Original Article

Association between Dietary Index for Gut Microbiota (DI-GM) and non-alcoholic fatty liver disease (NAFLD): Evidence from NHANES 1999–2018

Yan Xue MD[1] and Jianxian Zhang MD[2]

[1]*Department of Gastroenterology, The Second People's Hospital of Liaocheng City, Shandong, China*
[2]*Department of Gastrointestinal Surgery, The Second People's Hospital of Liaocheng City, Shandong, China*

Background and Objectives: Gut microbiota and liver are closely linked, and disruption of the gut–liver axis has been associated with various conditions, including non-alcoholic fatty liver disease (NAFLD). The Dietary Index for Gut Microbiota (DI-GM), a recently developed measure of gut microbiota variety, has not been researched in connection with NAFLD. **Methods and Study Design:** We conducted a cross-sectional analysis of 12,910 eligible participants aged ≥20 years from the National Health and Nutrition Examination Survey (NHANES) between 1999 and 2018 by adjusting for covariates. Dietary recall data were used to calculate the DI-GM (including components beneficial and unfavorable to gut microbiota). Multiple logistic regression and subgroup analyses were used. **Results:** A total of 12,910 patients were included in the study, of whom 4673 (36.2%) were identified as NAFLD. Each point increase in DI-GM was associated with an 8% decrease in the prevalence of NAFLD (OR = 0.92, 95% CI = 0.90, 0.94, p <0.001), the associations remained significant after adjusting for potential confounders (OR = 0.92, 95% CI = 0.89, 0.95, p <0.001). After grouping DI-GM, in the fully adjusted model, participants with DI-GM ≥ 6 were significantly negatively associated with the prevalence of NAFLD (OR = 0.71, 95% CI = 0.61, 0.82, p <0.001) compared to participants with DI-GM ≤3 group with adjustment for potential confounders. After subgroup analyses and sensitivity analyses, the relationship between DI-GM and NAFLD remained robust. **Conclusions:** Our findings indicate an inverse association between the newly proposed DI-GM and the presence of NAFLD in adult Americans, offering a novel perspective on NAFLD research.

Key Words: dietary index for gut microbiota (DI-GM), non-alcoholic fatty liver disease (NAFLD), NHANES

INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) is among the most common liver diseases, with a global prevalence of around 25%, impacting over 80 million individuals in the United States alone.[1–3] In recent years, the incidence of NAFLD in the US has been on the rise, posing a significant public health challenge.[4–6] Characterized by the abnormal accumulation of lipids in hepatic tissue exceeding 5% without significant alcohol consumption, NAFLD remains a condition with an incompletely understood pathogenesis.[7] The scarcity of effective therapeutic options underscores the necessity for further investigation into its risk factors to inform the development of effective treatment and prevention strategies.[8,9]

The gut microbiota encompasses the varied microbial community residing in the gastrointestinal tract.[10] In their recent analysis, Kase and colleagues, after scrutinizing 106 scholarly articles on the diet-microbiome nexus in adults, identified 14 nutrients that either foster or hinder gut microbial health. This work culminated in the creation of the Dietary Index for Gut Microbiota (DI-GM), an innovative metric aimed at evaluating the nutritional adequacy for preserving a robust gut microbiome. The DI-GM exhibits a positive association with the diversity of the gut microbiome and adeptly discerns dietary practices conducive or detrimental to microbial balance. Consequently, this index holds promise as a benchmark for evaluating diets that support gut microbiome equilibrium.[11]

The intestines and liver are intricately linked through a multifaceted interaction, with perturbations along the gut-liver axis correlating with numerous pathological conditions, such as NAFLD.[12] The pathogenesis of NAFLD is considered to be based on the "multiple hits" theory, where dysbiosis of the gut microbiota plays a key role.[13] This dysbiosis is implicated in both the initiation and advancement of hepatic pathologies via diverse pathways.[14] A wealth of research has highlighted disparities in the diversity of the gut microbiota among NAFLD sufferers. A comprehensive meta-analysis has identified alterations

Corresponding Author: Dr Jianxian Zhang, Department of Gastrointestinal Surgery, The Second People's Hospital of Liaocheng, No. 306 Jiankang Street, Linqing 252600, Shandong province, China
Tel: +86-18264521207
Email: zjx.2006@163.com
Manuscript received 27 November 2024. Initial review completed 18 January 2025. Revision accepted 16 May 2025.
doi: 10.6133/apjcn.202508_34(4).0015

in the prevalence of specific bacterial genera — *Escherichia*, *Prevotella*, *Streptococcus*, *Coprococcus*, *Faecalibacterium*, and *Ruminococcus* — as a hallmark of the intestinal microbiome in NAFLD.[15] Currently, multiple clinical trials are exploring interventions targeting the gut microbiota, encompassing the use of probiotics and prebiotics, fecal microbiota transplantation (FMT), and microbiome-directed therapies (MTT).[16–18] Evidence indicates that nutritional regimens shape the gut microbiota profile, drawing increased scrutiny to dietary modifications.[10,19–21] Intake components, including lipids, ethanol, sugars, dietary fiber, and broader nutritional strategies, profoundly influence the structure and operational dynamics of the gut microbiota, resulting in notable health repercussions.[14]

To our knowledge, studies investigating the association of DI-GM and NAFLD are lacking. Hence, the aim of this research is to explore the link between the DI-GM and NAFLD through analyzing adult participants' data from the National Health and Nutrition Examination Survey (NHANES), providing a novel perspective on the research of NAFLD.

METHODS

Study design and population

Data spanning 10 sequential NHANES data cycles, ranging from 1999 to 2018, were extracted from publicly accessible records. The NHANES constitutes an ongoing cross-sectional observational study, amassing health-related data from a non-institutionalized US population that is representative. This study's protocol was sanctioned by the Institutional Review Board of the National Center for Health Statistics (NCHS), ensuring that all participants had granted their written informed consent. By employing a stratified, multistage probability cluster sampling methodology, NHANES guarantees the collection of both extensive and dependable data.[22] This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline. Our study comprised adults aged 20 years or older who had undergone the interview procedure. From the analysis, we excluded pregnant participants, as well as those with deficient DI-GM information or for whom a conclusive NAFLD diagnosis was unattainable.

Ethical statements

Human studies were sanctioned by the National Center for Health Statistics' Research Ethics Review Board. Conducted in compliance with local laws and institutional protocols, participants gave written consent for study involvement.

