Accepted Article

DR. EDUARDO VILAR-GOMEZ (Orcid ID : 0000-0003-1435-4013)

DR. LAUREN NEPHEW (Orcid ID : 0000-0003-0837-0746)

DR. RAJ VUPPALANCHI (Orcid ID : 0000-0003-0637-1577)

DR. SAMER GAWRIEH (Orcid ID : 0000-0002-2056-4909)

DR. NAGA CHALASANI (Orcid ID : 0000-0003-4082-3178)

Article type : Original Article

High quality diet, physical activity and college education are associated with low risk of NAFLD among the U.S. population.

[1]EduardoVilar-Gomez, [1]Lauren D. Nephew, [1]Raj Vuppalanchi, [1]Samer Gawrieh, [1]Andrea Mladenovic, [2]Francis Pike, [1]Niharika Samala, [1]Naga Chalasani.

[1]Department of Medicine, Indiana University School of Medicine, Indianapolis, IN.

[2]Department of Biostatistics, Indiana University School of Medicine, Indianapolis, IN.

Keywords:

Vibration-controlled transient elastography; clinically significant fibrosis; healthy eating index; socio-economic status; physical activity.

Corresponding Author:

This is the author's manuscript of the article published in final edited form as:
Vilar-Gomez, E., Nephew, L. D., Vuppalanchi, R., Gawrieh, S., Mladenovic, A., Pike, F., Samala, N., & Chalasani, N. (2021). High quality diet, physical activity and college education are associated with low risk of NAFLD among the U.S. population. Hepatology. https://doi.org/10.1002/hep.32207

Naga Chalasani, MD., FAASLD
Indiana University School of Medicine, 702 Rotary Circle, Suite 225,
Indianapolis, IN 46202
E-mail: nchalasa@iu.edu
Fax: + 1-317-278-1949

Abbreviations:

NAFLD, non-alcoholic fatty liver disease; **T2DM**, type 2 diabetes mellitus; **PA**, physical activity; **DQ**, diet quality; **HEI**, healthy eating index; **DGA**, dietary guidelines for Americans; **SES**, socio-economic status; **NHANES**, National Health and Nutrition Examination Surveys; **NCHS**, National Center for Health Statistics; **MEC**, mobile examination center; **NH**, non-Hispanic; **VCTE**, vibration-controlled transient elastography; **EA**, educational attainment; **CAP**, controlled attenuation parameter; **LSM**, liver stiffness measurement; **GPAQ**, global physical activity questionnaire; **PIR**, poverty to income ratio; **DM**, diabetes mellitus; **BP**, blood pressure; **HbA1c**; glycosylated hemoglobin; **BMI**, body mass index; **STROBE**, The Strengthening the Reporting of Observational Studies in Epidemiology; **CSF**, clinically significant fibrosis.

Financial support:

This study was funded by departmental internal funding and did not influence the study's design, conduct, or reporting.

Conflicts of Interests:

There are none for this paper.
For full disclosure:
Dr. Chalasani has ongoing paid consulting activities (or had in the preceding 12 months) with Abbvie, Madrigal, Foresite labs, Altimmune, Zydus, Galectin, and Boehringer-Ingelheim. Dr. Chalasani receives research grant support from Exact Sciences and DSM. where his institution receives the funding. Dr. Chalasani has equity ownership in RestUp, Inc.

This article is protected by copyright. All rights reserved

Accepted Article

Dr. Gawrieh discloses consulting for TransMedics, Pfizer and receives research grant support from Zydus, Galmed, Viking and Sonic Incytes.

Drs. Vilar-Gomez, Nephew, Vuppalanchi, Mladenovic, Pike and Samala have nothing to disclose.

Author Contributions

All authors made substantial contributions to the intellectual content of the paper and approved the final version of the manuscript.

Conception and design: Chalasani, Vilar-Gomez.

Acquisition of data: Vilar-Gomez, Pike.

Analysis and interpretation of data: Vilar-Gomez, Pike, Nephew, Chalasani.

Drafting of the manuscript: Vilar-Gomez, Nephew, Chalasani.

Critical revision of the manuscript for intellectual content: Chalasani, Vilar-Gomez, Nephew, Vuppalanchi, Gawrieh, Mladenovic, Samala, Pike.

Statistical analysis: Vilar-Gomez, Pike.

Obtaining funding: Chalasani.

Supervision: Chalasani.

Naga Chalasani, MD had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Authors Emails

Eduardo Vilar-Gomez: evilar@iu.edu

Lauren Nephew: lnephew@iu.edu

Raj Vuppalanchi: rvuppala@iu.edu

Samer Gawrieh: sgawrieh@iu.edu

Andrea Mladenovic: amladeno@iu.edu

Francis Pike: fpike@iu.edu

Niharika Samala: nrsamala@iu.edu

Naga Chalasani: nchalasa@iu.edu

Acknowledgments

This article is protected by copyright. All rights reserved

Accepted Article

The authors thank the members of the National Center for Health Statistics, Centers for Disease Control and Prevention, for gathering the data and making it available for public use. The authors also thank the participants involved in the survey.

Accepted Article

Abstract

Background & Aims

The effects of diet quality (DQ), physical activity (PA), and socioeconomic status (SES) on the risk of nonalcoholic fatty liver disease (NAFLD) are unclear. We examined the association between DQ, PA, SES, and NAFLD risk.

Approach & Results

This is a cross-sectional analysis of the National Health and Nutrition Examination Surveys, 2017-2018, which included 3589 participants with reliable information on vibration-controlled transient elastography (VCTE) measurements, 24-h dietary recalls, PA, and SES. DQ was assessed by the Healthy Eating Index (HEI)-2015. PA was determined by the Global Physical Activity Questionnaire.

This article is protected by copyright. All rights reserved

SES was assessed by the educational attainment and family poverty to income ratio (PIR). Risk of NAFLD was considered by means of a composite outcome using VCTE measurements: non-NAFLD *vs* NAFLD without clinically significant fibrosis (CSF) *vs* NAFLD with CSF. The NAFLD risk was lower in physically active (≥600 MET min/week) vs. inactive participants (<600 MET min/week) (OR:0.71, P=0.043). A high-quality diet (HQD) (HEI>56.64) was associated with a lower risk of NAFLD (OR: 0.58, P<0.01) compared with a non-HQD. The lowest NAFLD risk was observed in those physically active with HQD (OR: 0.43, P<0.01). Body mass index (BMI) and waist circumference (WC) significantly mediated the effect of DQ and PA on NAFLD risk. Education (college or above) (OR: 0.65, P=0.034), but not PIR was associated with a reduced NAFLD risk. HQD and increased PA partially mediated the effect of education on NAFLD risk. The total effect of education on NAFLD risk mediated by DQ was 29% and by PA was 8%.

Conclusions

High-quality diet, increased physical activity, and college education were associated with lower NAFLD risk in the U.S. population.

Accepted Article

This article is protected by copyright. All rights reserved

Accepted Article

Introduction

In the U.S., approximately one-third of adults have nonalcoholic fatty liver disease (NAFLD), and the prevalence continues to grow in tandem obesity and type 2 diabetes mellitus (T2DM) epidemic.(1) Unhealthy lifestyles, including unhealthy diets and decreased time spent on physical activity (PA), have been strongly linked to the risk of NAFLD.(2, 3) Lifestyle interventions, which include dietary modifications and increased PA may be effective in the management of NAFLD.(4, 5)

Recent epidemiologic studies have stressed the importance of dietary patterns rather than single nutrients or foods to evaluate the association between diet and health outcomes.(6, 7) Diet is a complex mixture of nutrients and foods that interact with each other. The dietary pattern approaches consider contributions of various aspects of diet, including nutrient and food quality; therefore, they could reflect real-world dietary practices and integrate potentially interactive and cumulative associations of several food components.(8) In addition, dietary patterns more closely simulate real-world scenarios of nutrient and food combinations, which facilitates the translation of clinical findings into dietary recommendations. In the recent years, several dietary scores have been developed to explore the association between diet quality (DQ) and health outcomes. Healthier DQ, as determined by means of different dietary scores (Dietary Approaches to Stop Hypertension, Alternate Mediterranean Diet score, and Alternate Healthy Eating Index–2010), has been shown to reduce the risk of all-cause mortality, cardiovascular disease, cancer, and T2DM by approximately 22%, 22%, 15%, and 22%, respectively.(9, 10)

Most recently, a new iteration of the Healthy Eating Index (HEI) 2015 (11) has been reported to be strongly associated with a reduced risk of all-cause, cardiovascular, and cancer mortality in two large multiethnic cohorts.(12, 13). Given such consistent data on dietary patterns and health outcomes, the 2015-2020 Dietary Guidelines for Americans (DGA) acknowledge the importance of diet quality rather than the effect of single nutrients on health and recommend multiple healthy dietary patterns to provide dietary choices for all Americans with diverse cultural and personal food traditions or preferences.(14)

Aside from dietary recommendations, the Physical Activity Guidelines for Americans recommend that adults get at least 150 minutes of moderate-intensity aerobic PA or 75 minutes of vigorous-intensity PA, or an equivalent combination each week as it is associated with positive health

This article is protected by copyright. All rights reserved

outcomes.(15) It has been documented that increased PA is associated not only with a lower risk of NAFLD but also with reduced all-cause and cardiovascular mortality in individuals with NAFLD in the U.S. population.(3, 16)

Socioeconomic disparity in nutrition and PA is well documented (17), which explains some of the observed social inequalities in health. People with high socioeconomic status (SES) tend to be more physically active and have healthier dietary habits, hence contributing to their better health status.(18) Thus, evaluation of the relationships between DQ, PA, and SES in the context of NAFLD risk is necessary.

Data on the relationship between adherence to healthy dietary patterns, PA, SES, and the risk of NAFLD in the U.S. population are sparse. Thus, the aim of this study was to examine the association between DQ, PA, and the risk of NAFLD and its severity and identify relevant socioeconomic factors that may influence dietary and PA patterns.

