Protocol: Substituting red and processed meat, poultry, and fish with legumes and risk of non-alcoholic fatty liver disease in a large prospective cohort

Fie Langmann Daniel Borch Ibsen Luke W. Johnston
Christina Catherine Dahm

2024-05-28

Study information

Title: Substituting red and processed meat, poultry, and fish with legumes and risk of non-alcoholic fatty liver disease in a large prospective cohort

Description

Western diets has been shown to cause a multitude of non-communicable diseases. Non-alcoholic fatty liver disease (NAFLD) is the leading chronic disease globally and strongly associated with Western dietary patterns. Based on a combined environmental and health related focus, legumes are increasingly being recommended as a meat substitute, and this study therefore investigates the association between substituting legumes for meats, poultry, and fish and the risk of developing NAFLD.

This research will use the UK Biobank Resource under Application Number 81520.

Objectives and hypotheses

Background

The EAT-Lancet reference diet was introduced in 2019 as a globally sustainable and healthy diet emphasizing plant-based proteins instead of animal-based proteins, e.g., with a recommendation of 100 g legumes daily (1–3). Legume consumption minimizes the risk of non-alcoholic fatty liver disease (NAFLD) in animals, by reducing build-up of fats in the liver (4–8). NAFLD is the most prevalent chronic liver disease in the Western countries with a prevalence of 15-45 % (9,10). The disease is caused by Western diets high in red meat, fats, and sugars, obesity, physical inactivity, and smoking (9–12). As legumes are a source of both carbohydrates and proteins, research frequently compares legumes with other carbohydrate sources (13–17). Observed associations for foods may be due to the specific foods themselves. However, the associations could also represent the effect of the dietary pattern in which the specific food is consumed. When individuals limit intake of certain food groups, they will most often increase the intake of certain other food groups, in an otherwise stable diet (18,19). The association of such food intake will not only be caused by the increased consumption of one food, but rather the substitution effect including various foods. Replacing protein from animal sources with protein from plant sources has previously been associated with a substantially lower mortality rate and lower risk of NAFLD (5,20). Consumption of legumes in Western countries has been negligible to date and the impact of markedly increasing intakes of legumes on hepatobiliary and other diseases is understudied. Therefore, this study aims to investigate the effect of replacing meats, poultry, or fish with legumes and the risk of non-alcoholic steatohepatitis (NASH) or NAFLD contingent on potential confounding factors. As it might be more feasible for Western populations to include legumes and substitute dietary components that are not meats, this study will also aim to investigate the effect of a non-specific substitution of dietary non-meat components for legumes on the risk of NAFLD and NASH.

Hypothesis

- Replacing meats and poultry intakes with legumes is associated with a lower risk of NAFLD and NASH.
- Replacing fish intake with legumes is not associated with a lower risk of NAFLD and NASH.
- Replacing non-specific foods with legumes is not associated with a lower risk of NAFLD and NASH.

Design plan

Study type

Observational Study. Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, natural experiments, and regression discontinuity designs.

Blinding

No blinding is involved in this study.

Study design

Study population and setting

The initial recruitment of participants for the UK Biobank started in 2006 and ran until June 2010. Of 9.2 million people identified from the National Health Service registers and invited to participate in the study, 5.5% participated, approximately 500,000 participants, aged 37-73 years at baseline. The study protocol and more information are available elsewhere (21,22). All participants gave written, informed consent before baseline, and the study was approved by the National Information Governance Board for Health and Social Care and the National Health Service North West Multicentre Research Ethics Committee (reference number 06/MRE08/65).

Sampling plan

Existing data

Registration prior to analysis of the data

All data was collected in UK Biobank prior to access and data analyses.

Explanation of existing data

The UK Biobank is a large national cohort of participants in the UK, with data collected in a standardized format the the sole purpose of providing a data resource for researchers to use for health research. All information about collection procedures, limitations, and sources of bias are well established and described by the UK Biobank resource.

Data collection procedures

Study population and setting

At baseline, participants provided detailed information on several sociodemographic, physical, lifestyle, and health-related characteristics via self-completed touch-screen questionnaires and a computer assisted personal interview (23). Professionally trained staff did physical, anthropometric, and biomedical measures following standardized procedures (21). Diet was assessed through a touchscreen questionnaire at baseline and a 24-hour dietary assessment tool designed for the UK Biobank study. The 24-hour dietary assessment tool was completed by participants up to five times and 210,965 individuals completed one or more 24-hour dietary assessments (24).

