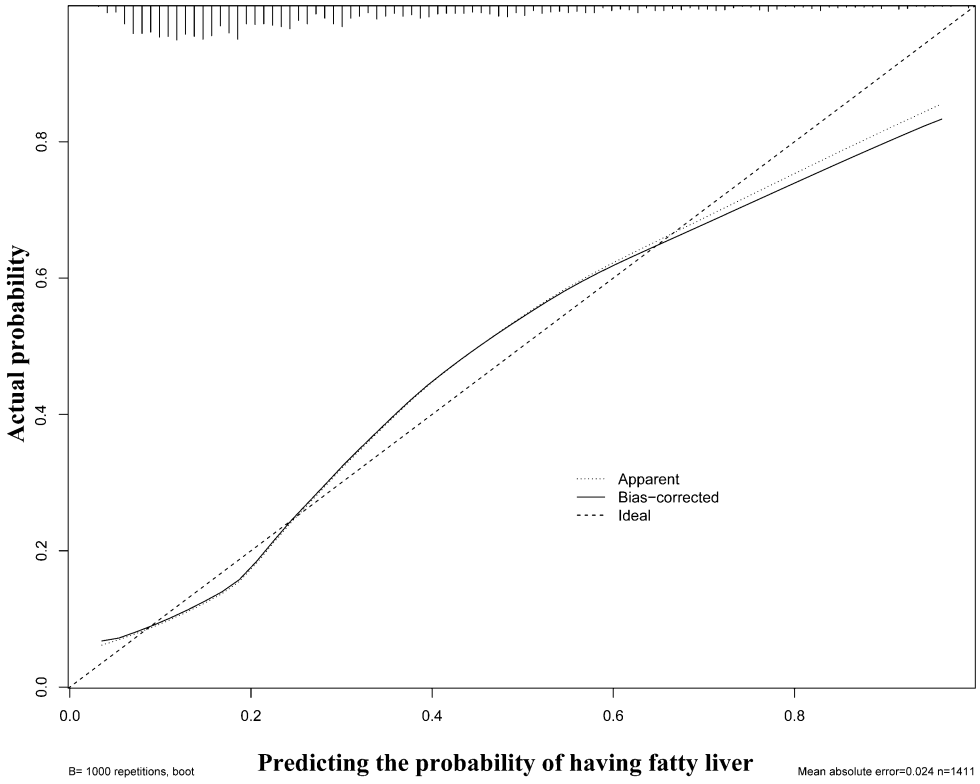


Figure 3: Calibration curve of nomogram on validation set.

A nomogram for predicting fatty liver was developed based on a multivariate logistic regression model incorporating seven variables: sex, age, BMI, ALT, UA, FPG, and TG, enabling a visual assessment of individual fatty liver risk (Figure 1). In the nomogram, each variable corresponds to a specific score, and the total score obtained by summing the individual scores can be used to estimate the probability of developing fatty liver. The calibration analysis demonstrated that the nomogram showed good agreement between predicted and observed outcomes in both the training

and test sets, indicating satisfactory model stability and predictive accuracy. Figure 2 presents the calibration curve of the nomogram in the training set, while Figure 3 illustrates the calibration curve in the test set.

## 4. Discussion

With the high prevalence of metabolism-related diseases, fatty liver disease has gradually become one of the most common chronic liver diseases in China [4, 5]. In its early stages, it typically lacks typical symptoms, and patients are often incidentally discovered during physical examinations. Fatty liver can progress to a series of severe diseases; therefore, early identification and intervention are particularly important [12, 13]. Based on data from a large-scale physical examination population, this study systematically evaluated the association between various routine clinical indicators and fatty liver, and introduced multiple machine learning algorithms to construct predictive models, aiming to provide a convenient and efficient risk assessment tool for the early screening of fatty liver.

In this study, univariate analysis results showed that the fatty liver group differed significantly from the normal group in multiple indicators including age, sex, BMI, ALT, UA, FPG, and TG, suggesting that metabolic abnormalities are an important basis for fatty liver. Further multivariate Logistic regression analysis indicated that BMI, ALT, UA, FPG, and TG are independent risk factors for fatty liver, which is basically consistent with previous research results. Ouyang Die et al. analyzed 992 examinees who participated in health check-ups at a Class A tertiary general hospital in Shanghai and found that sex, age, and body mass index are independent risk factors for fatty liver [14]. The research by Feng Junfang, Li Yunyun, and others also pointed out that ALT, blood glucose, and serum uric acid are risk factors for the occurrence of fatty liver [15, 16].

BMI is an important indicator reflecting the degree of individual obesity. In this study, the OR value of BMI was 1.197, suggesting it is one of the most significant predictive factors for fatty liver. Previous studies have confirmed that obesity is one of the main etiological factors for fatty liver; excessive accumulation of adipose tissue promotes insulin resistance, lipid metabolism disorders, and hepatic fat deposition [17,18]. ALT is a sensitive indicator of hepatocyte injury; elevated levels often indicate the progression of fatty liver disease to the stage of hepatitis [19]. Elevated UA reflects abnormal purine metabolism and oxidative stress, which may exacerbate hepatic steatosis by promoting fat generation and inflammatory responses [20]. FPG and TG reflect the status of glucose and lipid metabolism, respectively, and are both important components of metabolic syndrome. Their elevation suggests the presence of potential insulin resistance in patients with fatty liver [21, 22].