Assessment of NAFLD

While liver biopsy is recognized as the definitive diagnostic method for NAFLD, its invasive nature renders it impractical for population-based studies. Consequently, we utilized an alternative marker for NAFLD detection, the United States Fatty Liver Index (US FLI), formulated by Ruhl and Everhart.[23] This index has demonstrated superior accuracy compared to the Fatty Liver Index within the US population.[23] The USFLI relies on readily accessible parameters and enjoys robust validation through multiple scientific studies.[24–26] The calculation formula for USFLI is:

$$\text{USFLI} = e^{(-0.8073 \times \text{non-Hispanic black} + 0.3458 \times \text{Mexican American} + 0.0093 \times \text{age} + 0.6151 \times \ln(\text{GGT}) + 0.0249 \times \text{waist circumference} + 1.1792 \times \ln(\text{insulin}) + 0.8242 \times \ln(\text{glucose}) - 14.7812)} / (1 + e^{(-0.8073 \times \text{non-Hispanic black} + 0.3458 \times \text{Mexican American} + 0.0093 \times \text{age} + 0.6151 \times \ln(\text{GGT}) + 0.0249 \times \text{waist circumference} + 1.1792 \times \ln(\text{insulin}) + 0.8242 \times \ln(\text{glucose}) - 14.7812)}) \times 100.$$

NAFLD is diagnosed when the USFLI score is ≥30 and there is no presence of viral hepatitis (HBV or HCV) or a history of considerable alcohol consumption (>1 alcoholic drink/day for women or >2 alcoholic drinks/day for men), positive hepatitis B surface antigen, positive hepatitis C antibody.[27,28]

Assessment of dietary index for gut microbiota

According to the criteria set by Kase et al.[11, 14] food items and nutrients comprise the DI-GM. These include avocado, broccoli, chickpeas, coffee, cranberries, fermented dairy, fiber, green tea (not detailed in NHANES), soybeans, and whole grains as positive factors, while red meat, processed meat, refined grains, and diets with a high fat content (where fat accounts for 40% or more of total energy) are deemed to be negative components.[11] The dietary recall information from NHANES spanning 1999 to 2018 was employed to calculate the DI-GM. Details of the components and the scoring criteria for the DI-GM are presented in Supplementary Figure 1. For foods that are beneficial to the gut microbiota, a score of 1 was applied if the intake was at or above the sex-specific median; otherwise, it was 0. For items unfavorable to the gut microbiota, a score of 0 was given if the intake was at or above the sex-specific median or 40% for high-fat diets; otherwise, it was 1. These scores were added to derive the total DI-GM score, which varies from 0 to 13 (with beneficial to gut microbiota scores ranging from 0 to 9 and unfavorable to gut microbiota scores from 0 to 4), and then categorized into intervals of 0–3, 4, 5, and 6 or more.

Covariates

Based on both clinical experience and current literature, the following covariates were included: age, sex, race, marital status, family income-to-poverty ratio (PIR), education level, physical activity, smoke, alcohol intake, hypertension, diabetes, cardiovascular disease (CVD), glycated hemoglobin (HbA1c), alanine aminotransferase (ALT), aspartate aminotransferase (AST). Demographic measurements including body mass index (BMI) and blood lipid levels like high-density lipoprotein (HDL) and low-density lipoprotein (LDL) cholesterol were taken into account. Comprehensive details of the lab methods and variables are accessible on the NHANES site, ensuring the clarity and reproducibility of our study's procedures.[29–33] Age was analyzed as a continuous variable in logistic regression, while in subgroup analyses it was categorized as <45 years, and ≥45 years. We categorized the participants into the following 5 races and ethnicities: non-Hispanic White, non-Hispanic Black, Mexican American, other Hispanic and Other (including multiracial). Marital status was classified as married, never married, living with a partner, and other. We categorized family income into the following 3 levels based on the family poverty income ratio: low income (≤1.3), medium

income (>1.3 to 3.5), and high income (>3.5).[34] Education levels were categorized as less than high school, high school or equivalent, and above high school.[30] Physical activity encompasses the time (in minutes) that participants dedicate to various activities throughout the week, including walking, biking, household chores, work-related tasks, and recreational pursuits.[35] Smoking status was categorized into the following 3 groups: never smoked (or smoked <100 cigarettes), former smoker (smoked at least 100 cigarettes but has quit), and current smoker.[36] Alcohol intake was categorized as never (had <12 drinks in lifetime), former (had ≥12 drinks in 1 year and did not drink last year, or did not drink last year but drank ≥12 drinks in lifetime), current mild alcohol use (≤1 drink per day for females, ≤2 drinks per day for males).[37] Hypertension and diabetes mellitus were determined based on medication use, self-reported physician diagnosis, and relevant testing indicators including blood pressure measurements,[38] fasting blood glucose levels, and HbA1c levels.[39] CVD was determined based on self-reported diagnosis (any of the following conditions would be sufficient for diagnosis: coronary heart disease, angina, stroke, heart attack, or congestive heart failure).

Statistical analysis

Descriptive statistics were employed to summarize all data, including frequencies (as percentages) and means with their standard deviations. The Chi-square analysis assessed group differences for categorical data, and either the t-test or the Mann-Whitney U-test was applied to continuous data, based on suitability. Missing values in covariates were addressed using a multivariate single imputation method. This approach utilized an iterative imputer, with a Bayesian Ridge model serving as the estimator in each step of the round-robin imputation process.40 A multiple logistic regression analysis was performed to investigate the relationship between DI-GM score levels and the risk of NAFLD. We selected these confounders on the basis of clinical interest or their associations with the outcomes of interest or a change in effect estimate of more than 10%. In the multiple logistic regression analysis, three distinct models were applied, taking into account a range of sociodemographic and clinical variables. Model 1 was adjusted for age, sex. Model 2 incorporated all covariates from Model 1, with the addition of race, marital status, PIR, education, physical activity, BMI, smoke, alcohol intake; Model 3 was an extension of Model 2, further adjusted for HbA1c, ALT, AST, HDL, LDL, hypertension, cardiovascular disease, diabetes. Furthermore, potential modifications of the relationship between DI-GM and NAFLD were assessed, including the following variables: sex, age (< 45 and ≥45 years), Marital status, BMI (< 30 and ≥30kg/m^2), smoke, coronary heart disease, hypertension, and diabetes. The interaction among subgroups was assessed using the likelihood ratio test, and a sensitivity analysis was carried out after the exclusion of participants with incomplete covariate data. Statistical analyses were performed utilizing R and Free Statistics software version 2.0. In all tests, p <0.05 (2-sided) was considered to indicate statistical significance.

RESULTS

Characteristics of the study population

Our study involved a total of 55,081 participants aged ≥20 years from 1999 to 2018. Exclusion criteria of the analysis involved the pregnant data (n = 1,547), missing data on DI-GM and those ineligible for a diagnosis of NAFLD (n = 40,624). Ultimately, a total of 12,910 patients were included in the study following rigorous screening based on the predefined inclusion and exclusion criteria (Figure 1), of whom 4673(36.2%) were identified as NAFLD. Table 1 presents the characteristics of the study population stratified by DI-GM score. The average age of the study participants was 54.4 (±17.7) years, and 6463 (50.1%) individuals were female. Compared to individuals with lower DI-GM score, those with higher DI-GM tended to be older, had a higher proportion of females, non-Hispanic White, Married, had a high family income, had higher educational attainment, spend less time in physical activity, had lower BMI, and a lower prevalence of diabetes (all p <0.05). However, there were no significant differences observed among the four groups in terms of Hypertension, Cardiovascular diseases, ALT, AST and LDL levels (all p >0.05).