Assessment of diet

The Oxford WebQ was designed as an internet based 24-hour dietary assessment tool for measuring diet on repeated occasions. The questionnaire is a short set of food frequency questions on commonly eaten food groups in the British population on the day before. The questionnaire aims to measure the type and quantity of food and beverages consumed in the last 24 hours and estimate nutrients from the entered information through the UK Nutrient Databank Food Composition Tables (25,26). The Oxford WebQ was compared with interviewer-administered 24-hour dietary recalls and validated for macronutrients and total energy intake using recovery biomarkers and compared with a single food frequency questionnaire (25,27,28). Recently, the Oxford WebQ nutrient calculation was updated to provide more detailed information on nutrient intakes and to incorporate new dietary variables (26). Participants recruited between April 2009 and September 2010 completed the Oxford WebQ at baseline (n=70,747). The Oxford WebQ was not available until April 2009 and participants recruited before that date who provided a valid email address were invited to complete the four subsequent 24-hour dietary assessments online (29).

Legumes

Legume consumption will be estimated based on participants' reported diets from the self-administered online 24-hour dietary assessments, the Oxford WebQ. Consumption of legumes and pulses will be based on total weight by food group intakes estimated from participants' responses in the Oxford WebQ. Despite the high detail level of the Oxford WebQ, a single 24-hour dietary assessment cannot capture habitual intake of legumes in a UK setting (30). Therefore, this study will include varying numbers of 24-hour dietary assessments to ensure that we capture the usual intake of legumes. Usual intake will be estimated as a daily average across available 24-hour dietary assessments.

Meat, poultry, and fish

Consumption of red and processed meats, poultry, and fish will be based on total weight by food group intakes estimated from participants' responses to the Oxford WebQs. Red and processed meat will be defined as beef, pork, lamb, and other meats including offal, and processed meat including sausages, bacon, ham, and liver paté. Poultry will be defined as poultry with or without skin, and fried poultry with batter or breadcrumbs. Fish will be defined as oily fish, white fish and tinned tuna, fried fish with batter or breadcrumbs, and shellfish.

Non-alcoholic fatty liver disease

Incident cases of NASH and NAFLD will be assessed through linkage to the National Health Service registers where diagnosis after hospital admission or primary care visits are coded according to the International Classification of Diseases and Related Health Problems (ICD-10) (31). In 2023, NAFLD underwent a change in nomenclature to metabolic dysfunction-associated steatotic liver disease (MASLD) to shift focus towards the metabolic factors underlying the disease and not merely the lack of alcohol consumption and to aid the prevention and early diagnoses of MASLD (32). However, as data was collected prior to this change in nomenclature, cases will be classified based on the ICD-10-codes and diagnosis criteria for NAFLD and NASH. To avoid assuming that NAFLD and MASLD are the same, this study will define incident cases of NAFLD based on ICD-10-code for NAFLD, K76.0, at first admission to hospital, while incident cases of NASH are diagnosed based on ICD-10-code, K75.8 (33,34).

Covariates

Directed acyclic graphs were used to illustrate the potential and known association between covariates of the association between substituting meat, poultry or fish for legumes and development of NAFLD and NASH. Information on covariates will include information on all other dietary components based on total weight by food group intakes as g/week retrieved from the Oxford WebQ (fruits, vegetables, cereal products, dairy products, egg products, nuts, mixed dishes, condiments, added sugar and sweets, non-alcoholic beverages, and alcoholic beverages), sex (male, female), age (years), alcohol consumption (g ethanol/day), ethnic group (white, mixed or other, Asian, black), socioeconomic status (Townsend deprivation score [quintiles], educational level), geographical region of recruitment (ten UK regions), cohabitation (alone, with spouse/partner, with other non-partner, no answer), anthropometry (BMI [kg/m2]), physical activity (low [918 MET-min/week], moderate [918-3706 MET-min/week], high [3706 MET-min/week], and unknown) (35), smoking status (never, former, current 1-15 cigarettes per day, current 15 cigarettes per day, and smoking status unknown); metabolic risk factors (self-reported [yes or no/unknown] diagnoses of diabetes, hypertension, stroke, myocardial infarction, gallbladder disease, alcoholic liver disease or high cholesterol); family history of metabolic risk factors (self-reported [yes or no/unknown] diagnoses of diabetes, hypertension, stroke, or heart disease), and cancer (yes, no, unknown).