In recent years, the application of machine learning in the construction of medical prediction models has become increasingly widespread. Compared to traditional regression models, machine learning can more fully explore the nonlinear relationships and interaction effects between variables, thereby improving prediction accuracy [23]. In this study, five models were constructed: Logistic regression, SVM, Random Forest, KNN, and Neural Network, and their performance was evaluated on a test set. The results showed that all models possessed good predictive capabilities, with Logistic regression performing the best in terms of AUC (0.7975) and F1 score (0.5432), demonstrating good sensitivity and specificity. SVM and Neural Network followed, while KNN and Random Forest had a slight advantage in sensitivity but were slightly inferior in accuracy and AUC. As a traditional classification method widely used, Logistic regression has advantages such as strong interpretability and simple implementation [24]. Its optimal performance in this study indicates that the linear relationship between variables is relatively clear, also suggesting the important value of high-quality physical examination data in model building. Although deep learning models such as neural networks have stronger fitting capabilities, they have higher

requirements for data scale and structure, and their interpretability is poor, which limits their promotion in clinical practice. Dai Xiaoxia et al. also constructed a fatty liver prediction model based on machine learning, with model performance similar to this study, but they did not construct a nomogram, making it inapplicable in clinical settings [25].

This study constructed a nomogram for predicting fatty liver based on a Logistic regression model, integrating 7 easily obtainable physical examination indicators, enabling rapid quantitative assessment of individual fatty liver risk. Nomograms are characterized by strong intuitiveness and simple operation, and have been widely adopted in the risk assessment of various diseases [26, 27]. The nomogram constructed in this study demonstrated good calibration curves in both the training and validation sets, indicating a high degree of consistency between the predicted probabilities and the actual incidence rates, thus possessing clinical practical value. Combined with electronic health records, this tool can be rapidly implemented at physical examination sites to facilitate early identification and health interventions for high-risk populations of fatty liver, demonstrating broad potential for promotion.

This study has the following advantages: the data were derived from a real physical examination population, with a large sample size and strong representativeness; the included indicators were all routine physical examination items, facilitating simple data acquisition and model promotion; the combination with a nomogram improved the model's interpretability and clinical applicability. However, this study also has certain limitations: it was a single-center retrospective study with a strong regional sample bias, requiring further validation with multi-center data; some potential influencing factors (such as diet, exercise, genetics, etc.) were not included in the model, which may have caused certain interference to the prediction results; the diagnosis of fatty liver relied on B-ultrasound, which, although a commonly used method in physical examinations, carries a certain risk of missed diagnosis for mild fatty liver.

In summary, this study systematically evaluated the risk factors associated with fatty liver based on large-sample physical examination data. Prediction models were constructed using multivariate Logistic regression and various machine learning algorithms, and a nomogram was established to enable efficient identification of individual fatty liver risk.

## Acknowledgements

Self-raised Fund Scientific Research Project of the Guangxi Zhuang Autonomous Region Health Commission (No. Z-A20231117).

## References

*[1] Qu, J.C., Qin, J.T., Wang, A.P. and Others. (2024) Study on the correlation between fatty liver index and risk of incident diabetes in different glucose metabolism populations. Chinese Journal of Diabetes Mellitus, 32, 726-730.*

*[2] Tu, Y., Yang, C.J., Huang, Y. and Others. (2022) A study on the correlation between non-alcoholic fatty liver and cardiovascular diseases. Natural Magazine, 44, 149-154.*

*[3] Yip, T.C., Fan, J.G. and Wong, V.W. (2023) China's fatty liver crisis: A looming public health emergency. Gastroenterology, 165, 825-827.*

*[4] Eslam, M., Sarin, S.K., Wong, V.W. and Others. (2020) The Asian Pacific Association for the Study of the Liver clinical practice guidelines for the diagnosis and management of metabolic associated fatty liver disease. Hepatology International, 14, 889-919.*

*[5] Xu, H., Fang, D., Zhou, W.H. and Others. (2025) A retrospective cohort study on the association between Chinese visceral adiposity index and the risk of fatty liver. Chinese General Practice, 28, 1336-1341, 1366.*

*[6] Zhao, Q.W., Peng, D.L., Liu, S.S. and Others. (2024) An age-period-cohort effect analysis of liver cancer mortality burden caused by non-alcoholic fatty liver disease in China from 1990 to 2019. Journal of Hepatopancreaticobiliary Surgery, 36, 667-672.*

*[7] Mu, L. (2024) The clinical significance of ultrasound examination in the diagnosis of fatty liver. China Medical*