Association between DI-GM and NAFLD

Table 2 presents the results of the multivariable logistic regression analysis examining the association between DI-GM and NAFLD. An inverse association was observed after adjusting for potential confounders. Each point increase in DI-GM was associated with an 8% decrease in the prevalence of NAFLD (OR = 0.92, 95% CI = 0.90, 0.94, p <0.001), the associations remained significant after adjusting for potential confounders (OR = 0.92, 95% CI = 0.89, 0.95, p <0.001). After grouping DI-GM, in the fully adjusted model, participants with DI-GM ≥6 were significantly negatively correlated with the prevalence of NAFLD (OR = 0.71, 95% CI = 0.61, 0.82, p <0.001) compared to participants with DI-GM ≤3 group with adjustment for potential confounders (Table 2, Model 3).

Subgroup analyses

Figure 2 illustrates that no significant interactions were detected following stratification by age (< 45 and ≥45 years), sex, marital status, smoke, hypertension, coronary heart disease and diabetes. Owing to multiple testing, the p-value (0.011) for the interaction within the BMI subgroup (< 30kg/m^2 and ≥30 kg/m^2) might not reach statistical significance.

Sensitivity analyses

After excluding individuals with missing covariates (leaving 11,796 participants), the relationship between DI-GM and NAFLD remained robust in the sensitivity analysis after adjusting the model for multiple logistic analysis (Table 3).

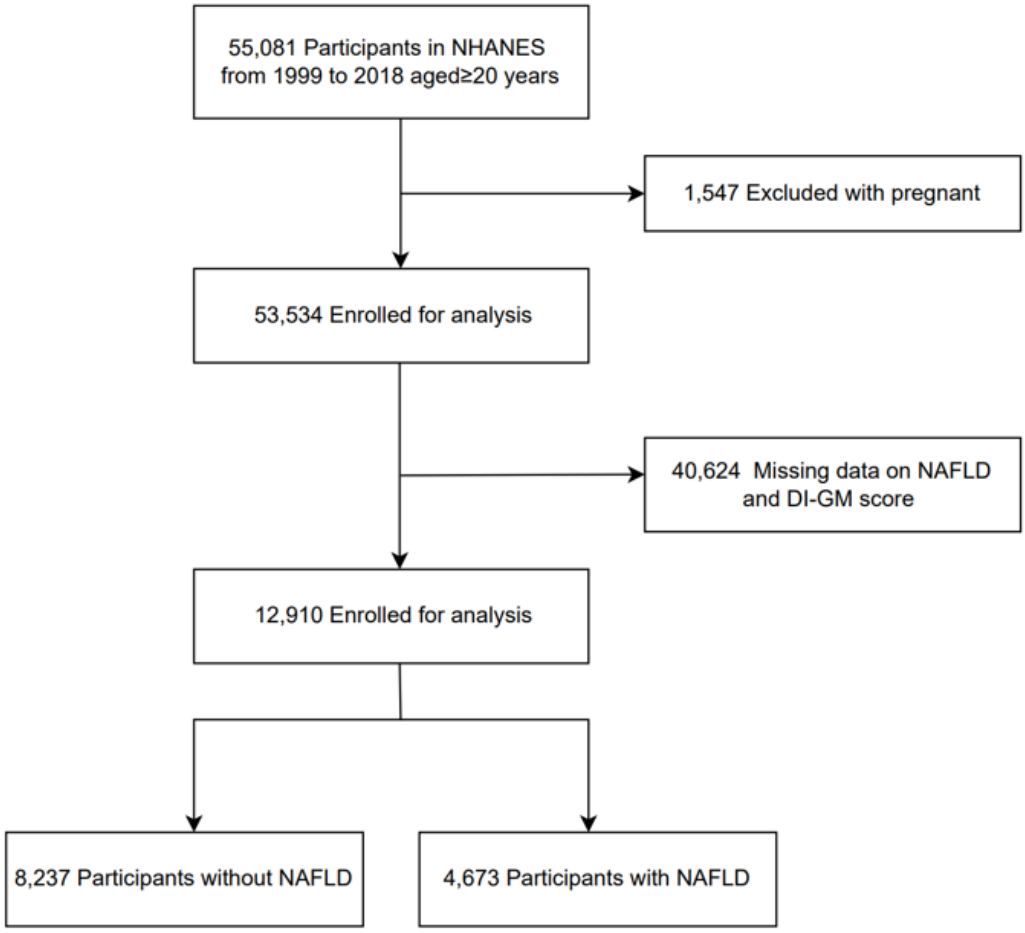

Figure 1. Flow chat of the screening and enrollment of study participants. NHANES, National Health and Nutrition Examination Survey; NAFLD, non-alcoholic fatty liver disease; DI-GM, the dietary index for gut microbiota.

DISCUSSION

In our study, we demonstrated that increases in DI-GM score, DI-GM ≥ 6 group were significantly and negatively associated with the prevalence of NAFLD after using single imputation, and the results remained robust after excluding participants with missing data. Similar patterns of association were observed for subsequent subgroup analysis. These findings have important clinical implications.

Our study observed negative associations of DI-GM with NAFLD in the context of population-based, which is consistent with findings from other observational studies. Studies have shown that changes in diet can induce shifts in the species composition of the gut microbiota and their potential role in NAFLD have historically been emphasized.[20,21,41,42] For instance, grains that are highly refined, categorized as unfavorable to gut microbiota within the DI-GM, constitute a main component of the Western dietary regimen. Overconsumption of such refined grains can result in hyperglycemia, a condition that correlates with inflammation within the gastrointestinal tract.[43] Fermented dairy, classified as beneficial to gut microbiota within the DI-GM, may play a crucial role. A meta-analysis of 19 clinical trials involving human subjects suggests that the consumption of fermented foods could be a key dietary strategy for either preventing or remedying imbalances in the gut microbiota.[44] A randomized controlled trial (RCT) indicated that a diet rich in fermented foods significantly enhanced microbial diversity within the gut. Moreover, it has been demonstrated that diets characterized by a high fat-to-carb ratio are associated with a reduction in gut microbiota diversity.45 Numerous studies have shown differences in gut microbiota diversity in patients with NAFLD. Research on Asian individuals with NAFLD, irrespective of obesity, revealed a reduction in microbial diversity and a shift in bacterial composition, with lower levels of *Ruminococcaceae* and higher levels of *Veillonellaceae*. These microbial alterations were correlated with the severity of hepatic fibrosis.[46] Wang L et al. concluded that in NAFLD patient, the alpha diversity of intestinal flora decreased, and the composition of intestinal flora changed (beta diversity, $p < 0.05$).[47] The paper examines fecal microbiota transplantation (FMT) as a treatment for restoring gut microbiota diversity, potentially benefiting NAFLD patients by improving their condition.[48] Additionally, Fan C et al revealed a robust correlation between dietary intake of live microbes and the prevalence of NAFLD in a cross-sectional analysis, which is consistent with our findings.[31]