Sample size

For this cohort study, only participants with two or more 24 h dietary assessments will be included in the analyses, while missing information on covariates will be filled in by multiple imputations where applicable. Individuals with alcohol intakes exceeding the 90th percentile will be excluded from the study to mitigate the risk of misclassifying the outcome by eliminating those with alcoholic fatty liver disease.

Sample size rationale

Variables

Measured variables

Exposures: g/week consumption of legumes, meats, poultry, and fish

Outcomes: non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH)

Covariates: g/week consumption of fruits, vegetables, cereal products, dairy products, egg products, nuts, mixed dishes, condiments, added sugar and sweets, non-alcoholic beverages, and alcoholic beverages. Sex, age, ethnic group, Townsend deprivation score, educational level, geographical region of recruitment (ten UK regions), cohabitation, BMI, physical activity, smoking status, own and family history of metabolic risk factors, and cancer.

Analysis plan

Statistical models

Main analyses

Standard summary statistics will be performed to describe the distribution of baseline characteristics and food consumption as an average g/day based on participants' 24-h WebQ responses. Multivariable adjusted Cox Proportional Hazards regression models will be used to estimate the hazard ratios for NAFLD based on replacing meats, poultry, or fish with legumes.

- Replacing red and processed meats, poultry, or fish with legumes (e.g., per 80 g/week)

Age will be used as the underlying time scale in the analyses. Follow-up time will begin with participants' last completed Oxford WebQ. As participants in UKB are still followed-up today, participants will be right censored at the date of the most recent registry update of full follow-up for the outcomes. Otherwise, censoring will occur at the event of death, loss to follow-up from the study, or date of diagnosis of NAFLD or NASH, whichever comes first. The substitution analyses will be conducted with different adjustment levels.

Model 1 will be minimally adjusted for strata of age at recruitment (>45, 45-49, 50-54, 55-59, 60-64, 65 years) and geographical region of recruitment (ten UK regions), sex, and intake of all other dietary components apart from the substitute components (red and processed meats; poultry; fish). When substituting g legumes/day, the unit for all dietary components will be g/day and the analyses will be adjusted for total amount of consumed foods in g/day.

Model 2 will be further adjusted for alcohol consumption (g/day), ethnic group (white, mixed, Asian, black), socioeconomic status (Townsend deprivation score, educational level), cohabitation (alone, with partner, with other non-partner, unknown), physical activity (low [918 MET-min/week], moderate [918-3706 MET-min/week], high [3706 MET-min/week], and unknown), smoking status (never, former, current 1-15 cigarettes per day, current 15 cigarettes per day, and unknown), own and family history of metabolic risk factors (yes or no/unknown), and cancer (yes, no, unknown).

Model 3 will further adjust for anthropometry (BMI 30 kg/m2), as obesity may either confound or mediate the association between replacing red and processed meats, poultry, or fish with legumes and risk of NAFLD

Secondary and sensitivity analyses

To evaluate the association between overall legume intake and NAFLD risk, hazard ratios associated with a 80 g/week increase in legumes will be estimated following adjustment levels in model 2. A non-specific substitution, adjusted like model 2, will also be conducted to evaluate the association between substituting 80 g/week of legumes for any other dietary component.

Peas are increasingly used as a plant-based meat alternative in the food industry (36). Despite this, peas are also included in the NHS 5 A Day recommendations for fruit and vegetables consumption in the UK (37). Therefore, in secondary analyses consumption of legumes will include participants self-reported intake of legumes and pulses together with consumed peas. The amount of peas consumed will be estimated based on participants' reported daily portions consumed with a portion size of peas weighing 80 g (38).

To evaluate the robustness of the main analyses, sensitivity analyses will include varying numbers of Oxford WebQ returns, and removal of participants with increased serum levels of alanine-aminotransferase (>40 U/L) and aspartate-aminotransferase. Sensitivity analyses will further include removal of participants in the highest 10% percentile of alcohol intake.

All analyses will be conducted in R with a significance level of 5%.