Material and Methods

Study design and participants

This study focuses on data from the National Health and Nutrition Examination Surveys (NHANES), administered by the U.S. National Center for Health Statistics (NCHS), between the years 2017-2018. Using a stratified, multistage, and probability-cluster design, the NHANES collected data representative of the noninstitutionalized civilian U.S. persons aged 1 to 80 years. Demographic characteristics, clinical history, and self-reported dietary and PA information were collected from participants via a structured household interview. Physical examinations, including anthropometric measurements and blood samples, were collected within mobile examination center (MEC). Detailed descriptions of the NHANES 2017-2018 data collection, analytical guidelines, and full datasets are publicly available elsewhere.(19, 20) The NHANES (2017-2018) included oversampling of selected populations, including Asian, Hispanic, non-Hispanic (NH) Black, and older adults, to increase statistics precision for meaningful subgroup analyses and more reliable variance estimates.(19)

Of the 9,254 participants included in the NHANES (2017-2018), 6401 aged 12 years and over were eligible for hepatic vibration-controlled transient elastography (VCTE) examinations. The

This article is protected by copyright. All rights reserved

elastography measurements were obtained in the NHANES MEC by an experienced technician, using the FibroScan® model 502 V2 Touch equipped with a medium (M) or extra-large (XL) wand (probe). Further information on quality assurance and quality control measures considered for VCTE procedures is publicly available elsewhere.(21)

For this cross-sectional analysis, adults aged 18 years or older were included. We excluded those meeting the following criteria: (1) elastography examination status ineligible (n=258), not performed (n=156) or partial (n=493), (2) younger than 18 years (n=748), (3) serologic positivity for viral hepatitis B (n=28) or C (n=42), (4) consumption of steatogenic medications for at least 3 months or more before study enrollment (n=65), (5) reporting significant alcohol intake defined as more than two or three standard drinks per day on average for both women and men, respectively (n=393) (1), and missing values on dietary information (n=359). The **Appendix Figure 1** shows the flow of patients through the study.

Our final study population consisted of 3589 participants who have complete and reliable information about race/ethnicity, education attainment (EA), self-reported 24-hour dietary recalls, alcohol consumption, PA, smoking status, and physical examination findings (height, weight, and waist circumference).

Definition and assessment of hepatic steatosis and clinically significant fibrosis (CSF)

In this study, we selected those CAP (controlled-attenuation parameter) and LSM (liver stiffness measurement) cutoffs that have been developed and validated in the U.S. population for detecting individuals with hepatic steatosis or CSF (stages of fibrosis ≥2).(22) An optimal CAP cutoff of ≥285 dB/m (sensitivity of 80% and specificity of 77%) is indicative of hepatic steatosis whereas an optimal LSM cutoff of ≥8.6 kPa (sensitivity of 66% and specificity of 80%) is suggestive of CSF.

A composite categorical (ordered) outcome, including (1) *non-NAFLD* (CAP <285 dB/m) vs (2) *NAFLD* (CAP ≥285 dB/m) *without CSF* (LSM <8.6 kPa) vs (3) *NAFLD* (CAP ≥285 dB/m) *with CSF* (LSM ≥8.6 kPa) was considered our primary outcome of interest. Two further composite secondary ordinal outcomes using LSM cutoffs at 90% sensitivity (≥5.6 kPa) and 90% specificity (≥6.5 kPa) for excluding or diagnosing CSF, respectively were defined as follows: ***LSM ≥5.6 kPa -*** (1) *non-NAFLD* (CAP <285 dB/m) vs (2) *NAFLD* (CAP ≥285 dB/m) *without CSF* (LSM <5.6 kPa) vs

This article is protected by copyright. All rights reserved

(3) *NAFLD* (CAP ≥285 dB/m) *with CSF* (LSM ≥5.6 kPa); ***LSM≥6.5 kPa -*** (1) *non-NAFLD* (CAP <285 dB/m) vs (2) *NAFLD* (CAP ≥285 dB/m) *without CSF* (LSM <6.5 kPa) vs (3) *NAFLD* (CAP ≥285 dB/m) *with CSF* (LSM ≥6.5 kPa).

Dietary Assessment and HEI-2015 calculation

Dietary intakes used in these analyses were obtained using data from the two 24-hour dietary recalls. (23) We used the latest iteration of the HEI (2015), which was created to assess alignment with the 2015–2020 U.S. DGA, and to examine relationships between DQ and health-related outcomes.(4) It consists of 13 components that are scored based on energy-adjusted food and nutrient intakes. Of the 13 components, 9 assess dietary adequacy (total fruits, whole fruits, total vegetables, greens and beans, whole grains, dairy, total protein foods, and fatty acids) and 4 assess moderation (refined grains, sodium, added sugar and saturated fats).(11) For the adequacy components, higher scores are associated with higher levels of intake, whereas for the moderation components, higher scores are linked to lower levels of consumption. Seven components are each scored on a 0 to 10 scale, and the six other components are each scored on a 0 to 5 scale. The HEI components are summed to obtain the total score with a maximum of 100 points; higher scores indicate better DQ and therefore closer compliance with 2015-2020 DGA recommendations. Further information is provided in the Appendix.

Physical activity assessment

Information on PA in NHANES has been self-reported by participants using the Global Physical Activity Questionnaire (GPAQ). It is a validated instrument for PA surveillance.(24) In order to explore dose-response effects of PA on risk of NAFLD, we created the following categories of total PA: <600, 600-7999 and ≥8000 MET minutes/week, based on previously published data.(25) We further classified PA into 2 levels: active (≥600 MET-minutes/week) and inactive (<600 MET-minutes/week). PA levels of ≥600 MET-minutes/week have been consistently associated with substantial health benefits.(15) A more detailed description about the GPAQ and how the PA was calculated is stated in the Appendix.

Socioeconomic status (SES) assessment

This article is protected by copyright. All rights reserved

For each participant, self-reported education and household incomes were used as indicators of SES. Educational attainment (less than high school graduate, high school graduate or GED, some college, or college graduate or above) and house income defined by the poverty income ratio (PIR), which is the ratio of family income to the federal poverty level (low income, PIR <1.30; middle income, PIR = 1.30-3.49; and high income, PIR ≥ 3.50). The PIR cutoff for participation in the Supplemental Nutrition Assistance Program is 1.30, and 3.50 provides relatively equal sample sizes for each PIR group.

Other covariates

Demographic variables included sex and age; race/ethnicity were categorized as NH-White, NH-Black, NH-Asian, and other (including multiple races). Mexican Americans and other Hispanic participants were combined to create the Hispanic group.

Diabetes mellitus was defined by a self-reported previous diagnosis, a glycosylated hemoglobin A1c (HbA1c) level of ≥6.5%, or a fasting plasma glucose level of ≥126 mg/dL. Hypertension was defined based on previous physician diagnosis, taking prescribed medicine to decrease blood pressure (BP), or BP of ≥140/90 mmHg. The body mass index (BMI) was calculated as the weight in kilograms divided by the height in meters squared. Smoking status was categorized as never, former, or current smokers. Participants were defined as former smokers if they smoked at least 100 cigarettes in one's lifetime, but they were not currently smoking cigarettes. Blood lipids and liver function tests were measured enzymatically; HbA1c was measured by high-performance liquid chromatography.

Details of the questionnaires, examination components and laboratory procedures can be found in the NHANES reference manuals and report.(20, 21)

The NHANES study was approved by the NCHS Institutional Review Board, and all participants provided written informed consent. For this analysis, we used publicly available data without personal identifiable information. The methods and results are in accordance with relevant NCHS regulations. (19)

Statistical analysis

This article is protected by copyright. All rights reserved

We summarized baseline characteristics of participants by three categories of NAFLD phenotypes (non-NAFLD vs NAFLD without CSF vs NAFLD with CSF), HEI scores tertiles (<42.05 [low-quality diet] vs 42.05-56.64 [borderline-quality diet] vs >56.64 [high-quality diet]) and PA (active vs inactive). Categorical variables were summarized as weighted percentages, and continuous variables as weighted means, with their confidence intervals (CIs). We calculated age-standardized prevalence estimates and 95% CIs for each category of NAFLD phenotypes. Differences in means or percentages of baseline characteristics among NAFLD phenotypes categories, HEI score tertiles or PA groups were tested using unadjusted linear or binary (two categories) or ordinal (three or more ordered categories) logistic regressions, respectively.

Ordinal logistic regressions were used to calculate ORs and 95% Cis of either HEI, PA, EA or PIR with risks across the NAFLD phenotypes categories. Multivariable ordinal logistic regression analyses were adjusted for potential confounders: age, sex, race/ethnicity, total energy intake, HEI, PA, smoking status, EA, PIR, and alcohol intake status. Participants with missing data on PIR (n=454) were excluded in the corresponding subgroup analyses and multivariable analysis. We performed interaction analyses by including two-way multiplicative terms of HEI-by-PA or EA-by-PIR in all ordinal logistic regression models. Since a significant interaction was found between HEI and PA ($P<0.05$), we explored the combinatorial effects of PA (active vs. inactive) with the highest tertile of HEI (>56.64 vs. ≤56.64) on age-adjusted prevalence and risk estimates of NAFLD. Finally, we calculated age-adjusted prevalence estimates of NAFLD phenotypes for the first and third tertile of each HEI component, and differences were compared via adjusted ordinal logistic regression.

To examine subpopulations susceptible to DQ- and PA-related disparities, sensitivity analyses were performed by sex, race/ethnicity, EA, and PIR. Since EA levels differed across NAFLD phenotypes, HEI tertiles or PA status, we sought to determine whether HEI or PA might explain partly the effect of EA on risk of NAFLD via a causal mediation analysis. The aim of the mediation analysis was to quantify how much of the association between EA and NAFLD risk is mediated by either HEI or PA. As expected, body mass index (BMI) and waist circumference (WC) were positively associated with risk of NAFLD (**Appendix Table 1**). Further, HEI and PA were negatively linked with either BMI or WC (**Appendix Table 2**). Thus, we sought to examine how much of the total effect of HEI or PA on NAFLD risk could be explained by changes in BMI or WC via causal

This article is protected by copyright. All rights reserved

mediation analyses. The causal mediation analysis was performed via CAUSALMED procedure (SAS Institute, Inc) or *PROCESS* macro (version 3.5, model 4) (SPSS, version 26).(26, 27) Additional details regarding our mediation models are given in the Appendix.

The statistical reliability of estimates were determined based on the relative standard error as recommended by NCHS.(28) All data analyses were performed by SAS 9.4 (SAS Institute, Cary, NC, U.S.), SPSS version 26 (Chicago, IL, USA) and Stata 16 (StataCorp, College Station, TX, U.S.), incorporating complex sampling design (primary sampling units, sampling strata and weights) into statistical modeling. Two-sided $P < 0.05$ indicates statistical significance.

