British Journal of Nutrition, page 1 of 11 doi:10.1017/S0007114524001442

Association between nutritional status indices and non-alcoholic fatty liver disease in older adults: insights from the National Health and Nutrition Examination Survey 2017–2018

Haisheng Chai†, Sicheng Gao†, Yaoyao Dai†, Jinhua Dai, Gang Zhao* and Junfeng Zhu*

Department of Hepatology, Yueyang Integrated Chinese and Western Medicine Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, People's Republic of China

(Submitted 19 March 2024 – Final revision received 28 May 2024 – Accepted 17 June 2024)

Abstract

While previous studies have identified a relationship between dietary intake and the risk of non-alcoholic fatty liver disease (NAFLD), the influence of overall nutritional status on NAFLD development has not been thoroughly investigated. This study sought to explore the association between different nutritional status indicators and NAFLD among the older adults. Nutritional status was evaluated using controlling nutritional status (CONUT), prognostic nutritional index (PNI) and nutritional risk index (GNRI), while NAFLD was identified based on a controlled attenuation parameter ≥ 285 dB/m, measured using transient elastography. The analysis included multivariate regression, receiver operating characteristic analysis, eXtreme Gradient Boosting and subgroup analysis to investigate the relationships between nutritional status indices and NAFLD. The study enrolled 1409 participants for the main analysis, with an NAFLD prevalence of 44·7 %. After accounting for potential confounders, a positive association between PNI and NAFLD was observed. Participants in the third and fourth quartiles of PNI showed increased odds of NAFLD compared with the lowest quartile (Q3: OR = 1·45, 95 % CI (1·03, 2·05); Q4: OR = 2·27, 95 % CI (1·59, 3·24)). Similarly, higher GNRI quartiles were significantly associated with greater odds of NAFLD (Q4 *v.* Q1: aOR = 1·84; 95 % CI (1·28, 2·65)). Conversely, higher CONUT values were linked to a reduced prevalence of NAFLD (OR = 0·65, 95 % CI (0·48, 0·87)). This study highlights that suboptimal nutritional status, indicating overnutrition as evaluated by PNI, GNRI and CONUT, is positively linked with the risk of NAFLD in individuals aged 50 years and above.

Keywords: **Nutritional status: Prognostic nutritional index: Controlling nutritional status: Geriatric nutritional risk index: Non-alcoholic fatty liver disease: National Health and Nutrition Examination Survey**

Non-alcoholic fatty liver disease (NAFLD), the most common liver disorder linked to obesity, is projected to become the leading cause of liver transplants in the USA in the near future[1]. Current estimates suggest that NAFLD affects between 34·1 % and 56·7 % of the American adult population[2]. The condition is characterised by abnormal accumulation of hepatic lipids, hepatocellular injury, necroinflammation and fibrosis, which may rapidly progress to cirrhosis and its subsequent complications[3]. Particularly, patients with advanced liver fibrosis due to NAFLD face a heightened risk of severe outcomes, including hepatocellular carcinoma, chronic liver failure and CVD[4]. Identifying and modifying risk factors for NAFLD is crucial for preventing or mitigating its progression, especially in the absence of effective pharmacological treatments[5].

Previous studies indicate that NAFLD prevention or treatment through intervention programmes might be influenced by behavioural factors, such as insufficient physical activity, low fitness levels (e.g. reduced muscle mass and grip strength), obesity and poor dietary habits[6]. Increasingly, nutritional status is being recognised as a constellation of modifiable risk factors[7]. Malnutrition, which can result from both undernutrition and overnutrition, is a widespread concern with significant negative effects on clinical and physical health[8] and is particularly prevalent among the older adults[9]. Epidemiological studies consistently link higher fat mass (FM) and lean body mass (LBM) with NAFLD, common indicator of nutritional status, with increased NAFLD prevalence in general populations, who are also at an elevated risk of disease progression[10].

Abbreviations: CONUT, controlling nutritional status; GNRI, geriatric nutritional risk index; NAFLD, non-alcoholic fatty liver disease; NHANES, National Health and Nutrition Examination Survey; PIR, poverty:income ratio; PNI, prognostic nutritional index; SHAP, Shapley Additive exPlanations; XGBoost, eXtreme gradient boosting.

* **Corresponding authors:** Gang Zhao, email zhaogangsh@vip.163.com; Junfeng Zhu, email zhujunfeng@shutcm.edu.cn

† These authors have contributed equally to this work.

Nonetheless, the connection between poor nutritional status and the risk of NAFLD has not been thoroughly investigated[11]. Recently, there has been an emergence of indices such as the controlling nutritional status (CONUT), geriatric nutritional risk index (GNRI) and prognostic nutritional index (PNI), which have been validated to assess the nutritional status of hospitalised patients and tumour patients. These indices encompass aspects of immunonutrition, including immunological and nutritional status, as well as chronic inflammation[12]. These indicators have been demonstrated to act as prognostic markers in a range of conditions, including cancer, autoimmune diseases and heart disease[13]. Given that NAFLD is associated with low-grade inflammation, malnutrition and immunological dysfunction[14], the potential links between these factors and NAFLD prevalence warrant investigation.

Given that dietary patterns adequately reflect the synergistic effects of various food groups or nutrients on health outcomes, recent research and systematic reviews suggest that dietary patterns, particularly Western diets, are associated with increased risks of chronic liver diseases such as NAFLD, cirrhosis and liver cancer[15], while Mediterranean and prudent dietary patterns are linked to reduced risks of cirrhosis[15]. Therefore, examining comprehensive dietary patterns and nutritional status could provide further understanding of the relationship between nutrition and NAFLD[16]. In this study, our goal is to clarify the connection between nutritional status indicators and the prevalence of NAFLD in the USA population. We also aim to assess the predictive value of these indicators for the disease, representing a novel approach in this area of research.

Materials and methods

Study samples

NHANES is a cross-sectional survey conducted nationwide by the Centers for Disease Control and Prevention to evaluate the health and nutritional status of non-institutionalised individuals in the USA. Using a sophisticated, multistage probability sampling technique, NHANES deliberately oversamples specific demographic groups that may be more vulnerable to health problems. This survey enrolls approximately 5000 participants each year. Although data collection is yearly, findings are disseminated biennially. Participants in NHANES complete an extensive questionnaire and provide urine and blood samples during a medical examination in a mobile clinic. This study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving human subjects/patients were approved by the Ethical Review Board of the National Center for Health Statistics. Written informed consent was obtained from all subjects[17].

This study utilised data from the 2017–2018 NHANES cycle and focused on adults aged 50 years or older who had undergone vibration-controlled transient elastography, a measure of liver stiffness indicating advanced liver fibrosis. Vibration-controlled transient elastography data were available exclusively in the 2017–2018 NHANES cycle. Exclusion criteria included pregnant women, individuals with a history of excessive alcohol consumption (≥ 10 g or 1 drink/day for females, ≥ 20 g or 2 drinks/day for males), those with implausible energy intake (< 600 or > 3500 kcal/d for women, < 800 or > 4200 kcal/d for men)[17], participants with incomplete dietary intake data or only one dietary recall and those lacking data on relevant demographic and clinical variables. The final sample comprised 1409 older adults who met all inclusion criteria and provided complete data for analysis. A flow chart detailing the study's methodology is available in Fig. 1.