The potential mechanisms linking DI-GM and NAFLD involve the impact of different dietary components on gut microbiota diversity. These changes in gut microbiota balance influence NAFLD development through the gut-liver axis. The imbalance of gut microbiota contributes to the occurrence and progression of liver diseases through various mechanisms, including intestinal barrier

Table 1. Baseline characteristics by categories of DI-GM

Variables	Total	1(DI-GM<3)	2 (DI-GM=4)	3 (DI-GM=5)	4 (DI-GM≥6)	*p* value
N	12910	3083	3258	3138	3431	
Age, years, Mean ± SD	54.4 ± 17.7	50.2 ± 18.0	53.1 ± 18.1	55.8 ± 17.2	58.2 ± 16.7	< 0.001
Sex, n (%)						< 0.001
Male	6447 (49.9)	1603 (52)	1687 (51.8)	1536 (48.9)	1621 (47.2)	
Female	6463 (50.1)	1480 (48)	1571 (48.2)	1602 (51.1)	1810 (52.8)	
Race, n (%)						< 0.001
Non-Hispanic White	6158 (47.7)	1265 (41)	1514 (46.5)	1517 (48.3)	1862 (54.3)	
Non-Hispanic Black	2520 (19.5)	891 (28.9)	687 (21.1)	520 (16.6)	422 (12.3)	
Mexican American	2048 (15.9)	487 (15.8)	545 (16.7)	550 (17.5)	466 (13.6)	
Other Hispanic	1015 (7.9)	222 (7.2)	229 (7)	274 (8.7)	290 (8.5)	
Other Race	1169 (9.1)	218 (7.1)	283 (8.7)	277 (8.8)	391 (11.4)	
Marital status, n (%)						< 0.001
Married	7472 (58.4)	1687 (55.4)	1826 (56.6)	1864 (59.9)	2095 (61.6)	
Never married	1666 (13.0)	505 (16.6)	471 (14.6)	355 (11.4)	335 (9.9)	
Living with partner	702 (5.5)	203 (6.7)	205 (6.4)	166 (5.3)	128 (3.8)	
Other	2946 (23.0)	650 (21.3)	724 (22.4)	729 (23.4)	843 (24.8)	
PIR, n (%)						< 0.001
≤1.30	3324 (28.1)	930 (32.6)	937 (31.4)	765 (26.7)	692 (22.1)	
1.3~3.50	4610 (39.0)	1204 (42.2)	1174 (39.3)	1134 (39.6)	1098 (35.1)	
>3.50	3895 (32.9)	718 (25.2)	876 (29.3)	966 (33.7)	1335 (42.7)	
Education, n (%)						< 0.001
< High school	3448 (26.7)	916 (29.7)	949 (29.2)	874 (27.9)	709 (20.7)	
High school or equivalent	2893 (22.4)	782 (25.4)	808 (24.8)	662 (21.1)	641 (18.7)	
> High school	6556 (50.8)	1383 (44.9)	1496 (46)	1600 (51)	2077 (60.6)	
Physical activity, minutes/week, Mean ± SD	2227 ± 4711	2379 ± 5151	2360 ± 5087	2138 ± 4569	2045 ± 3991	0.007
BMI (kg/m^2), Mean ± SD	29.0 ± 6.6	30.0 ± 7.1	29.2 ± 6.8	28.9 ± 6.5	28.1 ± 6.0	< 0.001
Smoke, n (%)						< 0.001
Never	7572 (58.7)	1797 (58.4)	1879 (57.7)	1804 (57.6)	2092 (61)	
Former	3543 (27.5)	760 (24.7)	860 (26.4)	890 (28.4)	1033 (30.1)	
Current	1785 (13.8)	522 (17)	518 (15.9)	440 (14)	305 (8.9)	
Alcohol intake, n (%)						< 0.001
Never	3766 (29.2)	997 (32.3)	1001 (30.7)	923 (29.4)	845 (24.6)	
Former	6538 (50.6)	1445 (46.9)	1580 (48.5)	1571 (50.1)	1942 (56.6)	
Mild	2606 (20.2)	641 (20.8)	677 (20.8)	644 (20.5)	644 (18.8)	

NAFLD, nonalcoholic fatty liver disease; DI-GM, dietary index for gut microbiota; NHANES, National Health and Nutrition Examination Survey; PIR, family income-to poverty ratio; BMI, body mass index; CVD, cardiovascular disease; DM, diabetes mellitus; HbA1c, glycated hemoglobin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

Table 1. Baseline characteristics by categories of DI-GM (cont.)

Variables	Total	1(DI-GM<3)	2 (DI-GM=4)	3 (DI-GM=5)	4 (DI-GM≥6)	*p* value
CVD, n (%)						0.1
No	11120 (86.1)	2691 (87.3)	2813 (86.3)	2672 (85.2)	2944 (85.8)	
Yes	1789 (13.9)	392 (12.7)	445 (13.7)	465 (14.8)	487 (14.2)	
DM, n (%)						0.02
No	10002 (77.5)	2339 (75.9)	2538 (77.9)	2415 (77)	2710 (79)	
Yes	2908 (22.5)	744 (24.1)	720 (22.1)	723 (23)	721 (21)	
HbA1c (%), Mean ± SD	5.8 ± 1.1	5.9 ± 1.2	5.8 ± 1.2	5.8 ± 1.0	5.8 ± 1.0	0.003
ALT(U/L), Mean ± SD	24.0 ± 23.2	24.8 ± 38.0	23.7 ± 13.8	23.9 ± 20.0	23.8 ± 13.6	0.235
AST(U/L), Mean ± SD	24.4 ± 16.6	24.3 ± 19.7	23.9 ± 8.7	24.8 ± 22.6	24.8 ± 12.0	0.066
HDL (mg/dL), Mean ± SD	52.7 ± 15.3	50.7 ± 14.5	52.2 ± 15.2	52.6 ± 15.0	55.0 ± 15.9	< 0.001
LDL (mg/ dL), Mean ± SD	116 ± 36.0	116 ± 36.6	116 ± 36.4	117 ± 36.0	116 ± 34.9	0.574

NAFLD, nonalcoholic fatty liver disease; DI-GM, dietary index for gut microbiota; NHANES, National Health and Nutrition Examination Survey; PIR, family income-to poverty ratio; BMI, body mass index; CVD, cardiovascular disease; DM, diabetes mellitus; HbA1c, glycated hemoglobin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

Table 2. Association between DI-GM and NAFLD in multiple logistic regression analyses model