Results

Baseline characteristics based on NAFLD phenotypes

Baseline characteristics of the study participants are shown by the category of NAFLD phenotypes in **Table 1**. Compared with those without NAFLD, participants with NAFLD were older or male, had higher BMI, and age-adjusted prevalence of diabetes and hypertension. Individuals without NAFLD were more likely to be physically active (71.9%) than those with NAFLD without (67.1%) or with (63.7%) CSF, P for trend =0.005. The weighted mean HEI of the sampled adults was 48.7 (95% CI: 47.2-50.3); being significantly higher in non-NAFLD (49.6) participants than those with NAFLD without (47.0) or with (43.8) CSF, P for trend <0.01.

Association between physical activity levels and age-adjusted prevalence of NAFLD phenotypes

PA levels were inversely associated with the age-adjusted prevalence of NAFLD without or with CSF in a non-linear dose-response manner (P for trend =0.007) (**Figure 1A**). The age-adjusted prevalence of NAFLD without or with CSF were significantly lower in participants with PA levels of ≥600 MET-min/week than those with PA levels of <600 MET-min/week. After PA levels of ≥600 MET min/week, the benefits of PA plateaued and then persisted unchanged across the middle and highest cutoffs. The age-adjusted prevalence of NAFLD without and with CSF were significantly lower among physically active (28.7% and 4.9%) than inactive participants (32.5% and 6.9%), P for trend =0.005 (**Figure 1B**).

Association between diet quality (DQ) levels and age-adjusted prevalence of NAFLD phenotypes

This article is protected by copyright. All rights reserved

DQ levels (<42.05, 42.05-56.64 and >56.64) were inversely associated with the age-adjusted prevalence of NAFLD without or with CSF in a dose-response manner (P for trend =0.009) (**Figure 1C**). The lowest age-adjusted prevalence of NAFLD without or with CSF were observed among individuals with high-quality diet (25.3% and 3.2%) as compared with borderline-quality diet (31.4% and 5.6%) and low-quality diet.

Association between diet quality (DQ) and PIR levels and age-adjusted prevalence of NAFLD phenotypes

Age-adjusted prevalence of NAFLD without or with CSF did not differ significantly between different levels of PIR (P for trend = 0.829) (**Appendix Figure 2**). Among different educational levels, participants attaining some college or above displayed significantly lower age-adjusted rates of NAFLD without and with CSF (25.3% and 3.3%) than those with high school graduate or GED (32.2% and 6.9%) and less than high school graduate (30.6% and 5.7%), P for trend = 0.025 (**Figure 1D**).

Relationships between PA, DQ, SES and racial and ethnic groups

Table 2 displays baseline characteristics of participants by HEI tertiles and PA (active vs inactive). Participants who met PA recommendations (>600 MET-min/week) or had borderline- and high-quality diets were more likely to have some college degrees or above and PIR at or above 3.5% of the poverty level than their counterparts who were physically inactive or had PIR lesser than 3.5%. PA was not significantly different across HEI tertiles, P for trend =0.617.

Hispanic and NH-Black displayed the lowest levels of EA and PIR among all racial/ethnic groups. Attaining a college degree or above was only reported in 16% and 23% of Hispanic and NH-Black respectively compared with 35% and 60% of NH-White and NH-Asian (**Appendix Figure 3**). Likewise, only 28% and 27% of Hispanic and NH-Black reported PIR at or above 350% of the poverty level compared with NH-White (50%) and NH-Asian (63%) (**Appendix Figure 4**).

Association of DQ, PA and SES with risk of NAFLD

Table 3 shows covariate-adjusted associations of DQ, PA and SES with NAFLD risk. When HEI was evaluated on a continuous scale, higher DQ was associated with lower odds of NAFLD (OR: 0.98, 95% CI: 0.97-0.99). Compared with a LQD, a HQD was linked with reduced odds of NAFLD (OR: 0.58, 95% CI: 0.41-0.81). Increased PA (≥600 MET-min/week) clearly reduced the risk of

This article is protected by copyright. All rights reserved

NAFLD (OR:0.71, 95% CI: 0.51-0.98). Among physically active participants, those with PA levels of ≥8000 MET-min/week displayed the lowest risk of NAFLD (OR: 0.67, 95% CI: 0.47-0.97).

Compared with the lowest EA (less than high school graduate), attainment of a college degree or above was associated with a lower risk of NAFLD (OR:0.65, 95% CI: 0.43-0.96). There was no statistically significant difference for risk of NAFLD by PIR (P=0.829). We did not find multiplicative interaction effects between PIR and EA on NAFLD risk (P=0.175). Of note, participants who had an education of college or above showed a reduced risk of NAFLD regardless of a PIR status of <3.5 (OR: 0.67, 95% CI: 0.37-0.98) or >3.5 (OR: 0.54, 95% CI: 0.36-0.81) (**Appendix Table 3**).

Sensitivity analyses using LSM cutoffs at 90% sensitivity and 90% specificity for excluding or detecting NAFLD individuals with CSF confirmed the associations of PA, DQ and EA with risk of NAFLD (**Appendix Tables 4 and 5**).

Combinatorial effects of DQ and PA on the risk of NAFLD

The lowest age-adjusted prevalence of NAFLD without or with CSF was observed in adults who met PA recommendations (≥600 MET-min/week) and had a HQD (HEI of >56.64) (25% and 3%), followed by those physically inactive but with a HQD (28% and 5%), and physically active but with a non-HQD (HEI of ≤56.64) (30% and 6%) (**Table 4**). As compared to physically inactive adults with a non-HQD, physically active individuals with a HQD displayed the lowest risk of NAFLD (OR: 0.43, 95% CI: 0.30-0.63), followed by physically inactive participants but with a HQD (OR: 0.64, 95% CI: 0.42-0.98) and those physically active but with a non-HQD (OR: 0.75, 95% CI: 0.54-1.00) (**Figure 2**). A higher PA level along with a HQD remained significantly associated with a reduced risk of NAFLD in all racial and ethnic groups and both sexes compared with those physically inactive and a non-HQD (**Figure 3**).

The protective effects of increased PA along with a HQD on NAFLD risk remained statistically significant in further sensitivity analyses considering LSM cutoffs at 90% of sensitivity and 90% of specificity for ruling out or ruling in NAFLD participants with CSF (**Appendix Table 6**).

Distribution of HEI component scores across the NAFLD phenotypes

This article is protected by copyright. All rights reserved

Greater consumptions of total vegetables, greens and beans, total and whole fruits, seafood and plant protein, and fatty acids (a ratio of poly- and mono-unsaturated to saturated fatty acids) were observed among adults without NAFLD vs those with NAFLD (**Appendix Table 7**). Likewise, the intake of refined grains was significantly lower in non-NAFLD vs NAFLD participants. Participants without NAFLD tended to consume less saturated fat and added sugar than those with NAFLD, although differences were not statistically significant.

Diet quality and PA as potential mediators of the effect of EA on NAFLD risk

We sought to explore how much DQ, and PA might explain the effect of education on NAFLD risk via a causal mediation analysis (**Appendix Table 8).** Both DQ and EA appeared to be clear mediators of the EA effects on NAFLD risk. The estimated proportion of the total effect of EA on NAFLD risk mediated by HEI was 29.1%, and by PA was 7.9%.

BMI or waist circumference as potential mediators of the effect of PA or DQ on NAFLD risk

As expected, both PA and DQ were negatively associated with BMI (<0.01) and WC (<0.01) (**Appendix Table 2**). In addition, we found strong interaction effects between either BMI or WC with DQ (P<0.01) and PA (P<0.01) and risk of NAFLD. These findings suggested that part the effect of DQ and PA on NAFLD risk could be potentially explained by a change in the BMI or WC. Thus, covariate-adjusted causal mediation analyses were conducted to explore these associations. There were significant indirect "mediation" effects of DQ on NAFLD risk through BMI (-0.047, 95% CI: -0.059 to -0.035) and WC (-0.056, 95% CI: -0.069 to -0.044). BMI and WC could account for 85% and 98% of the total effect of DQ on NAFLD risk. Likewise, there were significant indirect "mediation" effects of PA on NAFLD risk through BMI (-0.018, 95% CI: -0.038 to -0.002) and WC (-0.029, 95% CI: -0.050 to -0.009). BMI and WC could account for 34% and 60% of the total effect of PA on NAFLD risk. The magnitude the mediation effects were significantly greater for WC than BMI. **Table 5** summarizes the causal mediation analysis.

Discussion

This article is protected by copyright. All rights reserved

Accumulated evidence supports an association between healthy dietary patterns and a decreased risk of NAFLD.(2, 29) Results from recent epidemiologic and interventional studies suggest that improved DQ through a reduced content of processed meat (2) or increased consumption of plant-based protein (30) is associated with reduced risks of NAFLD. Studies have also demonstrated an inverse association between NAFLD risk and increased PA.(5) PA, specifically during leisure time and travel-to-work time, is associated with lower prevalence of NAFLD.(3, 5) An inverse relationship between SES and unhealthy behaviors such as poor nutrition and physical inactivity have been demonstrated to be associated with poor health outcomes. However, the combined effects of healthy diet along with PA and SES on NAFLD risk are poorly understood.

In this nationally representative sample of U.S. adults, we found that those who met PA recommendations or had the healthier diet were at reduced risk of NAFLD than those who did not. We identified an inverse and dose-dependent effect of DQ on NAFLD risk; higher HEI scores were associated with lower risks of NAFLD. Of note, HQD was linked with a 42% reduction in risk of NAFLD. We also noted that levels of PA of ≥600 MET-min/week were associated with a 29% lower risk of NAFLD. Among adults physically active, those with PA levels of ≥8000 MET-min/week displayed a 37% reduction in NAFLD risk. Among those physically active with a HQD, the risk of NAFLD was significantly reduced up to 57%. The benefits of a HQD along with increased PA extended to all racial/ethnic groups and both sexes. These results underscore the concept that improvements in DQ along with increased PA could significantly decrease the risk of NAFLD as compared with each lifestyle intervention separately.

Our analyses showed that most of the HEI-2015 components contributed to the association of DQ and NAFLD risk. The components with the largest differences in mean scores between tertile 1 and tertile 3 were identified as total vegetables, greens and beans, total fruits, whole fruits, whole grains, seafood and plant protein, fatty acids, and refined grains. These results reinforce the concept that the quality of what we eat may have profound effects on health outcomes and support the current DGA recommendations about the importance of healthy eating patterns as a whole instead of focusing on individuals' nutrients or foods.