Measurements and definition of non-alcoholic fatty liver disease

During the 2017–2018 NHANES cycle, participants aged 12 years and older underwent transient elastography exams. Detailed methodology of the NHANES transient elastography protocol is documented elsewhere[18]. Liver stiffness and controlled attenuation parameter scores were assessed using the FibroScan® model 502 V2 Touch (Echosens). A valid assessment necessitated at least ten reliable stiffness measurements, a fasting period of at least 3 h and a liver stiffness interquartile range/median of ≤ 30 %. Hepatic steatosis was diagnosed based on a controlled attenuation parameter threshold of 285 dB/m, which demonstrated optimal diagnostic performance with a specificity of 77 % and a sensitivity of 80 %[19].

Nutritional status evaluations

This research utilised three distinct indices to assess the nutritional status of the participants based on based on blood sample testing data: GNRI, PNI and CONUT, as referenced in previous studies[12]. The detailed formulas for each of these indices can be found in online Supplementary Table S1. Both the GNRI and PNI indices were divided into quartiles. The first quartile represented the lowest score, indicating a lower risk of overnutrition, while the fourth quartile, with the highest score, suggested an increased risk of overnutrition. In contrast, the CONUT index was categorised using a clinically established cut-off from prior research[20], dividing participants into two groups: a low CONUT group (< 2 scores), indicating a higher risk of overnutrition, and a high CONUT group (≥ 2 scores), denoting a lower risk.

Covariates

The selection of covariates for this study was guided by existing literature[15] and a pre-established directed acyclic graph (online Supplementary Fig. S1). The demographic and clinical variables included were age, sex (male, female), racial/ethnic background (Mexican American, Other Hispanic, Non-Hispanic White, Non-Hispanic Black and Other Race), educational attainment (less than high school, high school or equivalent, more than college) and BMI (categorised as normal weight, overweight, and obese). Other factors considered were the poverty:income ratio (PIR) (< 1·0, 1·0–3·0, > 3·0), smoking status (non-smoker, former smoker, current smoker), physical activity (yes, no), hypertension (yes, no), diabetes (yes, no), high cholesterol (yes, no) and total energy intake. Smoking status was classified into never smokers (less than 100 cigarettes lifetime), current smokers (more than 100 cigarettes lifetime) and former smokers (more than 100 cigarettes

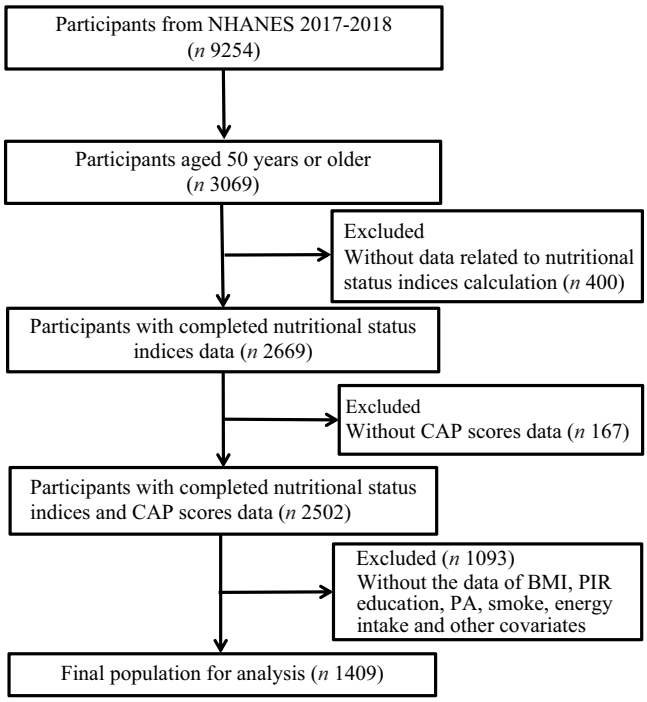

Fig. 1. Flow chart illustrating selection of the study population in NHANES from 2017 to 2018. NHANES, National Health and Nutrition Examination Survey.

lifetime and no longer smoking). Physical activity was assessed based on engagement in vigorous or moderate recreational activities[17]. Total energy intake was calculated as the average from two dietary recall interviews. Participants' food intake data were collected through two 24-hour dietary recall interviews, which were conducted in person 3–10 d later and by telephone. Dietary nutrition and energy intake were estimated using the Food and Nutrition Database for Dietary Studies. Interviews were conducted by trained interviewers at a mobile testing center. Interviewers used the USA Department of Agriculture automated multichannel approach to obtain more accurate and complete data for the NHANES study[21]. BMI was categorised into $< 25{\cdot}0$ kg/m^2 (normal weight), 25·0–30·0 kg/m^2 (overweight) and $\geq 30{\cdot}0$ kg/m^2 (obese). Hypertension, high cholesterol and diabetes status were determined through self-reported physician diagnosis.

Statistical analysis

The baseline characteristic analysis categorised participants into two groups based on the presence or absence of NAFLD. Categorical data were presented as frequencies and percentages, while continuous data were described using median values and interquartile ranges (P25–P75). Continuous data analysis utilised *t* tests or Mann–Whitney *U* tests, while categorical data analysis employed χ^2 tests. The study investigated the associations between nutritional status indicators (PNI, GNRI and CONUT) and the likelihood of NAFLD using both univariate and multivariate logistic regression models. Participants were divided into quartiles based on their GNRI and PNI scores, with the first quartile serving as a reference for comparison. OR and 95 % CI were calculated to assess the prevalence of NAFLD across different quartiles. The study employed multivariate logistic regression models with progressive adjustments: Model 1 adjusted for age and sex, model 2 included additional adjustments for race, PIR, BMI, hypertension, diabetes, high cholesterol and daily energy intake, and model 3 further adjusted for smoking status and physical activity levels.

Restricted cubic spline analysis was used to explore the nonlinear relationships between nutritional status indices and NAFLD. Knots were placed at the 5th, 25th, 50th, 75th and 95th percentiles. A likelihood ratio test was conducted to assess the nonlinearity of these relationships. The diagnostic performance of PNI, GNRI and CONUT for NAFLD was evaluated using the area under the receiver operating characteristic curve (AUROC). To address potential collinearity among variables, an eXtreme gradient boosting (XGBoost) model was applied to assess the importance of nutritional status indices. XGBoost, an optimised gradient boosting library, is known for its efficiency and flexibility[22]. All 1409 participants were included in the training dataset. A random 30 % subset of the data was then designated as the prediction set to evaluate the model's predictive accuracy. Variable selection in the XGBoost model was informed by variance inflation factor (VIF) and Relief-F analysis[23]. Shapley Additive exPlanations (SHAP) values were used to identify and visualise the key features influencing NAFLD. These values indicate the importance of individual predictors in determining the model outcome[24]. The model settings included a maximum of fifty iterations, a tree depth of 10 and a learning rate (eta) of 0·1, with cross-validation at each boosting iteration to prevent overfitting.