Variable	n. total	n. event_%	Crude model		Model 1†		Model 2‡		Model 3§	
			OR (95%CI)	*p* value	OR (95%CI)	*p* value	OR (95%CI)	*p* value	OR (95%CI)	*p* value
DI-GM	12910	4673 (36.2)	0.92 (0.90~0.94)	<0.001	0.88 (0.86~0.91)	<0.001	0.91 (0.88~0.94)	<0.001	0.92 (0.89~0.95)	<0.001
DI-GM group										
0-3	3083	1215 (39.4)	1 (Ref)		1 (Ref)		1 (Ref)		1 (Ref)	
4	3258	1219 (37.4)	0.92 (0.83~1.02)	0.103	0.87 (0.78~0.96)	0.006	0.88 (0.77~1)	0.053	0.91 (0.79~1.05)	0.210
5	3138	1155 (36.8)	0.9 (0.81~0.99)	0.035	0.81 (0.73~0.9)	<0.001	0.81 (0.71~0.93)	0.002	0.84 (0.73~0.98)	0.024
≥6	3431	1084 (31.6)	0.71 (0.64~0.79)	<0.001	0.61 (0.55~0.68)	<0.001	0.68 (0.59~0.77)	<0.001	0.71 (0.61~0.82)	<0.001
Trend test				<0.001		<0.001		<0.001		<0.001

OR, Odd Ratio; CI, Confidence interval; DI-GM, dietary index for gut microbiota; NAFLD, non-alcoholic fatty liver disease; PIR, family income-to poverty ratio; BMI, body mass index; CVD, cardiovascular disease; DM, diabetes mellitus; HbA1c, glycated hemoglobin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; HDL, high-density lipoprotein; LDL, low-density lipoprotein.
†Model 1: adjusted for Age, Sex.
‡Model 2: adjusted for Model 1+ Race, Marital status, PIR, Education, Physical activity, BMI, Smoke, Alcohol intake.
§Model 3: adjusted for Model 2+ HbA1c, ALT, AST, HDL, LDL, hypertension, CVD, DM.

Table 3. Association between DI-GM and NAFLD in multiple logistic regression analyses model (11,796 participants with complete data)

Variable	n. total	n. event_%	Crude model		Model 1†		Model 2‡		Model 3§	
			OR (95%CI)	*p* value	OR (95%CI)	*p* value	OR (95%CI)	*p* value	OR (95%CI)	*p* value
DI-GM	11,176	3869 (34.6)	0.92 (0.9~0.94)	<0.001	0.88 (0.86~0.91)	<0.001	0.89 (0.86~0.91)	<0.001	0.92 (0.89~0.96)	<0.001
DI-GM group										
0-3	2691	1010 (37.5)	1(Ref)		1(Ref)		1(Ref)		1(Ref)	
4	2802	1001 (35.7)	0.93 (0.83~1.03)	0.164	0.87 (0.78~0.96)	0.006	0.87 (0.78~0.97)	0.014	0.92 (0.78~1.07)	0.280
5	2697	955 (35.4)	0.91 (0.82~1.02)	0.106	0.81 (0.73~0.9)	<0.001	0.82 (0.73~0.92)	0.001	0.87 (0.74~1.02)	0.089
≥6	2986	903 (30.2)	0.72 (0.65~0.81)	<0.001	0.61 (0.55~0.68)	<0.001	0.62 (0.55~0.69)	<0.001	0.72 (0.62~0.85)	<0.001
Trend test				<0.001		<0.001		<0.001		<0.001

OR, Odd Ratio; CI, Confidence interval; DI-GM, dietary index for gut microbiota; NAFLD, non-alcoholic fatty liver disease; PIR, family income-to poverty ratio; BMI, body mass index; CVD, cardiovascular disease; DM, diabetes mellitus; HbA1c, glycated hemoglobin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

†Model 1: adjusted for Age, Sex.

‡Model 2: adjusted for Model 1+ Race, Marital status, PIR, Education, Physical activity, BMI, Smoke, Alcohol intake.

§Model 3: adjusted for Model 2+ HbA1c, ALT, AST, HDL, LDL, hypertension, CVD, DM.