Aside from HEI and PA, we also found that adults attaining a college degree or above had a lower risk of NAFLD as compared with less educated persons, and this effect remained statistically

This article is protected by copyright. All rights reserved

significant even after controlling for many relevant confounders. Conversely, there was no significant difference in the risk of NAFLD among household income groups, although those with a PIR at or above 3.5 of the poverty levels tended to have a lower age-adjusted prevalence of NAFLD. Not surprisingly, highest educational levels and household incomes above the poverty line were associated with either healthier diet or increased PA. These findings could suggest that DQ and PA might act as proxies between SES and NAFLD risk. For example, our mediation analysis identified that a third of the total effect of EA on NAFLD risk was mediated by HQD and increased PA.

In alignment with our findings, previous research reported that low education and limited economic resources may jointly contribute to unhealthy dietary habits; less educated people with a low income tend to consume more low-cost and energy-dense foods.(31) Furthermore, less educated persons have less awareness of the connection between a healthy diet and physical activity and the lowered risk of chronic diseases. This can result in less motivation to incorporate more healthy lifestyle habits.(31) In our study, education, rather than income was the dominant factor associated with NAFLD risk. The highly educated persons consumed healthier diets, regardless of income level. While healthier diets can be more costly, this may suggest a re-prioritization of household resources towards purchasing higher quality foods in persons with higher educational attainment. Additional studies will be needed to determine how highly educated but low-income households maintain high-quality diets.

Lastly, our mediation analyses showed that a large percentage (85%-98%) of the beneficial effects of DQ on NAFLD risk are potentially exerted via changes in BMI or WC. These findings are in alignment with previous studies reporting that a weight loss via dietary changes is associated with lower risk of steatosis, inflammation, and fibrosis in a dose-dependent manner.(4, 32) We also found that BMI or WC partly mediated the relationship between PA and NAFLD risk. High PA levels, including aerobic and resistance training-based exercises are associated with reduced body weight and visceral and liver fat.(5, 33) However, the amount and the intensity of PA that would be needed to reduce body weight or WC and their effects on NAFLD risk are less clear. High levels of aerobic exercise are able to reduce liver fat in the absence of weight loss, which suggest that PA might reduce NAFLD risk through additional pathways such as improving insulin sensitivity.(34) Our results highlight that BMI or WC might explain 34%-60% of the total effect of PA on NAFLD risk,

This article is protected by copyright. All rights reserved

Accepted Article

suggesting that other clinical or physiological factors might explain the beneficial effects of PA on NAFLD risk.

Achieving dietary adherence and prevention of long-term weight gain are two of the biggest challenges faced by care providers. One of the strategies to optimize the dietary adherence rates is to promote a diet plan that includes a broad spectrum of diet options designed to better match individual patient food preferences, lifestyles, and medical conditions. Previous reports have suggested that improving DQ, i.e., incorporating more vegetables, fruits, whole grains and reducing the content of processed foods, is associated with important benefits in preventing weight gain or promoting weight loss in adults of both sexes and even in those with genetic predisposition for obesity.(35, 36) Thus, the use of healthy dietary patterns in the clinical setting could be a reasonable alternative to promote weight loss and reduce risk of NAFLD and its progression, although more studies are needed to identify short- and long-term beneficial effects of these dietary interventions in the context of NAFLD.

Some study strengths must be highlighted. This is one of the first studies examining how adhering to recommended healthy eating patterns and PA might influence risk of NAFLD in a nationally representative sample of the US In addition, we examined the relationships between SES, dietary and physical activity habits to enhance understanding about the impact of socioeconomic inequalities on healthy lifestyle habits. DQ was determined by the HEI-2015, which is a well-validated and evolving tool for the evaluation of healthy dietary patterns in the U.S. population.(11) The HEI-2015 has been shown to be an useful tool to predict all-cause, cardiovascular and cancer mortality in two large multiethnic cohorts.(12, 13)

Potential limitations should also be considered. Our dietary and PA assessments were based on self-reports, and they are subject to measurement errors. However, we used the two 24-hour dietary data to reduce the effect of day-to-day variation in food intake and obtain more precise estimates of diet. Further, the questionnaire used to assess dietary intake in NHANES has been extensively validated against diet records and biomarkers. Although we controlled for many covariates that were considered relevant confounders, it is possible that there was still residual and unmeasured confounding. The cross-sectional study design may preclude us from drawing conclusions regarding causality between lifestyle factors and NAFLD risk.

This article is protected by copyright. All rights reserved

Accepted Article

In conclusion, greater adherence to healthy eating patterns and increased PA were associated with a substantially lower risk of NAFLD, and these benefits were largely consistent across sexes and racial/ethnic groups. Our findings may add support for the 2015-2020 DGA, which focus on healthy eating patterns rather than individual ingredients and nutrients to better account for diverse cultural and personal food traditions and preferences. Educational attainment over income seems to be an important determinant of lifestyle behaviors. Thus, food and health public policies that target DQ-, PA-, and health-related knowledge may address the growing burden of NAFLD.

References

1. Chalasani N, Younossi Z, Lavine JE, Charlton M, Cusi K, Rinella M, Harrison SA, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. Hepatology 2018;67:328-357.
2. Zelber-Sagi S, Ivancovsky-Wajcman D, Fliss Isakov N, Webb M, Orenstein D, Shibolet O, Kariv R. High red and processed meat consumption is associated with non-alcoholic fatty liver disease and insulin resistance. J Hepatol 2018;68:1239-1246.
3. Kim D, Vazquez-Montesino LM, Li AA, Cholankeril G, Ahmed A. Inadequate Physical Activity and Sedentary Behavior Are Independent Predictors of Nonalcoholic Fatty Liver Disease. Hepatology 2020;72:1556-1568.
4. Vilar-Gomez E, Martinez-Perez Y, Calzadilla-Bertot L, Torres-Gonzalez A, Gra-Oramas B, Gonzalez-Fabian L, Friedman SL, et al. Weight Loss Through Lifestyle Modification Significantly Reduces Features of Nonalcoholic Steatohepatitis. Gastroenterology 2015;149:367-378.e365; quiz e314-365.
5. Orci LA, Gariani K, Oldani G, Delaune V, Morel P, Toso C. Exercise-based Interventions for Nonalcoholic Fatty Liver Disease: A Meta-analysis and Meta-regression. Clin Gastroenterol Hepatol 2016;14:1398-1411.
6. Cespedes EM, Hu FB. Dietary patterns: from nutritional epidemiologic analysis to national guidelines. Am J Clin Nutr 2015;101:899-900.

This article is protected by copyright. All rights reserved

7. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;392:1789-1858.
8. Tapsell LC, Neale EP, Satija A, Hu FB. Foods, Nutrients, and Dietary Patterns: Interconnections and Implications for Dietary Guidelines. Adv Nutr 2016;7:445-454.
9. Sotos-Prieto M, Bhupathiraju SN, Mattei J, Fung TT, Li Y, Pan A, Willett WC, et al. Association of Changes in Diet Quality with Total and Cause-Specific Mortality. New England Journal of Medicine 2017;377:143-153.
10. Schwingshackl L, Hoffmann G. Diet quality as assessed by the Healthy Eating Index, the Alternate Healthy Eating Index, the Dietary Approaches to Stop Hypertension score, and health outcomes: a systematic review and meta-analysis of cohort studies. J Acad Nutr Diet 2015;115:780-800.e785.
11. Reedy J, Lerman JL, Krebs-Smith SM, Kirkpatrick SI, Pannucci TE, Wilson MM, Subar AF, et al. Evaluation of the Healthy Eating Index-2015. J Acad Nutr Diet 2018;118:1622-1633.
12. Hu EA, Steffen LM, Coresh J, Appel LJ, Rebholz CM. Adherence to the Healthy Eating Index-2015 and Other Dietary Patterns May Reduce Risk of Cardiovascular Disease, Cardiovascular Mortality, and All-Cause Mortality. J Nutr 2020;150:312-321.
13. Panizza CE, Shvetsov YB, Harmon BE, Wilkens LR, Le Marchand L, Haiman C, Reedy J, et al. Testing the Predictive Validity of the Healthy Eating Index-2015 in the Multiethnic Cohort: Is the Score Associated with a Reduced Risk of All-Cause and Cause-Specific Mortality? Nutrients 2018;10:452.
14. Millen BE, Abrams S, Adams-Campbell L, Anderson CA, Brenna JT, Campbell WW, Clinton S, et al. The 2015 Dietary Guidelines Advisory Committee Scientific Report: Development and Major Conclusions. Adv Nutr 2016;7:438-444.
15. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, George SM, et al. The Physical Activity Guidelines for Americans. Jama 2018;320:2020-2028.
16. Kim D, Murag S, Cholankeril G, Cheung A, Harrison SA, Younossi ZM, Ahmed A. Physical Activity, Measured Objectively, Is Associated With Lower Mortality in Patients With Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2021;19:1240-1247.e1245.

This article is protected by copyright. All rights reserved

17. Alkerwi A, Sauvageot N, Nau A, Lair ML, Donneau AF, Albert A, Guillaume M. Population compliance with national dietary recommendations and its determinants: findings from the ORISCAV-LUX study. Br J Nutr 2012;108:2083-2092.
18. Stringhini S, Carmeli C, Jokela M, Avendaño M, Muennig P, Guida F, Ricceri F, et al. Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1·7 million men and women. Lancet 2017;389:1229-1237.
19. Chen TC, Clark J, Riddles MK, Mohadjer LK, Fakhouri THI. National Health and Nutrition Examination Survey, 2015-2018: Sample Design and Estimation Procedures. Vital Health Stat 2 2020:1-35.
20. Centers for Disease Control and Prevention. NHANES 2017-2018 Overview. https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/overview.aspx?BeginYear=2017. Accessed July 23, 2021.
21. Centers for Disease Control and Prevention. NHANES 2017-2018 Procedure Manuals. https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/manuals.aspx?BeginYear=2017. Accessed July 23, 2021.
22. Siddiqui MS, Vuppalanchi R, Van Natta ML, Hallinan E, Kowdley KV, Abdelmalek M, Neuschwander-Tetri BA, et al. Vibration-Controlled Transient Elastography to Assess Fibrosis and Steatosis in Patients With Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2019;17:156-163 e152.
23. Centers for Disease Control and Prevention. NHANES 2017-2018. MEC In-Person Dietary Interviewers Procedures Manual. https://wwwn.cdc.gov/nchs/data/nhanes/2017-2018/manuals/2017_MEC_In-Person_Dietary_Interviewers_Manual.pdf021. Accessed July 23, 2021.
24. Cleland CL, Hunter RF, Kee F, Cupples ME, Sallis JF, Tully MA. Validity of the global physical activity questionnaire (GPAQ) in assessing levels and change in moderate-vigorous physical activity and sedentary behaviour. BMC Public Health 2014;14:1255.
25. Kyu HH, Bachman VF, Alexander LT, Mumford JE, Afshin A, Estep K, Veerman JL, et al. Physical activity and risk of breast cancer, colon cancer, diabetes, ischemic heart disease, and ischemic stroke events: systematic review and dose-response meta-analysis for the Global Burden of Disease Study 2013. Bmj 2016;354:i3857.