Stratified analyses were conducted to further examine the relationship between NAFLD and various demographic and health-related factors. These factors included age, sex, race,

education level, family income:poverty ratio, smoking status, physical activity, BMI, hypertension, high cholesterol and diabetes. Interactions between NAFLD and these factors were assessed using P values for interaction terms. All statistical analyses were performed using R version 4·2·1 (R Foundation for Statistical Computing), with a two-sided P value of $< 0{\cdot}05$ considered statistically significant.

Results

Population characteristics

The study included 1409 participants, with 630 undergoing liver ultrasound transient elastography, yielding an NAFLD prevalence of 44·7 %. The demographic and clinical characteristics are detailed in Table 1. The median (P25–P75) age of participants was 63·0 (58·0–71·0) years, distributed almost evenly between males (48·6 %) and females (51·4 %). The predominant racial group was non-Hispanic White (40·4 %), and the majority were never smokers (55·0 %). The median (P25–P75) BMI was 28·9 (25·6, 33·6) kg/m^2. Individuals with NAFLD were generally younger (mean age: 64·0 *v.* 65·0 years), predominantly male (53·3 % *v.* 44·8 %) and more likely to be classified as obese (62·5 % *v.* 29·4 %) compared with those without NAFLD. Moreover, the NAFLD group exhibited higher prevalence rates of chronic conditions such as hypertension, diabetes and high cholesterol. Physical activity levels were lower in the NAFLD group. Notably, lymphocyte counts were significantly higher in NAFLD participants (2·10 (1·70, 2·70) $\times 10^9$/l) compared with those without NAFLD (1·90 (1·50, 2·40) $\times 10^9$/l) (Table 1).

Association of nutritional indices with non-alcoholic fatty liver disease prevalence

The study applied nutritional indices including CONUT (categorised into 'low risk' < 2 and 'high risk' ≥ 2) and quartiles for PNI and GNRI to evaluate their association with NAFLD. Logistic regression and dose–response analyses were conducted to explore these relationships. Adjusting for covariates, PNI showed a positive association with NAFLD; individuals in the third and fourth quartiles of PNI had higher odds of NAFLD compared with the lowest quartile (Q3: OR = 1·45, 95 % CI (1·03, 2·05); Q4: OR = 2·27, 95 % CI (1·59, 3·24)). Similarly, being in the higher quartiles of GNRI was significantly associated with increased odds of NAFLD (Q4 *v.* Q1: aOR = 1·84; 95 % CI (1·28, 2·65)). Conversely, higher CONUT scores were linked to a decreased prevalence of NAFLD (OR = 0·65, 95 % CI (0·48, 0·87)) (Table 2). Multivariable adjustment in Model 3 did not reveal a significant nonlinear relationship between nutritional indices and NAFLD in restricted cubic spline regression ($P_{overall} < 0{\cdot}01$; $P_{nonlinear} > 0{\cdot}05$) (online Supplementary Fig. S2).

Receiver operating characteristic and eXtreme gradient boosting analysis of the importance of nutritional indices for non-alcoholic fatty liver disease

The receiver operating characteristic analysis was performed to assess the predictive ability of the nutritional indices for NAFLD. The area under the receiver operating characteristic curve (AUROC) for the PNI, when combined with demographic and clinical variables, was 0·774 (95 % CI: 0·751, 0·796), demonstrating a sensitivity of 0·746 and a specificity of 0·655. The CONUT index, in conjunction with demographic and clinical variables, yielded an AUROC of 0·767 (95 % CI: 0·744, 0·789), with a sensitivity of 0·824 and a specificity of 0·571. Additionally, the GNRI, alongside demographic and clinical variables, achieved an AUROC of 0·770 (95 % CI: 0·747, 0·792), with a sensitivity of 0·711 and a specificity of 0·684 (Fig. 2 and online Supplementary Table S2).

To demystify the predictive mechanism of the XGBoost model for NAFLD, we utilised SHAP values to illustrate the impact of each nutritional index on the model. The SHAP summary plot indicated that the nutritional indices (PNI, GNRI and CONUT) had SHAP values of 0·036, 0·025 and 0·012, respectively, highlighting their significance in the model (Fig. 3(a)). Further, the SHAP dependency analysis provided insight into how individual features influenced the XGBoost model's predictions (Fig. 3(b)), indicating that higher SHAP values of a feature corresponded to an increased likelihood of NAFLD.

Stratified analyses

A significant interaction was observed between the PNI and race regarding NAFLD presence ($P_{for\ interaction} = 0{\cdot}015$) (Fig. 4). Specifically, in the Mexican American subgroup compared with the reference group, the OR and 95 % CI for NAFLD in the fourth quartile were 1·21 (0·46, 3·24), while in the Non-Hispanic Black subgroup, the OR and 95 % CI were 1·91 (1·04, 3·52). A significant interaction was also identified between the GNRI score and hypertension in NAFLD incidence ($P_{for\ interaction} = 0{\cdot}015$), with a stronger association between GNRI score and NAFLD prevalence observed in hypertensive individuals compared with those without hypertension. Additionally, a notable interaction between BMI and GNRI in NAFLD incidence was found ($P_{for\ interaction} = 0{\cdot}008$). However, there were no significant interactions between nutritional indices and other stratified variables regarding NAFLD presence.

Discussion

To our knowledge, this is the first study to investigate the independent associations of various nutritional status indices with NAFLD in a nationally representative population of individuals aged over 50 years. Our analysis, after adjusting for potential confounders, indicates that relatively poorer nutritional status, as assessed by the PNI or GNRI, is significantly associated with the prevalence of NAFLD. Notably, a higher CONUT score exhibited an independent protective effect against the incidence of NAFLD. Further substantiation through receiver operating characteristic curve and XGBoost analyses underscores the significance of these nutritional indices in relation to NAFLD. Stratified analyses revealed that the associations varied across different subgroups based on race, BMI and hypertension, highlighting the nuanced nature of these relationships.