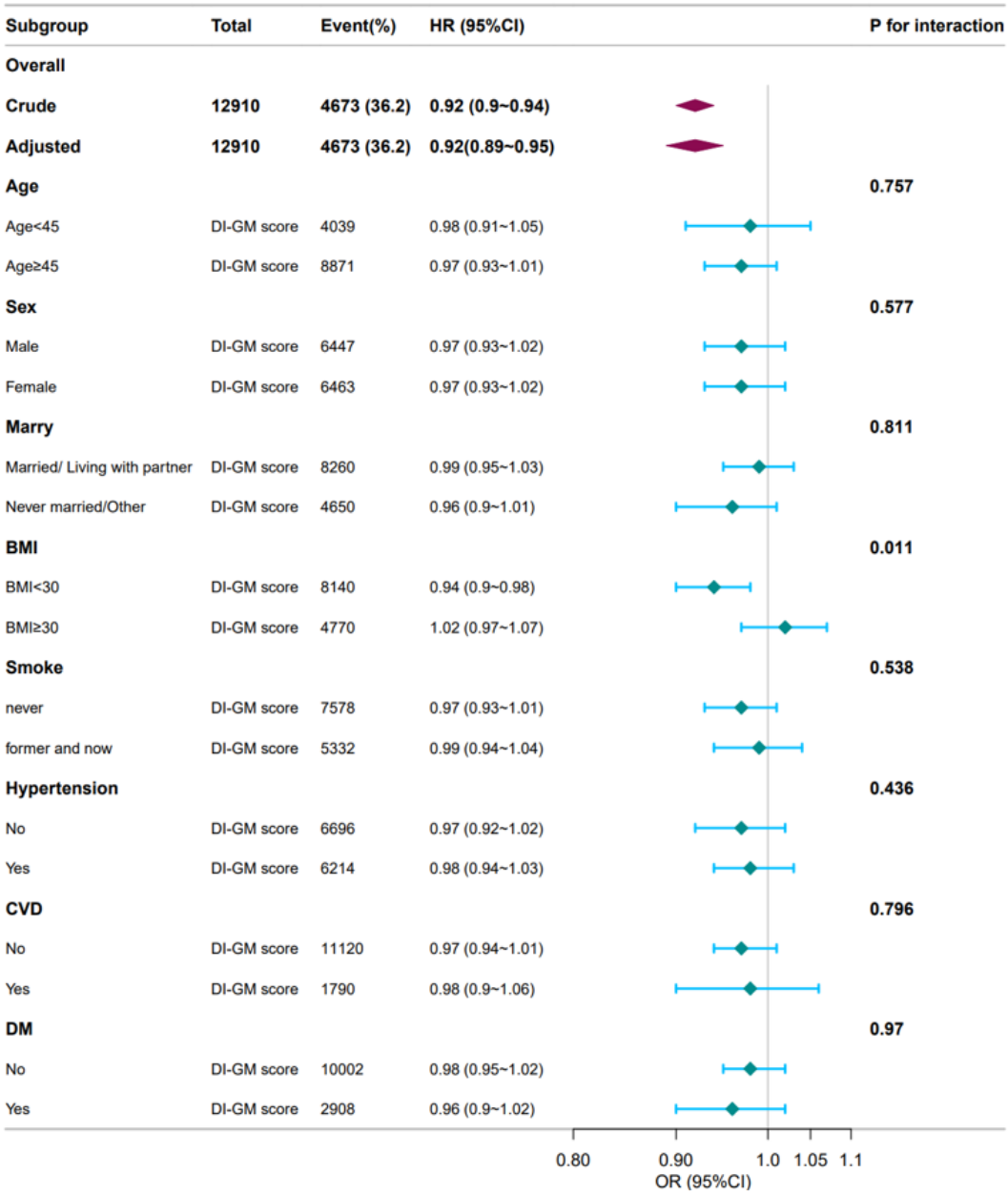

Figure 2. Association between DI-GM score and NAFLD in different subgroups. Adjusted for age, sex, race, marital status, PIR, education, Physical activity, BMI, smoke, alcohol intake, HbA1c, ALT, AST, HDL, LDL. CI, confidence interval, OR, odd ratio; DI-GM, dietary index for gut microbiota; NAFLD, non-alcoholic fatty liver disease; PIR, family income-to poverty ratio; BMI, body mass index; CVD, cardiovascular disease; DM, diabetes mellitus; HbA1c, glycated hemoglobin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; HDL, high-density lipoprotein; LDL, low-density lipoprotein

dysfunction, the impact of microbial metabolites, activation of inflammatory responses, modulation of the immune system, and alterations in bile acid (BA) metabolism.[49] Gut microbiota dysbiosis may compromise intestinal barrier function, enabling lipopolysaccharide (LPS), endotoxins, damage-associated molecular patterns to enter the circulation. Once in circulation, they trigger a signaling cascade in the liver, including the activation of toll-like receptor (TLR) and NLRP3 pathways, in turn stimulate the production of cytokines like tumor necrosis factor-alpha (TNF-α) and interleukin-1 beta (IL-1β). These inflammatory responses elevate serum-free fatty acid and triglyceride levels, leading to their accumulation in the liver and further inflammatory changes.[50] The gut microbiota produce endogenous ethanol, particularly when sugar rich foods are consumed,[51] which disrupt gut epithelial integrity and facilitate ethanol transport to the liver, thereby inducing oxidative stress and liver damage.[52] Gut microbiota dysbiosis can also contribute to the progression of NAFLD through modulating BA metabolism, including that reduced secondary BA synthesis impairs a G-protein-coupled BA receptor (TGR5)-dependent glucagon-like peptide-1 (GLP-1) secretion, exacerbating insulin resistance and lipogenesis;[53] Farnesoid X receptor (FXR) signaling suppression fibroblast growth factor (FGF)-19 secretion, upregulating hepatic Cholesterol 7α-hydroxylase (CYP7A1) activity and driving unchecked BA synthesis;[52,54,55] increased hydrophobic BA accumulation directly impairs hepatocyte mitochondrial function through oxidative stress induction and apoptotic pathway activation.[56]

Our study demonstrates that a higher DI-GM is associated with lower NAFLD prevalence, which has significant clinical implications. First, the findings suggest that improving DI-GM by increasing dietary fiber and fermented dairy product intake can boost gut microbiota diversity, potentially reducing NAFLD risk. Second, DI-GM can serve as a marker of dietary quality, aiding clinicians in designing more effective dietary intervention strategies, offer a promising approach for NAFLD prevention and treatment.

Based on our understanding, this research pioneers the exploration of the correlation between DI-GM, a metric of dietary patterns influencing gut microbiota diversity, and NAFLD. The stringent quality control measures and advanced sampling methodologies employed by NHANES in data acquisition have enabled us to assess the relationship within a substantial and varied cohort of United States adults. Furthermore, sensitivity analyses such as subgroup analyses enhanced the robustness and reliability of the findings. This study has several limitations. First, its cross-sectional design prevents the establishment of a causal link between DI-GM and NAFLD. Additional prospective studies and randomized controlled trials are necessary to confirm causality. Secondly, as in many studies, the possibility of confounding effects due to measurement error residuals from unmeasured variables or unknown confounders cannot be completely ruled out. Thirdly, while the DI-GM includes 14 types of food, the NHANES dietary data did not capture specific tea intake, omitting this from our analysis. Lastly, the DI-GM scores were determined from self-reported 24-hour dietary records, potentially introducing recall bias, and some covariates relied on self-reporting as well.

Conclusion

Our study revealed a robust association between DI-GM and the prevalence of NAFLD in a cross-sectional analysis. Given the strong association between diet, microbiota and NAFLD, future research and dietary interventions incorporating the DI-GM for individuals with NAFLD will be crucial in preventing and treating NAFLD.

ACKNOWLEDGEMENTS

The authors thank the NHANES database for providing the data source for this study.

AUTHOR DISCLOSURES

Yan Xue and Jianxian Zhang declare that they have no conflicts of interest. This study received no specific financial support.