*Device Information, 30, 108-110.*
*[8] Zhong, J.J., Li, W.T., Huang, Y.F. and Others. (2024) Design characteristics and methodological quality of machine learning prediction model studies in primary healthcare: A scoping review. Chinese General Practice, 27, 1271-1276.*
*[9] Cai, N.X., Ma, Y.N. and Wen, D.L. (2024) Application of machine learning in the prediction of non-alcoholic fatty liver. China Health Statistics, 41, 316-318.*
*[10] Chang, Q.X., Wang, X.M., Wang, C. and Others. (2022) A study on risk factors related to fatty liver in the elderly based on association rules. China Health Statistics, 39, 558-561.*
*[11] Chen, Z.B. (2015) Publication of the book "Chinese Guidelines for the Prevention and Treatment of Fatty Liver (Popular Science Edition)". Practical Hepatology, 18, 248.*
*[12] Marengo, A., Rosso, C. and Bugianesi, E. (2016) Liver cancer: Connections with obesity, fatty liver, and cirrhosis. Annual Review of Medicine, 67, 103-117.*
*[13] Polyzos, S.A., Chrysavgis, L., Vachliotis, I.D., Chartampilas, E. and Cholongitas, E. (2023) Nonalcoholic fatty liver disease and hepatocellular carcinoma: Insights in epidemiology, pathogenesis, imaging, prevention and therapy. Seminars in Cancer Biology, 93, 20-35.*
*[14] Ouyang, D., Xu, D.D., Sun, Y.R. and Others. (2024) Construction of a non-alcoholic fatty liver disease prediction model for non-obese populations based on lifestyle-related indicators. Journal of Nursing, 31, 1-7.*
*[15] Feng, J.F., Feng, J.Z., Liu, S.H. and Others. (2024) Correlation between serum uric acid/creatinine ratio and non-alcoholic fatty liver disease in physical examination populations. Chinese Health Management, 18, 58-61.*
*[16] Li, W.W. and Li, W. (2024) Value analysis of combined detection of serum inflammatory factors, blood glucose, and uric acid in the diagnosis of fatty liver in the elderly. Journal of Experimental and Laboratory Medicine, 42, 17-19.*
*[17] Feng, Y.P., Zhang, Y.X. and Chen, D. (2024) Correlation analysis of triglyceride-glucose index and its combination with body mass index with non-alcoholic fatty liver in medical staff. Chinese General Practitioner, 23, 1168-1173.*
*[18] Lei, T., Liu, Y., Bu, H.Y., et al. (2023) Correlation analysis between body mass index and fatty liver in a health examination population. Xinjiang Medical Journal, 53, 1344–1346, 1354.*
*[19] Gu, X.D., Zhang, X.Y. and Sun, X.H. (2017) Analysis of risk factors for elevated ALT in patients with non-alcoholic fatty liver. Modern Journal of Laboratory Medicine, 32, 140-143.*
*[20] Fan, J. and Wang, D. (2024) Serum uric acid and nonalcoholic fatty liver disease. Frontiers in Endocrinology, 15, 1455132.*
*[21] Arefhosseini, S., Aghajani, T., Tutunchi, H. and Ebrahimi-Mameghani, M. (2024) Association of systemic inflammatory indices with anthropometric measures, metabolic factors, and liver function in non-alcoholic fatty liver disease. Scientific Reports, 14, 12829.*
*[22] Kyhl, L.K., Nordestgaard, B.G., Tybjærg-Hansen, A., Smith, G.D. and Nielsen, S.F. (2025) VLDL triglycerides and cholesterol in non-alcoholic fatty liver disease and myocardial infarction. Atherosclerosis, 401, 119094.*
*[23] Fu, S.T., He, B. and Xu, J.C. (2025) Establishment and validation of a machine learning algorithm-based prediction model for liver cirrhosis-related hepatic encephalopathy. Chinese Journal of Laboratory Medicine, 48, 93-102.*
*[24] Chu, J., Yuan, Z.R. and Mu, X. (2024) Logistic regression analysis of risk factors for malnutrition in patients with hepatocellular carcinoma undergoing hepatic artery chemoembolization and construction of a nomogram prediction model. Chinese Medical Journal, 59, 1354-1358.*
*[25] Dai, X., Li, J., Wang, C.Y. and Others. (2022) Construction of a machine learning-based prediction model for fatty liver. Journal of Hubei University of Medicine, 41, 574-577.*
*[26] Gu, L., Chen, X., Zhang, Y. and Others. (2025) Analysis of risk factors for hospital infection in patients with stroke and construction of a nomogram prediction model. Journal of Jiangsu University (Medical Science), 35, 56-61.*
*[27] Yu, L., Wang, X., Zhang, X. and Others. (2025) Construction of a nomogram model for predicting the onset of hepatocellular carcinoma based on abnormal serum prothrombin and alpha-fetoprotein, and model evaluation. Chinese Journal of Hepatobiliary Surgery, 31, 1-5.*