The prevalence of NAFLD and the critical need for actionable strategies to curb its progression underscore the importance of

Table 1. Characteristics of participants with or without NAFLD in NHANES from 2017 to 2018

	Overall (*n* 1409)		Non-NAFLD (*n* 779)		NAFLD (*n* 630)		
Variables	Median	P25–P75	Median	P25–P75	Median	P25–P75	*P*-value*
Age, years	63·0	58·0–71·0	64·0	58·0–72·0	63·0	57·0–70·0	0·013
	n	%	*n*	%	*n*	%	
Sex, *n* (%)							0·002
Male	685	48·6	349	44·8	336	53·3	
Female	724	51·4	430	55·2	294	46·7	
Race, *n* (%)							0·001
Mexican American	160	11·4	69	8·9	91	14·4	
Other Hispanic	146	10·4	81	10·4	65	10·3	
Non-Hispanic White	570	40·5	304	39·0	266	42·2	
Non-Hispanic Black	313	22·2	197	25·3	116	18·4	
Other Race	220	15·6	128	16·4	92	14·6	
	Median	P25–P75	Median	P25–P75	Median	P25–P75	
BMI, kg/m^2	28·9	25·6–33·6	27·0	23·7–30·6	31·9	28·0–37·0	< 0·001
	n	%	*n*	%	*n*	%	
BMI, *n* (%)							< 0·001
Normal weight	302	21·4	255	32·7	47	7·46	
Overweight	484	34·4	295	37·9	189	30·0	
Obese	623	44·2	229	29·4	394	62·5	
Education, *n* (%)							0·808
Less than high school	266	18·9	151	19·4	115	18·3	
High school or equivalent	351	24·9	190·	24·4	161	25·6	
College or above	792	56·2	438·	56·2	354	56·2	
Poverty-to-income ratio, *n* (%)							0·808
< 1·0	217	15·4	117	15·0	100	15·9	
1·0–3·0	648	46·0	364	46·7	284	45·1	
> 3·0	544	38·6	298	38·3	246	39·0	
	Median	P25–P75	Median	P25–P75	Median	P25–P75	
Total energy, kcal	1764	1372–2277	1740	1361–2256	1785	1406–2318	0·119
	n	%	*n*	%	*n*	%	
Physical activity, *n* (%)							< 0·001
No	814	57·8	415	53·3	399	63·3	
Yes	595	42·2	364	46·7	231	36·7	
Smoking status, *n* (%)							0·001
Non-smoker	776	55·1	433	55·6	343	54·4	
Former smoker	452	32·1	226	29·0	226	35·9	
Current smoker	181	12·8	120	15·4	61	9·68	
Hypertension, *n* (%)							< 0·001
Yes	774	54·9	378	48·5	396	62·9	
No	635	45·1	401	51·5	234	37·1	
Diabetes, *n* (%)							< 0·001
Yes	375	26·6	136	17·5	239	37·9	
No	1034	73·4	643	82·5	391	62·1	
High cholesterol, *n* (%)							0·027
Yes	749	53·2	393	50·4	356	56·5	
No	660	46·8	386	49·6	274	43·5	
	Median	P25–P75	Median	P25–P75	Median	P25–P75	
Lymphocyte count(10^9/l)	2·00	1·60–2·60	1·90	1·50–2·40	2·10	1·70–2·70	< 0·001
Albumin (g/l)	40·0	38·0–42·0	40·0	38·0–42·0	40·0	38·0–42·0	0·353
Total cholesterol (mg/dl)	190	162–219	191	164–219	187	160–217	0·215

NAFLD, non-alcoholic fatty liver disease; NHANES, National Health and Nutrition Examination Survey.
Continues variables displayed as median (P25–P75) and categorical variables are displayed as numbers (percentages).
* Mann–Whitney *U* Test or $\chi 2$ test where appropriate.

dietary quality as a pivotal public health focus. Prior research has established a robust connection between nutritional status and NAFLD. For instance, a study leveraging the NHANES dataset with 10 052 participants showed a strong link between nutrition-related dietary inflammatory indices and NAFLD[25]. Traditionally, investigations into the nexus between nutritional status and NAFLD have primarily centered on diet-related indicators. A 2021 study, for example, delved into the relationship between dietary protein levels and NAFLD incidence[26]. Yet, existing data suggest that dietary variations alone might not fully account for changes in physiological markers in individuals with NAFLD. While some studies propose that a high-protein diet could reduce liver fat irrespective of the protein source, others indicate that an increase in dietary protein, especially certain amino acids, could exacerbate liver fat accumulation and disease severity[27]. Hence, integrating nutritional and inflammatory assessment parameters with physiological indices like serum albumin and lymphocyte count might offer a more direct reflection of the link between nutritional status and NAFLD, enhancing the reliability of the findings[21]. PNI, CONUT and

Table 2. Logistic regression analysis on the association between the nutritional indices and NAFLD

Nutritional Indices	NAFLD (*n* 630)		Non-NAFLD (*n* 779)		Model 1		*P*-value	Model 2		*P*-value	Model 3		*P*-value
	n	%	*n*	%	OR	95 % CI		OR	95 % CI		OR	95 % CI	
PNI													
Q1	134·	21·3	226·	29·0	Reference			Reference			Reference		
Q2	151·	24·0	194·	24·9	1·29	0·95, 1·75	0·102	1·31	0·93, 1·84	0·123	1·32	0·93, 1·86	0·115
Q3	164·	26·0	207·	26·6	1·26	0·93, 1·70	0·132	1·39	0·99, 1·95	0·060	1·45	1·03, 2·05	0·033
Q4	181·	28·7	152·	19·5	1·92	1·41, 2·60	< 0·001	2·17	1·52, 3·08	< 0·001	2·27	1·59, 3·24	< 0·001
$P_{\text{for trend}}$					< 0·001			< 0·001			< 0·001		
GNRI													
Q1	171·	27·1	237·	30·4	Reference			Reference			Reference		
Q2	156·	24·8	208·	26·7	1·01	0·76, 1·35	0·953	0·98	0·70, 1·36	0·895	0·98	0·70, 1·36	0·887
Q3	150·	23·8	181·	23·2	1·07	0·80, 1·45	0·633	1·18	0·84, 1·66	0·337	1·23	0·88, 1·74	0·229
Q4	153·	24·3	153·	19·6	1·27	0·94, 1·72	0·121	1·78	1·24, 2·56	0·002	1·84	1·28, 2·65	0·001
$P_{\text{for trend}}$					0·122			0·001			0·001		
CONUT													
< 2	489·	77·6	565·	72·5	Reference			Reference			Reference		
≥ 2	141·	22·4	214·	27·5	0·75	0·59, 0·97	0·030	0·66	0·49, 0·88	0·005	0·65	0·48, 0·87	0·004

NAFLD, non-alcoholic fatty liver disease; PNI, prognostic nutritional index; GNRI, geriatric nutritional risk index; COUNT, controlling nutritional status; Q1, quartile 1; Q2, quartile 2; Q3, quartile 3; Q4, quartile 4; PIR, poverty:income ratio.
P value for trend calculated treating the nutritional indices (quartile) as a continuous variable.
Model 1: Adjusted for age and sex.
Model 2: Model 1 + race, PIR, BMI, hypertension, diabetes, high cholesterol, total daily energy intake.
Model 3: Model 2 + smoking status and physical activity.