REFERENCES

1. Li J, Zou B, Yeo YH, Feng Y, Xie X, Lee DH, et al. Prevalence, incidence, and outcome of non-alcoholic fatty liver disease in Asia, 1999–2019: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol. 2019;4:389–98. doi: 10.1016/S2468-1253(19)30039-1.
2. Araújo AR, Rosso N, Bedogni G, Tiribelli C, Bellentani S. Global epidemiology of non-alcoholic fatty liver disease/non-alcoholic steatohepatitis: what we need in the future. Liver Int: Off J Int Assoc Study Liver. 2018;38:47–51. doi: 10.1111/liv.13643.
3. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-meta-analytic assessment of prevalence, incidence, and outcomes. Hepatol (Baltim Md,). 2016;64:73–84. doi: 10.1002/hep.28431.
4. Zou B, Yeo YH, Nguyen VH, Cheung R, Ingelsson E, Nguyen MH. Prevalence, characteristics and mortality outcomes of obese, nonobese and lean NAFLD in the United States, 1999–2016. J Intern Med. 2020;288:139–51. doi: 10.1111/joim.13069.
5. Mahady SE, George J. Predicting the future burden of NAFLD and NASH. J Hepatol. 2018;69:774–5. doi: 10.1016/j.jhep.2018.06.025.
6. Younossi ZM, Stepanova M, Younossi Y, Golabi P, Mishra A, Rafiq N, et al. Epidemiology of chronic liver diseases in the USA in the past three decades. Gut. 2020;69:564–8. doi: 10.1136/gutjnl-2019-318813.
7. Rinaldi L, Pafundi PC, Galiero R, Caturano A, Morone MV, Silvestri C, et al. Mechanisms of non-alcoholic fatty liver disease in the metabolic syndrome. A narrative review. Antioxid (Basel Switz). Antioxidants (Basel); 2021;10. doi: 10.3390/antiox10020270.
8. Shaaban HH, Alzaim I, El-Mallah A, Aly RG, El-Yazbi AF, Wahid A. Metformin, pioglitazone, dapagliflozin and their combinations ameliorate manifestations associated with NAFLD in rats via anti-inflammatory, anti-fibrotic, anti-oxidant and anti-apoptotic mechanisms. Life Sci. 2022;308:120956. doi: 10.1016/j.lfs.2022.120956.
9. Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. Nat Med. 2018;24:908–22. doi: 10.1038/s41591-018-0104-9.
10. Losno EA, Sieferle K, Perez-Cueto FJA, Ritz C. Vegan Diet and the Gut Microbiota Composition in Healthy Adults. Nutrients. 2021;13:2402. doi: 10.3390/nu13072402.
11. Kase BE, Liese AD, Zhang J, Murphy EA, Zhao L, Steck SE. The Development and Evaluation of a Literature-Based Dietary Index for Gut Microbiota. Nutrients. 2024;16:1045. doi: 10.3390/nu16071045.
12. Ore A, Akinloye OA. Phytotherapy as multi-hit therapy to confront the multiple pathophysiology in non-alcoholic fatty liver disease: a systematic review of experimental interventions. Med (Kaunas Lith). 2021;57:822. doi: 10.3390/medicina57080822.
13. Wiest R, Albillos A, Trauner M, Bajaj JS, Jalan R. Targeting the gut-liver axis in liver disease. J Hepatol. 2017;67:1084–103. doi: 10.1016/j.jhep.2017.05.007.
14. Hsu CL, Schnabl B. The gut-liver axis and gut microbiota in health and liver disease. Nat Rev Microbiol. 2023;21:719–33. doi: 10.1038/s41579-023-00904-3.
15. Li F, Ye J, Shao C, Zhong B. Compositional alterations of gut microbiota in nonalcoholic fatty liver disease patients: a systematic review and Meta-analysis. Lipids Health Dis. 2021;20:22. doi: 10.1186/s12944-021-01440-w.
16. Behrouz V, Aryaeian N, Zahedi MJ, Jazayeri S. Effects of probiotic and prebiotic supplementation on metabolic parameters, liver aminotransferases, and systemic inflammation in nonalcoholic fatty liver disease: A randomized clinical trial. J Food Sci. 2020;85:3611–7. doi: 10.1111/1750-3841.15367.
17. Liu C, Wang Y-L, Yang Y-Y, Zhang N-P, Niu C, Shen X-Z, et al. Novel approaches to intervene gut microbiota in the treatment of chronic liver diseases. FASEB j: off publ Fed Am Soc Exp Biol. 2021;35:e21871.doi: 10.1096/fj.202100939R.
18. Craven L, Rahman A, Nair Parvathy S, Beaton M, Silverman J, Qumosani K, et al. Allogenic Fecal Microbiota Transplantation in Patients With Nonalcoholic Fatty Liver Disease Improves Abnormal Small Intestinal Permeability:

A Randomized Control Trial. Am J Gastroenterol. 2020;115:1055–65. doi: 10.14309/ajg.0000000000000661.
19. O'Keefe SJD, Li JV, Lahti L, Ou J, Carbonero F, Mohammed K, et al. Fat, fibre and cancer risk in African Americans and rural Africans. Nat Commun. 2015;6:6342. doi: 10.1038/ncomms7342.
20. Maskarinec G, Hullar MAJ, Monroe KR, Shepherd JA, Hunt J, Randolph TW, et al. Fecal microbial diversity and structure are associated with diet quality in the multiethnic cohort adiposity phenotype study. J Nutr. 2019;149:1575–84. doi: 10.1093/jn/nxz065.
21. Reddel S, Putignani L, Del Chierico F. The Impact of Low-FODMAPs, Gluten-Free, and Ketogenic Diets on Gut Microbiota Modulation in Pathological Conditions. Nutrients. 2019;11:373. doi: 10.3390/nu11020373.
22. Zipf G, Chiappa M, Porter KS, Ostchega Y, Lewis BG, Dostal J. National health and nutrition examination survey: plan and operations, 1999-2010. Vital Health Stat, 1 Programs Collect Proced. 2013:1–37.
23. Ruhl CE, Everhart JE. Fatty liver indices in the multiethnic United States National Health and Nutrition Examination Survey. Aliment Pharmacol Ther. 2015;41:65–76. doi: 10.1111/apt.13012.
24. Zelber-Sagi S, Webb M, Assy N, Blendis L, Yeshua H, Leshno M, et al. Comparison of fatty liver index with noninvasive methods for steatosis detection and quantification. World J Gastroenterol. 2013;19:57–64. doi: 10.3748/wjg.v19.i1.57.
25. Chen L-W, Huang P-R, Chien C-H, Lin C-L, Chien R-N. A community-based study on the application of fatty liver index in screening subjects with nonalcoholic fatty liver disease. J Formos Med Assoc. 2020;119:173–81. doi: 10.1016/j.jfma.2019.03.016.
26. Carvalhana S, Leitão J, Alves AC, Bourbon M, Cortez-Pinto H. How good is controlled attenuation parameter and fatty liver index for assessing liver steatosis in general population: correlation with ultrasound. Liver International. 2014;34. doi: 10.1111/liv.12305.
27. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the american gastroenterological association, american association for the study of liver diseases, and american college of gastroenterology. Gastroenterology. 2012;142:1592–609. doi: 10.1053/j.gastro.2012.04.001.
28. Kim D, Kim W, Adejumo AC, Cholankeril G, Tighe SP, Wong RJ, et al. Race/ethnicity-based temporal changes in prevalence of NAFLD-related advanced fibrosis in the United States, 2005-2016. Hepatol Int. 2019;13:205–13. doi: 10.1007/s12072-018-09926-z.
29. Motamed N, Maadi M, Sohrabi M, Keyvani H, Poustchi H, Zamani F. Rural Residency has a Protective Effect and Marriage is a Risk Factor for NAFLD. Hepat Mon. 2016;16. doi: 10.5812/hepatmon.38357.
30. Fan C, He Y, Yang J, Da M. Association Between Live Microbe Intake and NAFLD: Evidence From NHANES 2003-2018. J Am Nutr Assoc. 2024;43:272–8. doi: 10.1080/27697061.2023.2270537.
31. Zhang X, Yang Q, Huang J, Lin H, Luo N, Tang H. Association of the newly proposed dietary index for gut microbiota and depression: the mediation effect of phenotypic age and body mass index. Eur Arch Psychiatry Clin Neurosci. 2024; doi: 10.1007/s00406-024-01912-x.
32. Vilar-Gomez E, Nephew LD, Vuppalanchi R, Gawrieh S, Mladenovic A, Pike F, et al. High-quality diet, physical activity, and college education are associated with low risk of NAFLD among the US population. Hepatology. 2022;75:1491–506. doi: 10.1002/hep.32207.
33. Younossi Z, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. Nat Rev Gastroenterol Hepatol. 2018;15:11–20. doi: 10.1038/nrgastro.2017.109.
34. Agricultural Research Service, US Department of Agriculture. What we eat in America data tables. Available at: https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/foodsurveys-research-group/docs/wweia-data-tables/.
35. Yun L, Vanderloo LM, Berry TR, Latimer-Cheung AE, O'Reilly N, Rhodes RE, et al. Political Orientation and Public Attributions for the Causes and Solutions of Physical Inactivity in Canada: Implications for Policy Support. Front Public Health. 2019;7:153. doi: 10.3389/fpubh.2019.00153.
36. Ruan Z, Lu T, Chen Y, Yuan M, Yu H, Liu R, et al. Association between psoriasis and nonalcoholic fatty liver disease among outpatient US adults. JAMA Dermatol. 2022;158:745. doi: 10.1001/jamadermatol.2022.1609.
37. Rattan P, Penrice DD, Ahn JC, Ferrer A, Patnaik M, Shah VH, et al. Inverse association of telomere length with liver disease and mortality in the US population. Hepatol Commun. 2022;6:399–410. doi: 10.1002/hep4.1803.
38. Zheng Y, Wang J, Wang Y, Xu K, Chen X. The Hidden Dangers of Plant-Based Diets Affecting Bone Health: A Cross-Sectional Study with U.S. National Health and Nutrition Examination Survey (NHANES) Data from 2005-2018. Nutrients. 2023;15:1794. doi: 10.3390/nu15071794.
39. American Diabetes Association. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43:S14–31. doi: 10.2337/dc20-S002.
40. Bernard D, Doumard E, Ader I, Kemoun P, Pagès J, Galinier A, et al. Explainable machine learning framework to predict personalized physiological aging. Aging Cell. 2023;22:e13872. doi: 10.1111/acel.13872.
41. Yun E-J, Imdad S, Jang J, Park J, So B, Kim J-H, et al. Diet Is a Stronger Covariate than Exercise in Determining Gut Microbial Richness and Diversity. Nutrients. 2022;14:2507. doi: 10.3390/nu14122507.
42. Valdes AM, Walter J, Segal E, Spector TD. Role of the gut microbiota in nutrition and health. BMJ. 2018;k2179. doi: 10.1136/bmj.k2179.
43. González Olmo BM, Butler MJ, Barrientos RM. Evolution of the Human Diet and Its Impact on Gut Microbiota, Immune Responses, and Brain Health. Nutrients. 2021;13:196. doi: 10.3390/nu13010196.
44. Stiemsma LT, Nakamura RE, Nguyen JG, Michels KB. Does consumption of fermented foods modify the human gut microbiota? J Nutr. 2020;150:1680–92. doi: 10.1093/jn/nxaa077.
45. Wastyk HC, Fragiadakis GK, Perelman D, Dahan D, Merrill BD, Yu FB, et al. Gut microbiota-targeted diets modulate human immune status. Cell. 2021; 184: 4137–53. doi: 10.1016/j.cell.2021.06.019.
46. Lee G, You HJ, Bajaj JS, Joo SK, Yu J, Park S, et al. Distinct signatures of gut microbiome and metabolites associated with significant fibrosis in non-obese NAFLD. Nat Commun. 2020;11. doi: 10.1038/s41467-020-18754-5.
47. Wang L, Xu J, You N, Shao L, Zhuang Z, Zhuo L, et al. Characteristics of intestinal flora in nonobese nonalcoholic fatty liver disease patients and the impact of ursodeoxycholic acid treatment on these features. Lipids. 2024;59:193–207. doi: 10.1002/lipd.12410.