Inference criteria

All analyses will be evaluated based on two-sided p-values. Values below 5% are classified as statistically significant. Inference on relevance and significance and the evaluation of, whether a result is meaningful will be based on the size of the estimate, and confidence intervals containing 1 or 0, for ratios or absolute measures respectively.

References

1. Loken D B. Diets for a better future [Internet]. The EAT-Lancet Commission on Food, Planet, Health; 2020. Available from: https://eatforum.org/diets-for-a-better-future-report/

2. Springmann M, Spajic L, Clark MA, Poore J, Herforth A, Webb P, et al. The healthiness and sustainability of national and global food based dietary guidelines: Modelling study. BMJ [Internet]. 2020;370:m2322. Available from: https://www.bmj.com/content/bmj/370/bmj.m2322.full.pdf

3. Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, et al. Food in the anthropocene: The EAT-lancet commission on healthy diets from sustainable food systems. Lancet. 2019;393(10170):447–92.

4. Hong J, Kim S, Kim HS. Hepatoprotective effects of soybean embryo by enhancing adiponectin-mediated AMP-activated protein kinase pathway in high-fat and high-cholesterol diet-induced nonalcoholic fatty liver disease. J Med Food. 2016;19(6):549–59.

5. Zhang S, Yan Y, Meng G, Zhang Q, Liu L, Wu H, et al. Protein foods from animal sources and risk of nonalcoholic fatty liver disease in representative cohorts from north and south china. Journal of Internal Medicine [Internet]. 2023;293(3):340–53. Available from: https://www.embase.com/search/results?subaction=vi http://dx.doi.org/10.1111/joim.13586

6. Bouchenak M, Lamri-Senhadji M. Nutritional quality of legumes, and their role in cardiometabolic risk prevention: A review. J Med Food. 2013;16(3):185–98.

7. Son Y, Jang MK, Jung MH. Vigna nakashimae extract prevents hepatic steatosis in obese mice fed high-fat diets. J Med Food. 2014;17(12):1322–31.

8. Rigotti A, Marzolo MP, Ulloa N, González O, Nervi F. Effect of bean intake on biliary lipid secretion and on hepatic cholesterol metabolism in the rat. Journal of lipid research. 1989;30:1041–8.

9. Benedict M, Zhang X. Non-alcoholic fatty liver disease: An expanded review. World J Hepatol. 2017;9(16):715–32.

10. 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. Hepatology. 2016;64(1):73–84.

11. Younossi ZM, Loomba R, Rinella ME, Bugianesi E, Marchesini G, Neuschwander-Tetri BA, et al. Current and future therapeutic regimens for nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Hepatology. 2018;68(1):361–71.

12. Al-Dayyat HM, Rayyan YM, Tayyem RF. Non-alcoholic fatty liver disease and associated dietary and lifestyle risk factors. Diabetes Metab Syndr. 2018;12(4):569–75.

13. Sievenpiper JL, Kendall CW, Esfahani A, Wong JM, Carleton AJ, Jiang HY, et al. Effect of non-oil-seed pulses on glycaemic control: A systematic review and meta-analysis of randomised controlled experimental trials in people with and without diabetes. Diabetologia. 2009;52(8):1479–95.

14. Jayalath VH, Souza RJ de, Sievenpiper JL, Ha V, Chiavaroli L, Mirrahimi A, et al. Effect of dietary pulses on blood pressure: A systematic review and meta-analysis of controlled feeding trials. Am J Hypertens. 2014;27(1):56–64.

15. Belski R, Mori TA, Puddey IB, Sipsas S, Woodman RJ, Ackland TR, et al. Effects of lupin-enriched foods on body composition and cardiovascular disease risk factors: A 12-month randomized controlled weight loss trial. Int J Obes (Lond). 2011;35(6):810–9.

16. Linlawan S, Patcharatrakul T, Somlaw N, Gonlachanvit S. Effect of rice, wheat, and mung bean ingestion on intestinal gas production and postprandial gastrointestinal symptoms in non-constipation irritable bowel syndrome patients. Nutrients. 2019;11(9).

17. Lee YP, Puddey IB, Hodgson JM. Protein, fibre and blood pressure: Potential benefit of legumes. Clin Exp Pharmacol Physiol. 2008;35(4):473–6.

18. Ibsen DB, Laursen ASD, Würtz AML, Dahm CC, Rimm EB, Parner ET, et al. Food substitution models for nutritional epidemiology. The American Journal of Clinical Nutrition. 2020;113(2):294–303.