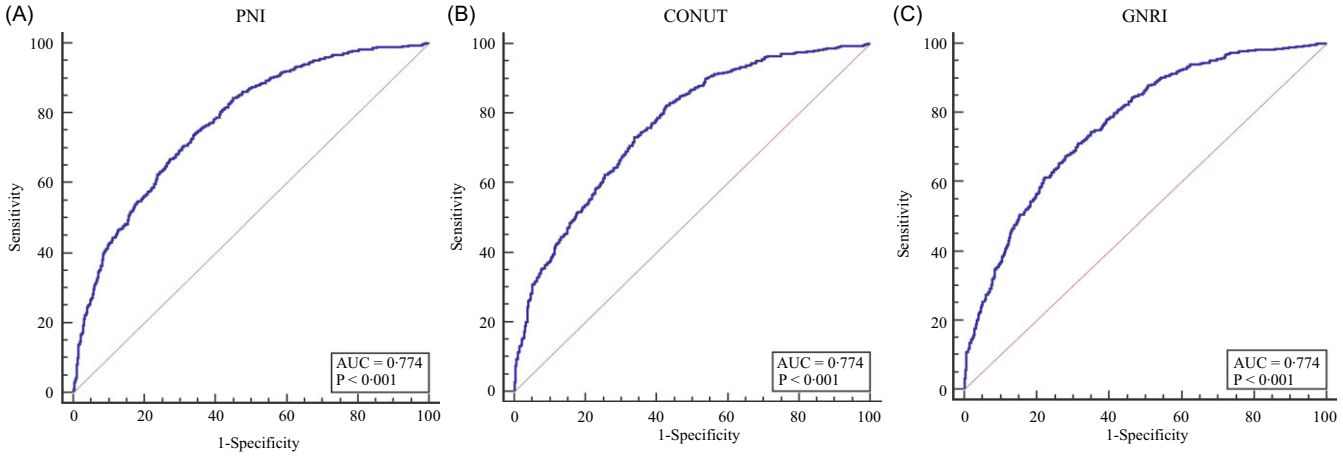

Fig. 2. The AUROC of nutritional indices for NAFLD. AUROC/AUC, area under the receiver operating characteristics; A stands for PNI; B stands for COUNT; C stands for GNRI. AUROC, area under the receiver operating characteristic curve; CONUT, controlling nutritional status; GNRI, geriatric nutritional risk index; NAFLD, non-alcoholic fatty liver disease; PIR, PNI, prognostic nutritional index. All models were adjusted for race, PIR, BMI, hypertension, diabetes, high cholesterol, total daily energy intake, smoking status and physical activity.

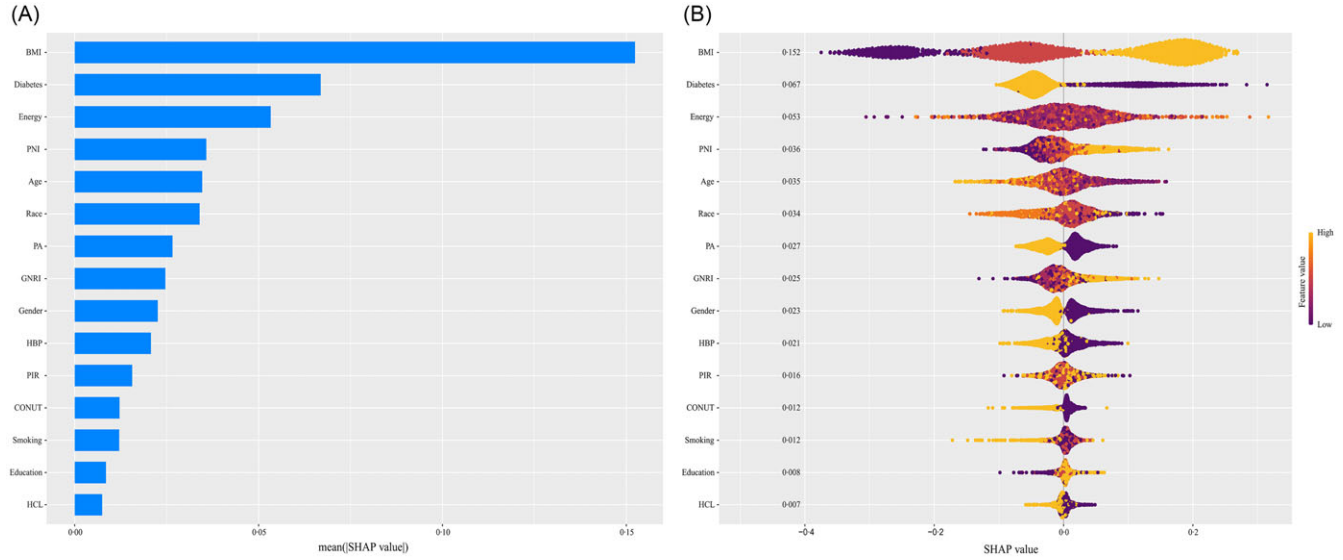

Fig. 3. The XGBoost algorithm determines the relative importance of each variable on NAFLD and assigns a variable importance score to each variable. (A) Importance matrix and SHAP summary plot showing nutritional indices and baseline characteristics contributing to the XGBoost model. The X-axis represents the importance score, which is the relative importance of variables used to distribute the data; the Y-axis represents the variables chosen. (B) SHAP summary plot for the gradient boosted trees trained on the NAFLD prediction task. The colours represent feature values for numeric features: red for larger values and blue for smaller. The thickness of the line comprised of individual dots is determined by the number of examples at a given value. All models were adjusted for race, PIR, BMI, hypertension, diabetes, high cholesterol, total daily energy intake, smoking status and physical activity. CONUT, controlling nutritional status; GNRI, geriatric nutritional risk index; HCL, high cholesterol; NAFLD, non-alcoholic fatty liver disease; PIR, poverty-to-income ratio; PNI, prognostic nutritional index; SHAP, Shapley Additive exPlanations; XGBoost, eXtreme gradient boosting.

GNRI are widely recognised in clinical settings for evaluating nutritional status and predicting outcomes in cancer and various chronic conditions such as hypertension[28]. Our study corroborates the significant relationship between these indices and NAFLD, contributing valuable insights into the intricate interplay between nutrition and liver health.

In this study, all three nutritional status indicators – PNI, GNRI and CONUT – incorporate serum albumin levels in their assessments, reflecting the established role of albumin as an indicator of nutritional status, particularly in post-surgical and post-chemotherapy patients[29]. Numerous studies have identified decreased serum albumin levels as a risk factor for NAFLD[30]. Albumin performs critical physiological functions, including immunomodulation, endothelial stabilisation and antioxidant and anti-inflammatory effects, and it interacts with various drugs, toxins and molecules. Typically, low albumin levels are indicative of malnutrition related to inflammation[31]. The early detection of NAFLD is crucial; however, traditional biochemical markers like albumin levels and liver enzymes might not show significant changes in the initial stages of NAFLD,