48. Philips CA, Ahamed R, Rajesh S, Abduljaleel JKP, Augustine P. Long-term Outcomes of Stool Transplant in Alcohol-associated Hepatitis-Analysis of Clinical Outcomes, Relapse, Gut Microbiota and Comparisons with Standard Care. J Clin Exp Hepatol. 2022;12:1124–32. doi: 10.1016/j.jceh.2022.01.001.
49. Rahman K, Desai C, Iyer SS, Thorn NE, Kumar P, Liu Y, et al. Loss of Junctional Adhesion Molecule A Promotes Severe Steatohepatitis in Mice on a Diet High in Saturated Fat, Fructose, and Cholesterol. Gastroenterology. 2016;151(4):733-746.e12. doi: 10.1053/j.gastro.2016.06.022.
50. Rennert C, Heil T, Schicht G, Stilkerich A, Seidemann L, Kegel-Hübner V, et al. Prolonged lipid accumulation in cultured primary human hepatocytes rather leads to ER stress than oxidative stress. Int J Mol Sci. 2020;21:7097. doi: 10.3390/ijms21197097.
51. Zhong S, Li L, Liang N, Zhang L, Xu X, Chen S, et al. Acetaldehyde Dehydrogenase 2 regulates HMG-CoA reductase stability and cholesterol synthesis in the liver. RedoxBiol.2021;41:101919.doi:10.1016/j.redox.2021.101919.
52. Haslam DB. Nonalcoholic steatohepatitis and the intestinal microbiota.Hepatology.2017;65:4013.doi:10.1002/hep.28864.
53. Liu J, Wu A, Cai J, She Z-G, Li H. The contribution of the gut-liver axis to the immune signaling pathway of NAFLD. Front Immunol.2022;13:968799.doi:10.3389/fimmu.2022.968799.
54. Schubert K, Olde Damink SWM, Von Bergen M, Schaap FG. Interactions between bile salts, gut microbiota, and hepatic innate immunity. Immunol Rev. 2017;279:23–35. doi: 10.1111/imr.12579.
55. Marra F, Svegliati-Baroni G. Lipotoxicity and the gut-liver axis in NASH pathogenesis. J Hepatol. 2018;68:280–95. doi: 10.1016/j.jhep.2017.11.014.
56. Li Z, Yuan H, Chu H, Yang L. The crosstalk between gut microbiota and bile acids promotes the development of non-alcoholic fatty liver disease. Microorganisms. 2023;11:2059. doi: 10.3390/microorganisms11082059.

Supplementary figure

Component	Included Foods within the Component	Scoring
Beneficial to gut microbiota		
Avocados	Avocados	For each component, a score of 1 if consumption at or above the sex-specific median, else 0
Broccoli	Broccoli	
Chickpea	Chickpeas	
Coffee	Coffee	
Cranberries	Cranberries	
Fermented dairy	Yogurt, cheese, kefir, sour cream, buttermilk	
Fiber	Not applicable	
Green tea	Green tea	
Soybean	Soy products--Soy milk, Tofu	
Whole grains	Grains defined as whole grains, containing the entire grain kernel—the bran, germ, and endosperm	
Unfavorable to gut microbiota		
High-fat diet (% energy)	Not applicable	0 if consumption at or above 40% energy from fat, else 1 For each remaining component, a score of 0 if consumption at or above the sex-specific median, else 1
Processed meat	Frankfurters, sausages, corned beef, and luncheon meat that are made from beef, pork, or poultry	
Red meat	Beef, veal, pork, lamb, and game meat; excludes organ meat and cured meat	
Refined grains	Refined grains that do not contain all of the components of the entire grain kernel	

Supplementary Figure 1. Components and the scoring criteria for the DI-GM