This article is protected by copyright. All rights reserved

26. VanderWeele TJ, Vansteelandt S. Mediation Analysis with Multiple Mediators. Epidemiol Methods 2014;2:95-115.
27. Hayes AF. An Index and Test of Linear Moderated Mediation. Multivariate behavioral research 2015;50:1-22.
28. Parker JD, Talih M, Malec DJ, Beresovsky V, Carroll M, Gonzalez JF, Hamilton BE, et al. National Center for Health Statistics Data Presentation Standards for Proportions. Vital Health Stat 2 2017:1-22.
29. Hassani Zadeh S, Mansoori A, Hosseinzadeh M. Relationship between dietary patterns and non-alcoholic fatty liver disease: A systematic review and meta-analysis. J Gastroenterol Hepatol 2021;36:1470-1478.
30. Yaskolka Meir A, Rinott E, Tsaban G, Zelicha H, Kaplan A, Rosen P, Shelef I, et al. Effect of green-Mediterranean diet on intrahepatic fat: the DIRECT PLUS randomised controlled trial. Gut 2021.
31. Hiza HA, Casavale KO, Guenther PM, Davis CA. Diet quality of Americans differs by age, sex, race/ethnicity, income, and education level. J Acad Nutr Diet 2013;113:297-306.
32. Promrat K, Kleiner DE, Niemeier HM, Jackvony E, Kearns M, Wands JR, Fava JL, et al. Randomized controlled trial testing the effects of weight loss on nonalcoholic steatohepatitis. Hepatology 2010;51:121-129.
33. Keating SE, Hackett DA, Parker HM, O'Connor HT, Gerofi JA, Sainsbury A, Baker MK, et al. Effect of aerobic exercise training dose on liver fat and visceral adiposity. J Hepatol 2015;63:174-182.
34. Johnson NA, Sachinwalla T, Walton DW, Smith K, Armstrong A, Thompson MW, George J. Aerobic exercise training reduces hepatic and visceral lipids in obese individuals without weight loss. Hepatology 2009;50:1105-1112.
35. Wang T, Heianza Y, Sun D, Huang T, Ma W, Rimm EB, Manson JE, et al. Improving adherence to healthy dietary patterns, genetic risk, and long term weight gain: gene-diet interaction analysis in two prospective cohort studies. BMJ 2018;360:j5644.
36. Satija A, Malik V, Rimm EB, Sacks F, Willett W, Hu FB. Changes in intake of plant-based diets and weight change: results from 3 prospective cohort studies. Am J Clin Nutr 2019;110:574-582.

This article is protected by copyright. All rights reserved

Table 1. Baseline characteristics across the NAFLD phenotypes based on VCTE measurements: Non-NAFLD (CAP <285 dB/m) * vs NAFLD (CAP ≥285 dB/m) without clinically significant fibrosis (LSM <8.6 kPa) † vs NAFLD (CAP ≥285 dB/m) with clinically significant fibrosis (LSM ≥8.6 kPa), n=3859. Age-adjusted estimates.

	Non-NAFLD *Weighted mean or percentage* (95% CI) N=2437	**NAFLD without clinically significant fibrosis** *Weighted mean or percentage* (95% CI) N=1192	**NAFLD with clinically significant fibrosis** *Weighted mean or percentage* (95% CI) N=230	P value ‡
VCTE measurements				
CAP (dB/m)	224 (222-227)	327 (323-331)	351 (345-357)	<0.001
LSM (kPa)	4.85 (4.69-5.01)	5.34 (5.23-5.45)	17.80 (13.95-21.65)	<0.001
Demographics				
Age, years	45.03 (43.22-46.84)	46.47 (45.84-47.11)	47.20 (45.74-48.66)	<0.001
Gender,				<0.001
Male, n=1873	44.2 (42.0-46.3)	54.8 (49.1-60.4)	62.7 (50.9-73.2)	
Female, n=1986	55.8 (53.7-57.96)	45.2 (39.6-50.8)	37.3 (26.8-49.0)	
Race/ethnicity				0.386
Non-Hispanic white, n=1331	61.9 (56.2-67.3)	60.0 (52.4-67.1)	63.9 (54.3-72.6)	
Non-Hispanic black, n=877	12.8 (9.6-16.9)	8.0 (5.1-12.2)	7.1 (3.3-14.6)	

This article is protected by copyright. All rights reserved

Non-Hispanic Asian, n=537	6.1 (4.3-8.6)	5.5 (3.7-8.0)	3.0 (1.7-5.4)	
Hispanic, n=910	14.3 (11.0-18.5)	22.4 (16.1-30.3)	20.6 (14.5-28.5)	
Other or multiracial, n=204	4.9 (3.4-6.7)	4.1 (2.8-6.1)	5.4 (2.5-11.3)	
Educational attainment				0.001
Less than high school graduate, n=708	11.8 (8.9-15.4)	12.6 (8.8-17.7)	13.0 (8.4-19.4)	
High school graduate or GED, n=2124	53.7 (45.7-61.4)	61.2 (56.1-66.0)	68.7 (56.9-78.5)	
Some college or above, n=1027	34.5 (27.4-42.4)	26.2 (21.1-32.1)	18.3 (10.3-30.3)	
Family income to poverty ratio				0.829
<1.30, n=974	20.5 (17.9-23.4)	19.9 (17.2-23.0)	22.7 (15.3-32.3)	
1.30-3.49, n=1411	35.2 (30.7-39.9)	39.8 (33.6-46.3)	42.3 (30.3-55.3)	
≥3.50, n=1020	44.3 (39.4-49.3)	40.2 (33.7-47.1)	35.0 (23.3-48.7)	
Comorbidities				
Body mass index (kg/m^2)	27.0 (26.4-27.6)	33.6 (32.8-34.3)	40.2 (38.7-41.7)	<0.001
Waist circumference (cm)	93.2 (91.8-94.7)	110.4 (108.6-112.2)	126.6 (123.9-129.3)	<0.001
Hypertension, n=1677	28.2 (24.8-31.8)	46.5 (40.7-52.3)	55.1 (43.4-66.2)	<0.001
Diabetes mellitus, n=777	7.3 (5.9-8.9)	19.1 (17.1-21.3)	45.6 (37.1-54.5)	<0.001
Smoking status				0.653
Never, n=2356	62.7 (58.1-67.2)	58.8 (54.2-63.2)	61.7 (51.8-70.7)	
Former, n=892	21.3 (18.7-24.1)	25.9 (21.4-31.1)	27.6 (21.4-34.8)	
Current, n=611	16.0 (13.0-19.5)	15.3 (12.6-18.3)	10.7 (5.9-18.4)	

Accepted Article

This article is protected by copyright. All rights reserved

Non-heavy drinkers, n=2497	73.1 (70.5-75.6)	73.1 (69.4-76.4)	71.2 (63.4-77.9)	0.322
Physical activity				
≥600 MET-minutes/week, n=2452 ***§***	71.9 (68.9-74.6)	67.1 (62.0-71.7)	63.7 (55.4-70.4)	0.005
Diet quality				
Healthy Eating Index (2015) scores (0-100)	49.59 (47.92-51.25)	47.00 (45.29-48.67)	43.82 (40.43-47.20)	0.008
Lab panel				
Total cholesterol (mg/dl)	188.6 (185.6-191.7)	190.8 (181.6-199.9)	182.5 (163.8-201.2)	0.039
Triglycerides (mg/dl)	118.0 (112.2-123.7)	160.9 (138.3-183.4)	174.5 (125.1-224.0)	<0.001
High-density lipoprotein (mg/dl)	56.1 (55.2-57.1)	48.1 (46.6-49.4)	43.7 (41.6-45.9)	<0.001
Low-density lipoprotein (mg/dl)	109.1 (106.5-111.6)	109.6 (105.4-113.8)	107.4 (98.9-115.8)	0.938
Fasting glucose (mg/dl)	104.2 (102.7-105.8)	117.7 (112.6-122.9)	138.8 (120.7-156.9)	<0.001
Glycosylated hemoglobin (%)	5.47 (5.43-5.51)	5.90 (5.76-6.04)	6.47 (6.03-6.90)	<0.001
Alanine aminotransferase (U/L)	19.5 (18.9-20.1)	26.9 (25.5-28.4)	38.6 (30.5-46.7)	<0.001
Aspartate aminotransferase (U/L)	20.6 (20.1-21.2)	21.7 (20.7-22.6)	29.7 (25.8-33.6)	0.003
Gamma-Glutamyl Transferase (U/L)	23.2 (21.8-24.6)	32.5 (27.3-37.7)	53.6 (44.3-62.9)	<0.001
Albumin (g/dl)	4.12 (4.08-4.17)	4.01 (3.95-4.06)	3.92 (3.79-4.06)	0.004
Platelet x 10^9/L	242 (236-249)	248 (241-255)	243 (223-263)	0.034
Total bilirubin (mg/dl)	0.49 (0.46-0.51)	0.43 (0.42-0.45)	0.47 (0.42-0.53)	0.014
HS-CRP (mg/dl)	127.7 (118.8-136.5)	164.1 (143.5-184.6)	187.3 (149.3-225.2)	<0.001

Abbreviations: VCTE, vibration-controlled transient elastography; CAP, controlled attenuation parameter; LSM, liver stiffness measure; HS-CRP, high-sensitive c-reactive protein.

This article is protected by copyright. All rights reserved

* CAP cutoff of 285 selected based on Youden's index which yields a sensitivity of 80% and specificity of 77%.