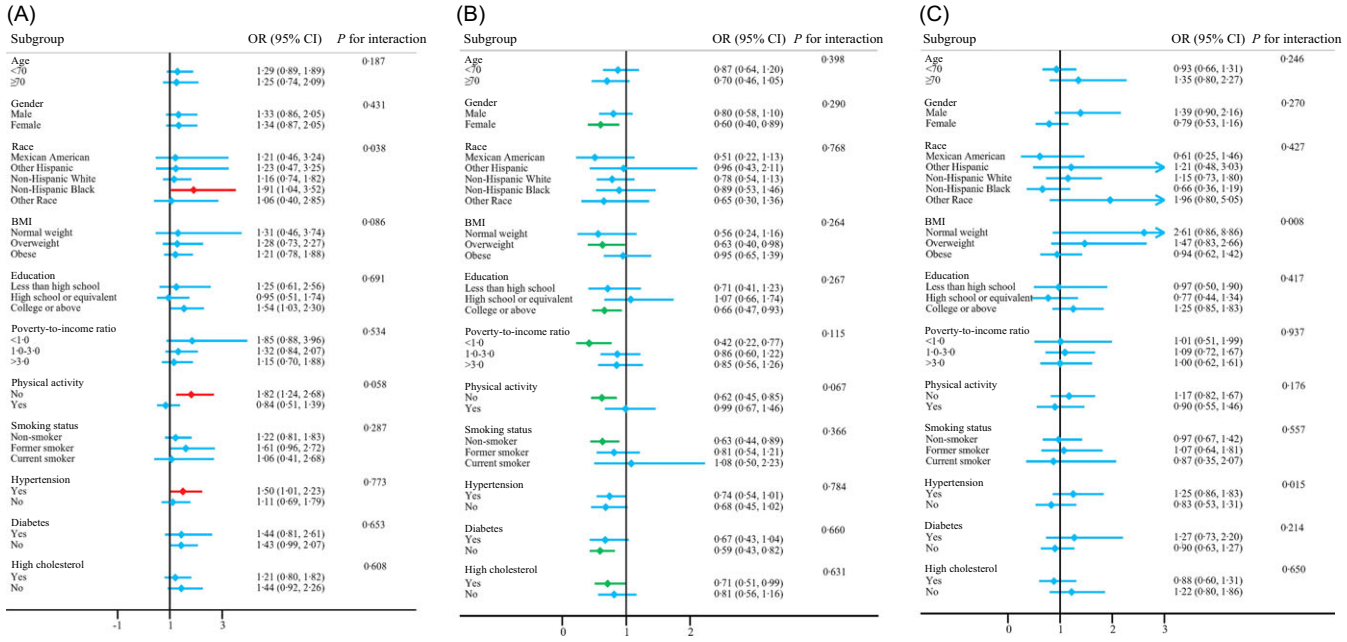

Fig. 4. Associations of nutritional indices with NAFLD in various subgroups in NHANES 2017–2018. (A): PNI; (B): CONUT; (3) GNRI; the colours red, green and blue stand for positive, null and negative significant association, respectively. CONUT, controlling nutritional status; GNRI, geriatric nutritional risk index; NAFLD, nonalcoholic fatty liver disease; NHANES, National Health and Nutrition Examination Survey; PNI, prognostic nutritional index. All models were adjusted for race, PIR, BMI, hypertension, diabetes, high cholesterol, total daily energy intake, smoking status and physical activity.

rendering them nearly indistinguishable from those in individuals without the disease[32]. As the disease progresses to fibrosis and eventually cirrhosis, there is a marked reduction in liver function[33]. Notably, individuals with liver fibrosis or cirrhosis exhibit significantly lower serum albumin levels due to impaired hepatocyte function, nutritional deficiencies, malabsorption and other factors[34]. This correlation was supported by a multicentre study revealing a significant negative relationship between serum albumin levels and liver fibrosis or cirrhosis complications, such as varices and hepatic encephalopathy. Moreover, conditions like cirrhosis not only diminish albumin synthesis but also induce specific changes in albumin structure and functionality. Serum albumin is another major component of the PNI and CONUT formulas, and the above observations may also partially explain the relationship between the nutritional status index and the prevalence of NAFLD in the present study. Future larger prospective studies as well as mechanistic studies are needed to elucidate the association.

The susceptibility to NAFLD is highly variable, influenced by a spectrum of factors, including environmental elements such as dietary habits and physical activity, as well as genetic and epigenetic risk factors[35]. The rise in living standards and shifts in dietary patterns have led to widespread excessive caloric intake and suboptimal nutritional habits, exacerbating the issue of global nutritional surplus[36]. Nutritional surplus, a condition where energy intake surpasses the body's requirements, leading to energy storage and fat accumulation, is a critical contributor to the onset and progression of NAFLD. Dietary influences play a significant role, where an excess caloric intake, particularly from diets rich in saturated fats and carbohydrates (characterised by higher PNI and GNRI scores and lower CONUT scores), is associated with NAFLD development[37]. The potential mechanism might include the liver's heightened metabolism of glucose and fats under excessive sugar and fat intake, resulting in hepatic fat accumulation[38]. This fat deposition induces cellular stress and inflammation, serving as a primary pathophysiological foundation for NAFLD[39]. Furthermore, chronic low-grade inflammation is crucial in the progression from NAFLD to more severe conditions like liver fibrosis and cirrhosis, damaging liver tissue and function through the activation of inflammatory pathways[40]. Insulin resistance also plays a role by diminishing the liver's and other tissues' response to insulin, promoting fat accumulation in the liver and triggering inflammatory processes[41]. Genetic predispositions significantly influence NAFLD risk, with certain populations displaying higher genetic susceptibility, making them more prone to NAFLD even with lower body weights and healthier diets[42]. In conclusion, nutritional surplus drives NAFLD pathogenesis through various mechanisms. Therefore, NAFLD prevention and treatment strategies should encompass comprehensive measures to control nutrient intake, ameliorate metabolic disorders and mitigate inflammatory responses[43], aiming to diminish NAFLD's adverse impacts on both public and individual health.

The blood lymphocyte count, incorporated as a component of the nutritional status indices, exhibited a correlation with NAFLD prevalence in our study. Few studies have scrutinised the association between lymphocyte count and NAFLD[44]. Earlier research has investigated the diagnostic utility of inflammatory indices as markers for NAFLD and its severity[45]. Nutritional intake may offer protection against NAFLD, independent of the

diet type. Specifically, pro-inflammatory nutrients such as total, saturated and *trans*-fats may contribute to NAFLD pathogenesis by fostering low-grade systemic inflammation[46]. Excessive consumption can lead to increased hepatic TAG deposition, triggering a cascade of metabolic disturbances including insulin resistance, hyperinsulinaemia, hepatic inflammation, oxidative stress, mitochondrial dysfunction, an imbalance in pro-inflammatory cytokines and fibrosis – all associated with NAFLD[47]. Moreover, an unbalanced diet is a known risk factor for obesity, type 2 diabetes and metabolic syndrome, conditions intimately linked with NAFLD[48].