19. Ibsen DB, Dahm CC. Food substitutions revisited. The American Journal of Clinical Nutrition. 2022;116(5):1195–8.

20. Song M, Fung TT, Hu FB, Willett WC, Longo VD, Chan AT, et al. Association of animal and plant protein intake with all-cause and cause-specific mortality. JAMA Internal Medicine. 2016;176(10):1453–63.

21. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3):e1001779.

22. Biobank U. UK biobank: Protocol for a large-scale prospective epidemiological resource [Internet]. Vol. 2021. UK Biobank Coordinating Centre; 2007. Available from: https://www.ukbiobank.ac.uk/media/gnkeyh2q/study-rationale.pdf

23. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. Am J Epidemiol. 2017;186(9):1026–34.

24. Biobank U. Data-field 105010 [Internet]. 2021. Available from: https://biobank.ndph.ox.ac.uk/crystal/field.cgi?id=105010

25. Liu B, Young H, Crowe FL, Benson VS, Spencer EA, Key TJ, et al. Development and evaluation of the oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. Public Health Nutr. 2011;14(11):1998–2005.

26. Perez-Cornago A, Pollard Z, Young H, Uden M van, Andrews C, Piernas C, et al. Description of the updated nutrition calculation of the oxford WebQ questionnaire and comparison with the previous version among 207,144 participants in UK biobank. Eur J Nutr. 2021;60(7):4019–30.

27. Greenwood DC, Hardie LJ, Frost GS, Alwan NA, Bradbury KE, Carter M, et al. Validation of the oxford WebQ online 24-hour dietary questionnaire using biomarkers. Am J Epidemiol. 2019;188(10):1858–67.

28. Carter JL, Lewington S, Piernas C, Bradbury K, Key TJ, Jebb SA, et al. Reproducibility of dietary intakes of macronutrients, specific food groups, and dietary patterns in 211 050 adults in the UK biobank study. J Nutr Sci. 2019;8:e34.

29. Kelly RK, Watling CZ, Tong TYN, Piernas C, Carter JL, Papier K, et al. Associations between macronutrients from different dietary sources and serum lipids in 24 639 UK biobank study participants. Arterioscler Thromb Vasc Biol. 2021;41(7):2190–200.

30. FAO. Dietary assessment: A resource guide to method selection and application in low ressource settings. [Internet]. Food; Agriculture Organiszation of the United Nations; 2018. Available from: https://www.fao.org/3/i9940en/I9940EN.pdf

31. Digital N. NHS data model and dictionary international classification of diseases (ICD) [Internet]. Vol. 2021. The National Health Service,; 2021. Available from: https://datadictionary.nhs.uk/supporting_information/international_classification_of_diseases___icd_.html

32. Thornton J. Associations rename fatty liver disease to reduce stigma. BMJ [Internet]. 2023;382:p1587. Available from: http://www.bmj.com/content/382/bmj.p1587.abstract

33. Biobank U. Hospital inpatient data [Internet]. UK Biobank; 2020. Available from: http://www.ukbiobank.ac.uk/

34. Zhang Z, Burrows K, Fuller H, Speliotes EK, Abeysekera KWM, Thorne JL, et al. Non-alcoholic fatty liver disease and vitamin d in the UK biobank: A two-sample bidirectional mendelian randomisation study. Nutrients. 2023;15(6).

35. Cassidy S, Chau JY, Catt M, Bauman A, Trenell MI. Cross-sectional study of diet, physical activity, television viewing and sleep duration in 233 110 adults from the UK biobank; the behavioural phenotype of cardiovascular disease and type 2 diabetes. BMJ Open [Internet]. 2016;6(3). Available from: https://bmjopen.bmj.com/content/6/3/e010038

36. Maningat CC, Jeradechachai T, Buttshaw MR. Textured wheat and pea proteins for meat alternative applications. Cereal Chemistry. 2022;99(1):37–66.

37. Services UNH. 5 a day portion sizes [Internet]. Vol. 2022. UK NHS; 2022. Available from: https://www.nhs.uk/live-well/eat-well/5-a-day/portion-sizes/

38. Services UNH. 5 a day: What counts? [Internet]. Vol. 2022. UK NHS; 2021. Available from: https://www.nhs.uk/live-well/eat-well/5-a-day/5-a-day-what-counts/