† LSM cutoff of 8.6 selected based on Youden's index which yields a sensitivity of 66% and specificity of 80%.

‡ Unadjusted P values for comparison between the three groups. These were calculated via linear, binary, or ordinal logistic regression when appropriate.

§ Participants with PA ≥600 MET-minutes/week are considered as physically active.

Accepted Article

This article is protected by copyright. All rights reserved

Table 2. Baseline characteristics of participants (n=3859) across tertiles of Healthy Eating Index score and physical activity status. Age-adjusted estimates.

	Healthy Eating Index score (tertiles)				**Physical activity status ***		
	Low-quality diet Tertile 1 (<42.05) *Weighted mean or percentage* (95% CI)	***Borderline-quality diet*** Tertile 2 (42.05-56.64) *Weighted mean or percentage* (95% CI)	***High-quality diet*** Tertile 3 (>56.64) *Weighted mean or percentage* (95% CI)	**P value †**	***Inactive*** *Weighted mean or percentage* (95% CI)	***Active*** *Weighted mean or percentage* (95% CI)	**P value †**
Number	1287	1286	1286		1407	2452	
				0.006			0.007
Male, n=1873	54.5 (49.3-59.6)	45.7 (41.6-49.8)	44.8 (40.4-49.3)		36.6 (32.1-41.3)	53.4 (50.2-56.6)	
Female, n=1986	45.5 (40.4-50.6)	54.3 (50.2-58.4)	55.2 (50.7-59.6)		63.4 (58.7-67.9)	46.6 (43.4-49.8)	
Educational attainment				<0.001			0.011
Less than high school graduate, n=708	12.1 (9.5-15.3)	12.3 (9.2-16.2)	11.4 (8.4-15.2)		15.1 (11.7-19.3)	10.4 (7.9-13.5)	
High school graduate or GED, n=2124	65.7 (58.6-72.2)	59.8 (53.9-65.5)	42.4 (35.6-49.6)		58.4 (51.1-65.3)	56.0 (48.5-63.2)	
Some college or above, n=1027	22.2 (16.2-29.6)	27.9 (22.4-34.1)	46.2 (38.9-53.7)		26.5 (21.0-32.8)	33.6 (26.8-41.2)	
Family income to poverty ratio				0.033			0.020
<1.30, n=974	23.5 (20.9-26.3)	18.8 (15.8-22.3)	19.1 (15.9-22.7)		25.8 (21.9-30.2)	18.6 (16.3-21.0)	
1.30-3.49, n=1411	38.8 (33.9-43.9)	37.9 (31.4-44.9)	32.3 (26.5-38.6)		35.3 (30.4-40.4)	36.5 (31.6-41.8)	
≥3.50, n=1020	37.7 (31.6-44.3)	43.3 (36.4-50.4)	48.6 (42.5-54.7)		38.9 (32.7-45.5)	44.9 (40.0-49.8)	
Physically active, n=2452	69.4 (66.5-72.1)	67.2 (61.2-72.6)	73.5 (69.3-77.4)	0.617	-	-	-

Accepted Article

This article is protected by copyright. All rights reserved

VCTE measurements							
CAP (dB/m)	270 (265-275)	262 (255-270)	251 (244-257)	0.012	267 (262-273)	259 (253-264)	0.004
LSM (kPa)	6.11 (5.84-6.39)	5.61 (5.17-6.04)	5.12 (4.90-5.33)	<0.001	5.78 (5.49-6.06)	5.55 (5.28-5.82)	0.049
				0.009			0.005
CAP <285, n=2437	59.9 (56.1-63.6)	63.0 (56.9-68.6)	71.5 (67.1-75.6)		60.6 (54.9-65.9)	66.4 (62.5-70.0)	
CAP ≥285 / LSM<8.6, n=1192	32.0 (28.1-36.1)	31.4 (25.7-37.7)	25.3 (20.9-30.3)		32.5 (27.9-37.5)	28.7 (25.0-32.7)	
CAP ≥285 / LSM≥8.6, n=230	8.1 (5.5-11.9)	5.6 (3.5-8.8)	3.2 (1.9-5.3)		6.9 (4.8-9.6)	4.9 (3.4-7.0)	
Anthropometrics							
Body mass index (kg/m^2)	30.7 (30.2-31.1)	29.9 (29.1-30.7)	27.5 (26.9-28.1)	<0.001	30.3 (29.4-31.1)	29.3 (28.6-30.0)	0.010
Waist circumference (cm)	103.3 (102.1-104.6)	100.5 (98.6-102.5)	94.6 (93.3-95.8)	<0.001	101.5 (99.5-103.5)	99.2 (97.3-101.1)	0.001
Lab panel							
Triglycerides (mg/dl)	149.9 (140.4-159.3)	142.4 (131.5-153.3)	128.3 (122.9-133.6)	0.006	151.3 (139.5-162.9)	137.8 (128.8-146.8)	0.003
Cholesterol (mg/dl)	187.9 (184.1-191.8)	190.4 (185.4-195.3)	185.5 (182.8-188.2)	0.611	189.6 (185.1-194.1)	188.4 (184.2-192.7)	0.137
High-density lipoprotein (mg/dl)	49.7 (48.7-50.7)	52.8 (51.0-54.6)	55.9 (54.4-57.4)	<0.001	52.0 (50.7-53.2)	52.9 (51.6-54.1)	0.864
Low-density lipoprotein (mg/dl)	108.7 (103.9-113.5)	109.9 (105.9-113.9)	105.9 (102.9-108.9)	0.490	107.9 (103.8-112.1)	108.3 (105.1-111.7)	0.909
Alanine aminotransferase (U/L)	22.8 (21.8-23.8)	22.3 (21.2-23.4)	21.9 (21.1-22.7)	0.049	21.8 (20.6-22.9)	22.8 (22.2-23.5)	0.065
Aspartate aminotransferase (U/L)	21.8 (20.4-21.8)	20.8 (19.9-21.5)	22.1 (21.1-23.1)	0.279	20.5 (19.8-22.2)	21.6 (21.1-22.1)	0.071
Gamma-Glutamyl Transferase (U/L)	30.4 (26.9-33.8)	26.3 (25.2-27.4)	24.2 (22.7-25.9)	0.010	28.9 (26.2-31.6)	26.9 (25.5-28.3)	0.167
Alkaline phosphatase (U/L)	79.8 (77.9-81.8)	75.4 (73.8-77.1)	73.2 (71.3-75.0)	0.002	79.3 (76.9-81.6)	75.3 (73.2-77.4)	0.003
HS-CRP (mg/dl)	4.45 (4.00-4.89)	3.63 (3.22-4.03)	3.22 (2.60-3.83)	<0.001	4.47 (3.82-5.11)	3.42 (2.99-3.85)	0.002

Abbreviations: VCTE, vibration controlled transient elastography CAP, controlled attenuation parameter; LSM, liver stiffness measure; HS-CRP, high-sensitive c-reactive protein.

This article is protected by copyright. All rights reserved

* The active participants included those who met the recommended levels of PA of ≥600 MET minutes/week. The inactive participants were those who did not achieve the recommended levels of PA (MET minutes/week <600). (32)

† Unadjusted P values calculated via linear, binary, or ordinal logistic regression when appropriate.

Accepted Article

This article is protected by copyright. All rights reserved

Table 3. Risk across the NAFLD phenotypes by diet quality, physical activity, and socioeconomic status.

	Multivariable analysis *			
	Age-adjusted OR (95% CI)	P value	***Model 1*** OR (95% CI)	P value
Healthy Eating Index (2015) score				
Continuous scale	0.98 (0.97-0.99)	<0.001	0.98 (0.97-0.99)	0.001
Tertiles				
<42.05 (*Low-quality diet)*	1 (ref.)	-	1 (ref.)	-
42.05-56.64 (*Borderline-quality diet)*	0.87 (0.66-1.16)	0.338	0.87 (0.63-1.20)	0.378
>56.64 (*High-quality diet)*	0.56 (0.43-0.73)	<0.001	0.58 (0.41-0.81)	0.004
Physical activity (MET-min/week)				
Continuous scale	0.99 (0.99-0.99)	0.016	0.99 (0.99-0.99)	0.001
Cutoffs				
<600	1 (ref.)		1 (ref.)	
600-7999	0.67 (0.51-0.88)	0.008	0.71 (0.50-0.99)	0.050
≥8000	0.67 (0.51-0.89)	0.008	0.67 (0.47-0.97)	0.036
Physically active (≥600)	0.76 (0.59-0.99)	0.048	0.71 (0.51-0.98)	0.043
Educational attainment				
Less than high school graduate	1 (ref.)	-	1 (ref.)	-
High school graduate or GED	1.14 (0.77-1.67)	0.476	1.15 (0.85-1.56)	0.338
Some college or above	0.69 (0.44-0.99)	0.044	0.65 (0.43-0.96)	0.034
Family income to poverty ratio				
<1.30	1 (ref.)	-	1 (ref.)	-
1.30-3.49	1.07 (0.78-1.46)	0.633	1.14 (0.85-1.52)	0.351
≥3.50	0.88 (0.67-1.15)	0.328	1.02 (0.75-1.38)	0.885

Abbreviation: ref., reference.

Accepted Article

This article is protected by copyright. All rights reserved

* Ordinal logistic regression model.

Model 1: age plus sex, race and ethnicity, total energy intake, education level, family income to poverty ratio, non-heavy alcohol intake, physical activity, smoking status, as well as Healthy Eating Index score.

Accepted Article

This article is protected by copyright. All rights reserved

Table 4. Age-adjusted estimates of selected baseline feature across different combinatorial subgroups based on diet quality (high- vs non-high) and physical activity (active vs inactive).