To our knowledge, this study is pioneering in elucidating the impact of nutritional status on liver health by examining the association between nutritional indices and NAFLD incidence. Leveraging the capabilities of machine learning in clinical contexts, we employed the XGBoost model in conjunction with receiver operating characteristic curve analysis to substantiate the significance of nutritional status indices in predicting NAFLD, thereby enhancing the credibility of our results. At the same time, our study is comparable to the prevalence of NAFLD observed using other indicators of nutritional status (BMI, WC and body fat) or dietary quality indicators[49], suggesting that our findings are significantly representative. Nevertheless, our study has limitations. The cross-sectional design constrains our capacity to establish causal relationships. Moreover, dietary data were collected through 24-hour recalls, which may not accurately reflect long-term dietary habits. However, the dietary intake assessment tool used in NHANES has undergone extensive validation against dietary records and biomarkers, and we adjusted for socio-economic status, racial groups and other variables. What is more, given that previous studies have suggested that reduced serum albumin concentrations may also be due to hepatic dysfunction[49], there is a need to explore the association between nutritional indices of concern and varying degrees of fibrosis severity in studies with large population samples. Finally, certain factors potentially influencing nutritional status indices, like gastrointestinal diseases, were not included due to data unavailability[50]. Further research is necessary to delve deeper into the nexus between nutritional status indices and NAFLD.

Conclusions

In conclusion, our findings indicate that poorer nutritional status, signifying overnutrition as gauged by PNI, GNRI and CONUT, is positively associated with NAFLD prevalence among individuals aged over 50 years. This suggests the potential to focus on screening for NAFLD among older adults with signs of overnutrition, incorporating them into early-stage clinical management. Enhancing overall nutritional status could reduce NAFLD prevalence. However, more extensive prospective studies are essential to determine if a causal relationship exists between nutritional status and NAFLD prevalence.

Acknowledgements

The authors thank the staff and the participants of the NHANES study for their valuable contributions.

This work was supported by grants from National Natural Science Foundation of China (82074386), National Natural Science Foundation of China (82305334), preclinical study of a new Chinese herbal medicine for the treatment of ascites of liver cirrhosis (spleen and kidney yang deficiency type) with the clinical formula of Qigui Xiaogu Cataplasm (No. 23S21900100), Traditional Chinese Medicine/Chinese and Western Medicine Advantage Specialty Construction Specialty for Department of Hepatology (No. YW(2023-2024)-01-03), Construction of special disease alliance of traditional Chinese medicine in East China area and municipal level, Shanghai special disease alliance of traditional Chinese medicine for liver cirrhosis ascites (water sickness) and Clinical Research Plan of SHDC (No. SHDC2020CR3095B).

H. C.: Conceptualisation and writing – original draft. S. G.: Methodology and software. Y. D.: Methodology, software, and writing – original draft. J. D.: Data curation and resources. G. Z.: Validation and supervision. J. Z.: Supervision, writing – review and editing and project administration.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

This study involving human participants was rigorously evaluated and received the necessary approvals from the National Center for Health Statistics and the Institutional Review Board. Prior to participating in this study, all patients provided written consent after receiving comprehensive information about the purpose and procedures of the study.

The datasets generated and/or analysed during the current study are available in https://www.cdc.gov/nchs/nhanes/index.htm.

Supplementary material

For supplementary material/s referred to in this article, please visit https://doi.org/10.1017/S0007114524001442

References

1. Chalasani N, Younossi Z, Lavine JE, *et al*. (2018) The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of liver diseases. *Hepatology* **67**, 328–357.
2. Younossi ZM, Tampi R, Priyadarshini M, *et al*. (2019) Burden of illness and economic model for patients with nonalcoholic steatohepatitis in the United States. *Hepatology* **69**, 564–572.
3. Altajar S & Baffy G (2020) Skeletal muscle dysfunction in the development and progression of nonalcoholic fatty liver disease. *J Clin Transl Hepatol* **8**, 414–423.
4. Peng H, Wang S, Wang M, *et al*. (2022) Nonalcoholic fatty liver disease and cardiovascular diseases: a Mendelian randomization study. *Metabolism* **133**, 155220.
5. (2016) EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *Diabetologia* **59**, 1121–1140.
6. Chen L, Fan Z & Lv G (2022) Associations of muscle mass and grip strength with severe NAFLD: a prospective study of 333 295 UK Biobank participants. *J Hepatol* **77**, 1453–1454.