	Physically inactive Non-high-quality diet † *Weighted mean or percentage* (95% CI) N=928	**Physically active Non-high-quality diet †** *Weighted mean or percentage* (95% CI) N=1645	**Physically inactive High-quality diet ‡** *Weighted mean or percentage* (95% CI) N=479	**Physically active High-quality diet ‡** *Weighted mean or percentage* (95% CI) N=807	P value *
VCTE measurements					
CAP (dB/m)	270 (263-278)	264 (258-269)	258 (248-267)	248 (240-257)	0.001
LSM (kPa)	5.99 (5.63-6.34)	5.73 (5.44-6.02)	5.19 (4.80-5.58)	5.07 (4.74-5.37)	0.001
					0.001
CAP <285, n=2437	58.3 (51.6-64.7)	63.5 (59.2-67.5)	67.5 (59.2-74.5)	72.8 (66.5-78.4)	
CAP ≥285 / LSM <8.6, n=1192	34.0 (28.4-40.1)	30.4 (25.7-35.6)	27.5 (21.4-34.6)	24.6 (19.0-31.1)	
CAP ≥285 / LSM ≥8.6, n=230	7.7 (4.8-11.9)	6.1 (4.3-8.4)	5.0 (2.7-7.8)	2.6 (1.3-5.0)	
Anthropometrics					
Body mass index (kg/m^2)	30.7 (29.7-31.6)	30.1 (29.5-30.8)	28.7 (27.7-29.8)	27.4 (26.8-28.1)	<0.001
Waist circumference (cm)	102.6 (100.5-104.7)	101.3 (99.7-102.9)	96.8 (94.8-98.9)	94.3 (92.9-95.8)	<0.001
Labs					
Triglycerides (mg/dl)	149.3 (138.4-160.1)	142.3 (134.7-149.9)	140.7 (129.6-145.8)	123.2 (116.9-129.4)	0.001

This article is protected by copyright. All rights reserved

Total cholesterol (mg/dl)	185.8 (182.4-189.2)	190.2 (185.1-195.1)	189.8 (186.1-193.4)	184.2 (181.3-187.2)	0.179
Low-density lipoprotein (mg/dl)	107.2 (103.5-110.9)	110.1 (105.6-114.7)	110.6 (101.7-119.6)	104.7 (101.7-107.8)	0.109
High-density lipoprotein (mg/dl)	50.7 (49.3-52.1)	51.7 (50.5-52.9)	53.5 (52.3-54.7)	56.3 (54.5-58.1)	0.001
Alanine aminotransferase (U/L)	21.4 (19.9-22.9)	23.2 (22.3-23.9)	22.9 (21.4-24.4)	21.9 (20.8-23.2)	0.485
Aspartate aminotransferase (U/L)	19.9 (19.3-20.6)	21.2 (20.7-21.7)	20.9 (20.1-21.9)	22.1 (21.0-23.1)	0.128
Gamma-Glutamyl Transferase (U/L)	28.6 (26.0-31.2)	28.4 (25.4-29.5)	26.6 (23.2-29.9)	23.2 (21.4-25.1)	0.004
Alkaline phosphatase (U/L)	80.8 (78.8-82.8)	75.8 (74.1-77.6)	74.3 (72.4-76.3)	72.8 (70.5-75.1)	0.001
HS-CRP (mg/dl)	4.68 (3.87-5.49)	3.65 (3.30-4.01)	3.88 (2.81-4.95)	2.94 (2.12-3.76)	<0.001

Abbreviations: VCTE, vibration controlled transient elastography; CAP, controlled attenuation parameter; LSM, liver stiffness measure; HEI, healthy eating index; HS-CRP, high-sensitive c-reactive protein.

* Unadjusted P value calculated via linear or ordinal logistic regression when appropriate.

† A non-high-quality diet refers to a healthy eating index score of ≤56.64.

‡ A high-quality diet refers to a healthy eating index score of >56.64.

Accepted Article

This article is protected by copyright. All rights reserved

Table 5. Role of body mass index (kg/m^2) or waist circumference (cm) as mediators between DQ (HEI tertiles) or PA (active vs inactive) and risk of NAFLD without or with CSF. Results based on covariate-adjusted causal mediation analyses.

Mediators	**Diet quality (HEI tertiles)**			
	Bootstrapped estimates (n=10,000)			
Body mass index (kg/m^2)	Estimates *	95% confidence intervals	SEs	P value
Total effect of HEI (tertiles) †	-0.055	-0.079 to -0.031	0.012	<0.001
Direct effect of HEI (tertiles) ‡	-0.008	-0.029 to 0.013	0.011	0.461
Indirect effect of HEI (tertiles) through BMI §	-0.047	-0.059 to -0.035	0.006	<0.001
Percentage mediated	*85.5*	-	-	-
Waist circumference (cm)	Estimates *	95% confidence intervals	SEs	P value
Total effect of HEI (tertiles) †	-0.056	-0.080 to -0.032	0.012	<0.001
Direct effect of HEI (tertiles) ‡	-0.001	-0.020 to 0.021	0.011	0.955
Indirect effect of HEI (tertiles) through WC §	-0.055	-0.069 to -0.044	0.006	<0.001
Percentage mediated	*98.2*	-	-	-
Mediators	**PA (active vs inactive)**			
Body mass index (kg/m^2)	Estimates *	95% confidence intervals	SEs	P value
Total effect of PA (active) †	-0.053	-0.094 to -0.012	0.021	0.010
Direct effect of PA (active) ‡	-0.035	-0.070 to 0.001	0.018	0.056
Indirect effect of PA (active) through BMI §	-0.018	-0.038 to -0.002	0.010	0.047
Percentage mediated	*34.0*	-	-	-
Waist circumference (cm)	Estimates *	95% confidence intervals	SEs	P value
Total effect of PA (active) †	-0.048	-0.089 to -0.007	0.021	0.021
Direct effect of PA (active) ‡	-0.019	-0.054 to 0.016	0.018	0.293
Indirect effect of PA (active) through WC §	-0.029	-0.050 to -0.009	0.011	<0.001
Percentage mediated	*60.4*	-	-	-

Abbreviations: CSF, clinically significant fibrosis; SEs, standard errors; HEI, Healthy Eating Index; PA, physical activity; BMI, body mass index; WC, waist circumference.

* Bootstrapped β coefficients.

This article is protected by copyright. All rights reserved

Negative estimates [bootstrapped β coefficients]) indicate that both HEI and PA are inversely associated with the risk of NAFLD without or with clinically significant fibrosis.

Analysis adjusted for age, sex, race and ethnicity, total energy intake, non-heavy alcohol intake, physical activity, smoking status, educational levels, and family income to poverty ratio.

Estimates are calculated taking the non-NAFLD group as reference.

† Effect of HEI tertiles or PA (active) on the NAFLD risk including the mediator (BMI or WC) effect.

‡ Effect of HEI tertiles or PA (active) on the NAFLD risk excluding the mediator (BMI or WC) effect.

§ Effect of HEI tertiles or PA (active) on the NAFLD risk through the mediator (BMI or WC).

Figure legends

Figure 1. Age-adjusted prevalence of NAFLD phenotypes by physical activity levels and diet quality (Healthy Eating Index) tertiles and educational levels.

A) Physical activity levels (<600, 600-7999 and ≥8000 MET-min/week).

B) Physical activity levels (<600 vs ≥600 MET-min/week).

C) Diet quality (Healthy Eating Index) tertiles (<42.05, 42.05-56.64 and >56.64).

D) Educational levels.

Abbreviations: NAFLD, non-alcoholic fatty liver disease; CSF, clinically significant fibrosis.

Numbers at the bottom of bars represent the weighted percentage. Bar whiskers represent the 95% confidence level.

Figure 2. Adjusted risk across the NAFLD phenotypes by physical activity levels and diet quality.

Abbreviations: CAP, controlled attenuation parameter; LSM, liver stiffness measure; HEI, healthy eating index.

This article is protected by copyright. All rights reserved

Accepted Article

Risk of NAFLD without or with clinically significant fibrosis was computed via ordinal logistic regressions models. All analyses were adjusted for age, sex, race, total energy intake, education level, family income to poverty ratio, non-heavy alcohol intake and smoking status.

The reference group are participants with a non-HQD (HEI ≤56.64) and physically inactive.

Figure 3. Adjusted risk across the NAFLD phenotypes among participants physically active with HQD vs those physically inactive with non-HQD. Sensitivity analysis by racial and ethnic and gender subgroups.

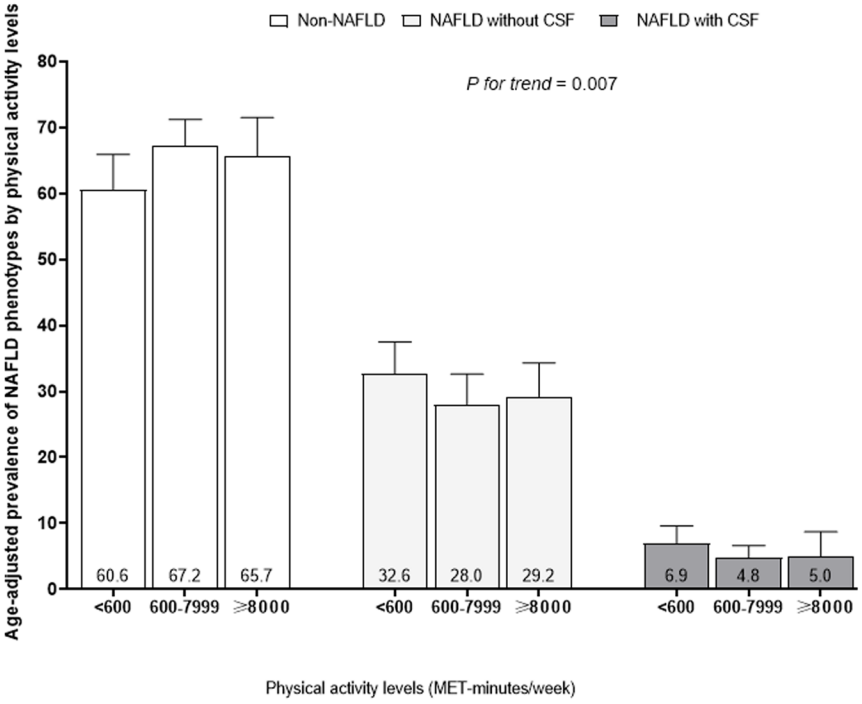
Abbreviations: HQD, high-quality diet; NH, non-Hispanic.

The reference group are participants with a non-HQD (HEI ≤56.64) and physically inactive.

Risk of NAFLD without or with clinically significant fibrosis was computed via ordinal logistic regressions model. All analyses were adjusted for age, total energy intake, education levels, family income to poverty ratio, non-heavy alcohol intake and smoking status.