7. George ES, Forsyth A, Itsiopoulos C, *et al.* (2018) Practical dietary recommendations for the prevention and management of nonalcoholic fatty liver disease in adults. *Adv Nutr* **9**, 30–40.
8. Cederholm T, Jensen GL, Correia M, *et al.* (2019) GLIM criteria for the diagnosis of malnutrition - a consensus report from the global clinical nutrition community. *J Cachexia Sarcopenia Muscle* **10**, 207–217.
9. Zhang X & Edwards BJ (2019) Malnutrition in older adults with cancer. *Curr Oncol Rep* **21**, 80.
10. Xing J, Guan X, Zhang Q, *et al.* (2021) Triglycerides mediate Body Mass Index and nonalcoholic fatty liver disease: a population-based study. *Obes Facts* **14**, 190–196.
11. Chen G, Fan L, Yang T, *et al.* (2024) Prognostic nutritional index (PNI) and risk of non-alcoholic fatty liver disease and advanced liver fibrosis in US adults: evidence from NHANES 2017–2020. *Heliyon* **10**, e25660.
12. Ignacio de Ulíbarri J, González-Madroño A, de Villar NG, *et al.* (2005) CONUT: a tool for controlling nutritional status. First validation in a hospital population. *Nutr Hosp* **20**, 38–45.
13. Ning Y, Pan D, Guo J, *et al.* (2023) Association of prognostic nutritional index with the risk of all-cause mortality and cardiovascular mortality in patients with type 2 diabetes: NHANES 1999–2018. *BMJ Open Diabetes Res Care* **11**, e003564.
14. Bauer KC, Littlejohn PT, Ayala V, *et al.* (2022) Nonalcoholic fatty liver disease and the gut-liver axis: exploring an under-nutrition perspective. Gastroenterology 162, 1858–1875.e1852.
15. Pan J, Hu Y, Pang N, *et al.* (2023) Association between dietary niacin intake and nonalcoholic fatty liver disease: NHANES 2003–2018. *Nutrients* **15**, 4128.
16. Katsiki N, Stoian AP & Rizzo M (2022) Dietary patterns in non-alcoholic fatty liver disease (NAFLD): stay on the straight and narrow path! *Clin Investig Arterioscler* **34**, S24–S31.
17. Huang Z (2022) Association between blood lead level with high blood pressure in US (NHANES 1999–2018). *Front Public Health* **10**, 836357.
18. Vilar-Gomez E, Vuppalanchi R, Mladenovic A, *et al.* (2023) Prevalence of High-risk Nonalcoholic Steatohepatitis (NASH) in the United States: results from NHANES 2017–2018. *Clin Gastroenterol Hepatol* **21**, 115–124.e117.
19. Siddiqui MS, Vuppalanchi R, Van Natta ML, *et al.* (2019) Vibration-controlled transient elastography to assess fibrosis and steatosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol* **17**, 156–163.e152.
20. Yamaura T, Arizumi F, Maruo K, *et al.* (2022) The impact of Controlling Nutritional Status (CONUT) score on functional prognosis in hospitalized elderly patients with acute osteoporotic vertebral fractures. *BMC Geriatr* **22**, 1002.
21. Peng P, Chen L, Shen Q, *et al.* (2023) Prognostic Nutritional Index (PNI) and Controlling Nutritional Status (CONUT) score for predicting outcomes of breast cancer: a systematic review and meta-analysis. *Pak J Med Sci* **39**, 1535–1541.
22. Zhang C, Zhang Y, Shi X, *et al.* (2019) On incremental learning for gradient boosting decision trees. *Neural Process Lett* **50**, 957–987.
23. Ibrahim AA, Ridwan RL, Muhammed MM, *et al.* (2020) Comparison of the CatBoost classifier with other machine learning methods. *Int J Adv Comput Sci Appl* **11**, 11.
24. Lundberg SM, Erion G, Chen H, *et al.* (2020) From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* **2**, 56–67.
25. Petermann-Rocha F, Wirth MD, Boonpor J, *et al.* (2023) Associations between an inflammatory diet index and severe non-alcoholic fatty liver disease: a prospective study of 171 544 UK Biobank participants. *BMC Med* **21**, 123.
26. Charatcharoenwitthaya P, Tansakul E, Chaiyasoot K, *et al.* (2021) Dietary composition and its association with newly diagnosed nonalcoholic fatty liver disease and insulin resistance. *Nutrients* **13**, 4438.
27. Ruiz-Margáin A, Macías-Rodríguez RU, Ríos-Torres SL, *et al.* (2018) Effect of a high-protein, high-fiber diet plus supplementation with branched-chain amino acids on the nutritional status of patients with cirrhosis. *Rev Gastroenterol Mex (Engl Ed)* **83**, 9–15.
28. Shiroma K, Tanabe H, Takiguchi Y, *et al.* (2023) A nutritional assessment tool, GNRI, predicts sarcopenia and its components in type 2 diabetes mellitus: a Japanese cross-sectional study. *Front Nutr* **10**, 1087471.
29. Zhu M, Chen L, Kong X, *et al.* (2022) Controlling Nutritional Status (CONUT) as a novel postoperative prognostic marker in breast cancer patients: a retrospective study. *Biomed Res Int* **2022**, 3254581.
30. Takahashi H, Kawanaka M, Fujii H, *et al.* (2023) Association of serum albumin levels and long-term prognosis in patients with biopsy-confirmed nonalcoholic fatty liver disease. *Nutrients* **15**, 2014.
31. Eckart A, Struja T, Kutz A, *et al.* (2020) Relationship of nutritional status, inflammation, and serum albumin levels during acute illness: a prospective study. *Am J Med* **133**, 713–722.e717.
32. Liu CF & Chien LW (2023) Predictive role of Neutrophil-Percentage-to-Albumin Ratio (NPAR) in nonalcoholic fatty liver disease and advanced liver fibrosis in nondiabetic US adults: evidence from NHANES 2017–2018. *Nutrients* **15**, 1892.
33. Roehlen N, Crouchet E & Baumert TF (2020) Liver fibrosis: mechanistic concepts and therapeutic perspectives. *Cells* **9**, 875.
34. Di Stefano G, Fiume L, Domenicali M, *et al.* (2006) Doxorubicin coupled to lactosaminated albumin: effects on rats with liver fibrosis and cirrhosis. *Dig Liver Dis* **38**, 404–408.
35. Buzzetti E, Pinzani M & Tsochatzis EA (2016) The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NAFLD). *Metab* **65**, 1038–1048.
36. Buyukdere Y & Akyol A (2023) From a toxin to an obesogen: a review of potential obesogenic roles of acrylamide with a mechanistic approach. *Nutr Rev* **82**, 128–142.
37. Yki-Järvinen H, Luukkonen PK, Hodson L, *et al.* (2021) Dietary carbohydrates and fats in nonalcoholic fatty liver disease. *Nat Rev Gastroenterol Hepatol* **18**, 770–786.
38. Broadfield LA, Duarte JAG, Schmieder R, *et al.* (2021) Fat induces glucose metabolism in nontransformed liver cells and promotes liver tumorigenesis. *Cancer Res* **81**, 1988–2001.
39. Li J, Wang T, Liu P, *et al.* (2021) Hesperetin ameliorates hepatic oxidative stress and inflammation via the PI3K/AKT-Nrf2-ARE pathway in oleic acid-induced HepG2 cells and a rat model of high-fat diet-induced NAFLD. *Food Funct* **12**, 3898–3918.
40. Milić S, Lulić D & Štimac D (2014) Non-alcoholic fatty liver disease and obesity: biochemical, metabolic and clinical presentations. *World J Gastroenterol* **20**, 9330–9337.
41. Ter Horst KW, Vatner DF, Zhang D, *et al.* (2021) Hepatic insulin resistance is not pathway selective in humans with nonalcoholic fatty liver disease. *Diabetes Care* **44**, 489–498.
42. Flessa CM, Nasiri-Ansari N, Kyrou I, *et al.* (2022) Genetic and diet-induced animal models for Non-Alcoholic Fatty Liver Disease (NAFLD) research. *Int J Mol Sci* **23**, 15791.
43. Mascaró CM, Bouzas C, Montemayor S, *et al.* (2022) Effect of a six-month lifestyle intervention on the physical activity and fitness status of adults with NAFLD and metabolic syndrome. *Nutrients* **14**, 1813.
44. Song Y, Guo W, Li Z, *et al.* (2022) Systemic immune-inflammation index is associated with hepatic steatosis: evidence from NHANES 2015–2018. *Front Immunol* **13**, 1058779.

British Journal of Nutrition

45. Valibeygi A, Davoodi A, Dehghan A, *et al.* (2023) Dietary inflammatory index (DII) is correlated with the incidence of non-alcoholic fatty liver disease (NAFLD): Fasa PERSIAN cohort study. *BMC Nutr* **9**, 84.
46. Monserrat-Mesquida M, Quetglas-Llabrés MM, Bouzas C, *et al.* (2023) Plasma fatty acid composition, oxidative and inflammatory status, and adherence to the Mediterranean Diet of patients with Non-Alcoholic Fatty Liver Disease. *Antioxidants (Basel)* **12**, 1554.
47. Wang Y, Chen C, Chen J, *et al.* (2022) Overexpression of NAG-1/GDF15 prevents hepatic steatosis through inhibiting oxidative stress-mediated dsDNA release and AIM2 inflammasome activation. *Redox Biol* **52**, 102322.
48. Montemayor S, Mascaró CM, Ugarriza L, *et al.* (2022) Adherence to Mediterranean Diet and NAFLD in patients with metabolic syndrome: the FLIPAN study. *Nutrients* **14**, 3186.
49. Chen L, Fan Z, Sun X, *et al.* (2023) Diet-derived antioxidants and nonalcoholic fatty liver disease: a Mendelian randomization study. *Hepatol Int* **17**, 326–338.
50. Li D, Li Y, Yang S, *et al.* (2022) Diet-gut microbiota-epigenetics in metabolic diseases: from mechanisms to therapeutics. *Biomed Pharmacother* **153**, 113290.

British Journal of Nutrition