Accepted Article

This article is protected by copyright. All rights reserved

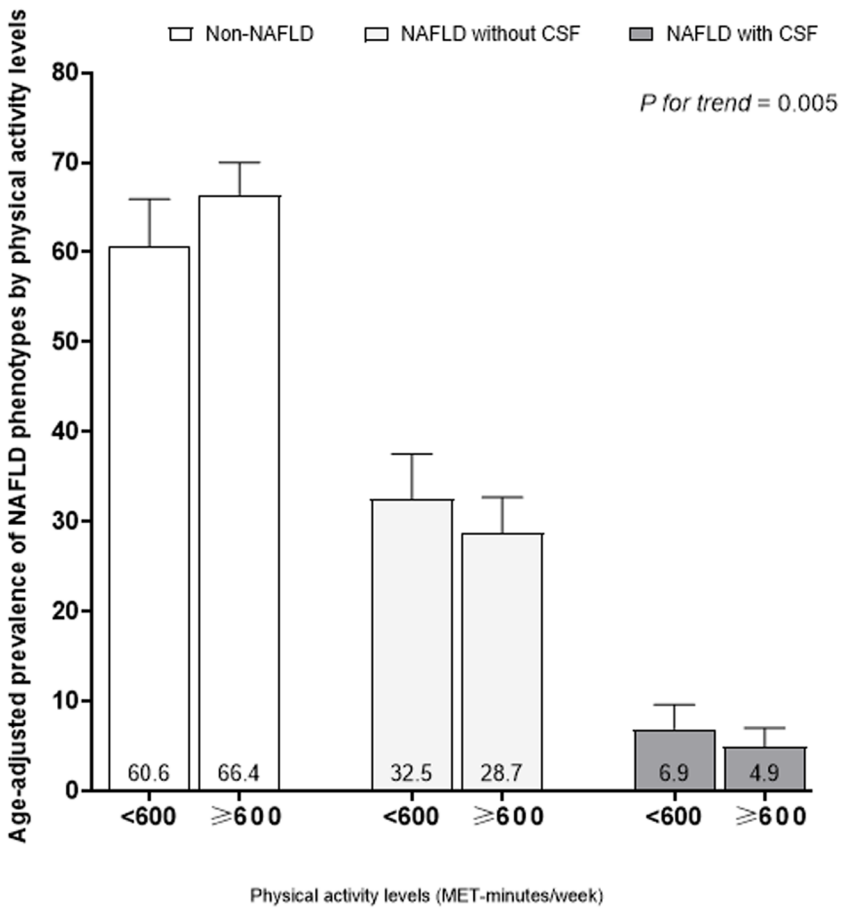


hep_32207_f1a.tif

Accepted Article

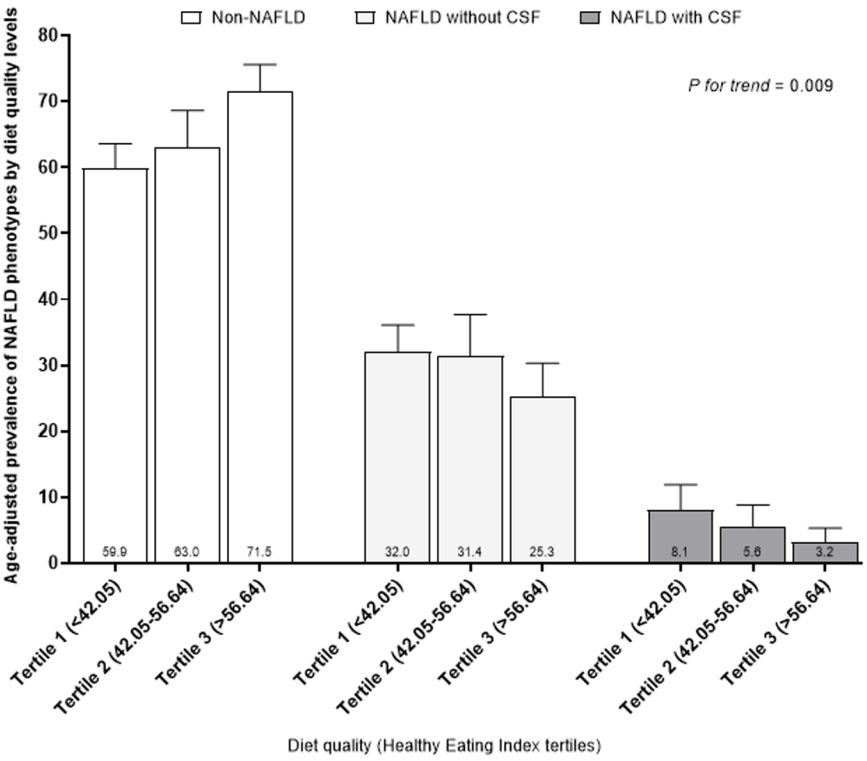
This article is protected by copyright. All rights reserved

Accepted Article



hep_32207_f1b.tif

This article is protected by copyright. All rights reserved

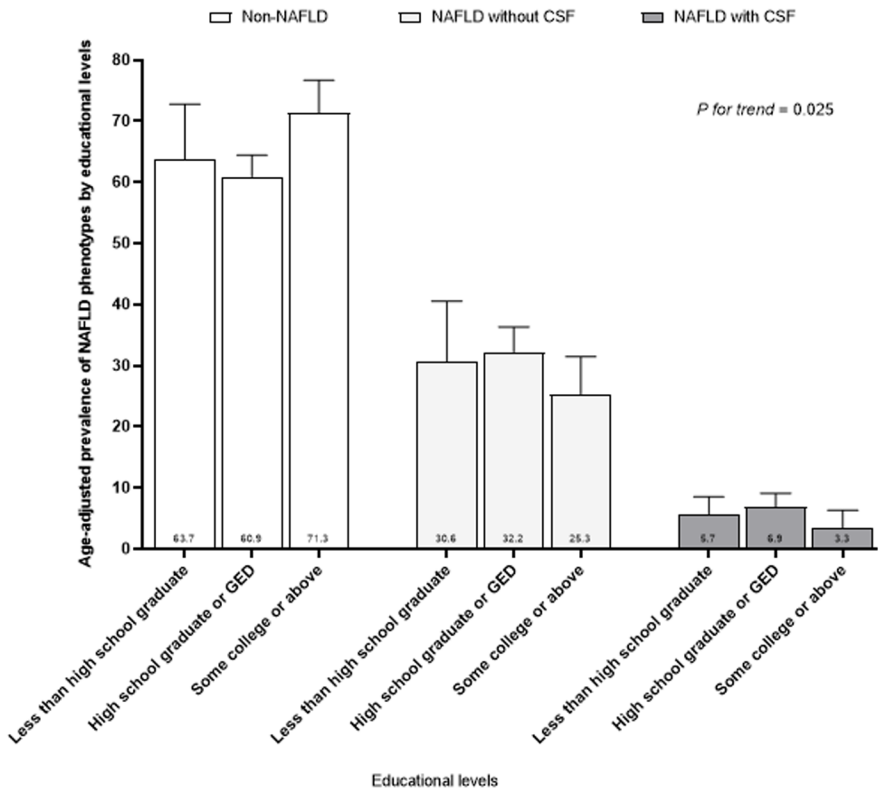


hep_32207_f1c.tif

Accepted Article

This article is protected by copyright. All rights reserved

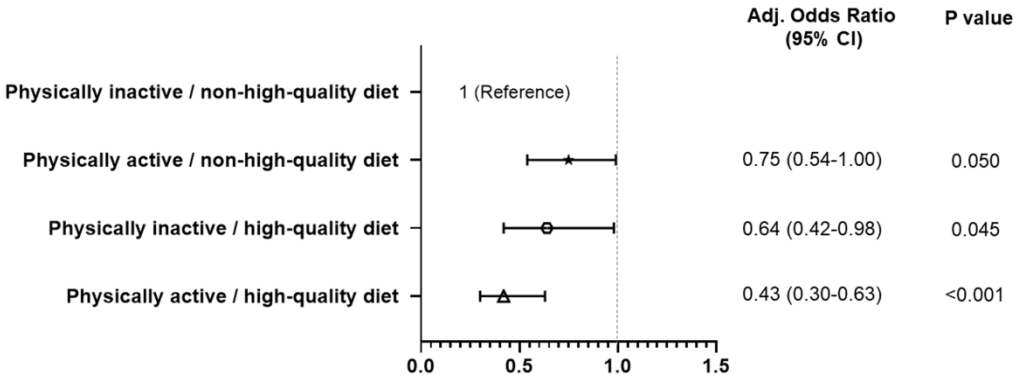
Accepted Article



hep_32207_f1d.tif

This article is protected by copyright. All rights reserved

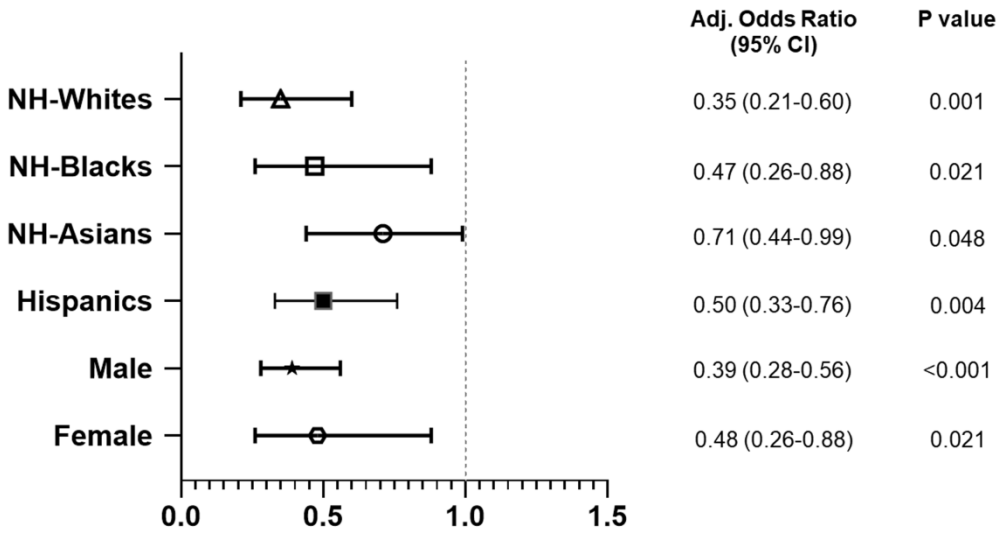
Accepted Article



hep_32207_f2.tif

This article is protected by copyright. All rights reserved

Accepted Article



hep_32207_f3.tif

This article is protected by copyright. All rights reserved