BMJ Open

Diet quality indices, genetic risk and risk of cardiovascular disease and mortality: a longitudinal analysis of 77,004 UK Biobank participants

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045362
Article Type:	Original research
Date Submitted by the Author:	29-Sep-2020
Complete List of Authors:	Livingstone, Katherine; Deakin University, Institute for Physical Activity and Nutrition Abbott, Gavin; Deakin University, Institute for Physical Activity and Nutrition Bowe, Steven; Deakin University - Geelong Waterfront Campus, Deakin Biostatistics Unit Ward, Joey; University of Glasgow, Institute of Health and Wellbeing Milte, CM ; Deakin University, Institute for Physical Activity and Nutrition McNaughton, Sarah; Deakin University - Geelong Waterfront Campus, Institute for Physical Activity and Nutrition
Keywords:	NUTRITION & DIETETICS, EPIDEMIOLOGY, PUBLIC HEALTH, Coronary heart disease < CARDIOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

TITLE

Diet quality indices, genetic risk and risk of cardiovascular disease and mortality: a longitudinal analysis of 77,004 UK Biobank participants

AUTHORS

Katherine M. Livingstone[1]*, Gavin Abbott[1], Steven J. Bowe[2], Joey Ward[3], Catherine Milte[1], Sarah A. McNaughton[1]

AFFILIATIONS

1, Deakin University, Institute for Physical Activity and Nutrition Research, School of Exercise and Nutrition Sciences, Geelong, VIC 3220, Australia, k.livingstone@deakin.edu.au, gavin.abbott@deakin.edu.au, catherine.milte@deakin.edu.au sarah.mcnaughton@deakin.edu.au

2, Deakin Biostatistics Unit, Deakin University, Geelong, VIC 3220, Australia, s.bowe@deakin.edu.au

3, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK, Joey.Ward@glasgow.ac.uk

CONTACT INFO:

Dr Katherine Mary Livingstone,

Deakin University, Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Melbourne, Melbourne Burwood Campus, 221 Burwood Highway, Victoria 3125, Australia

Tel: +61 3 92445416 Email: k.livingstone@deakin.edu.au

WORD COUNT: 4639

ABBREVIATIONS: Body Mass Index, BMI; Cardiovascular Disease, CVD; Healthy Diet Indicator, HDI; Myocardial Infraction, MI; Mediterranean Diet Score, MDS; Polygenic Risk Score, PRS; Recommended Food Score, RFS; Single Nucleotide Polymorphism, SNP

ABSTRACT

Objectives: To examine associations of three diet quality indices and a polygenic risk score with incidence of all-cause mortality, CVD-related mortality, myocardial infarction (MI) and stroke.

Design: Prospective cohort study.

Setting: The UK Biobank, UK.

Participants: 77,004 men and women (40-70 years) recruited between 2006 and 2010.

Main outcome measures: Cox proportional hazard ratios (HR) were used to estimate effects of diet quality and genetic risk on risk of all-cause mortality, CVD-related mortality, MI and stroke. Dietary intake, assessed using the Oxford WebQ, was used to calculate the Recommended Food Score (RFS), Healthy Diet Indicator (HDI) and Mediterranean Diet Score (MDS). A polygenic risk score was created from 300 single nucleotide polymorphisms associated with CVD to examine moderation effects.

Results: New deaths due to CVD (n=364) and all-cause (n=2,409), and MI (n=1,141) and stroke (n=748) events were identified during mean follow-ups of 7.9 and 7.8 years, respectively. The adjusted HR associated with one-point higher RFS was 0.96 (0.94, 0.98) for all-cause mortality, 0.94 (0.90, 0.98) for CVD-related mortality, 0.97 (95% CI: 0.95, 1.00) for MI and 0.94 (95% CI: 0.91, 0.98) for stroke. The adjusted HR for all-cause mortality associated with one-point higher HDI and MDS was 0.97 (0.93 to 0.99) and 0.95 (0.91 to 0.98), respectively. The adjusted HR associated with one-point higher MDS was 0.93 (95% CI: 0.87, 1.00) for stroke. There was little evidence of associations between HDI and risk of CVD-related mortality, MI or stroke. There was only evidence of an interaction between diet quality and genetic risk score for MI.

Conclusion: Higher diet quality (RFS, HDI and MDS) predicted lower risk of all-cause mortality, independent of genetic risk. Higher RFS was also associated with lower risk of CVD-related mortality and MI. These findings demonstrate the benefit of following a healthy diet, regardless of genetic risk.

Strengths and limitations of this study

- This large prospective population-based cohort included repeat dietary assessments, using a validated questionnaire, and hospital register data on CVD incidence and mortality.
- The creation of three contrasting diet quality indices informs the best practice design and implementation of food-based diet quality indices for assessing diet-disease relationships.
- The polygenic genetic risk score was created using 300 SNPs known to be associated with CVD and all-cause mortality.
- Although the present analysis is likely to be subject to self-selection bias associated with the number of participants who completed the dietary assessment and the low response rate, associations between demographic and behavioural risk factors and mortality in the UK Biobank have been shown to be comparable to those from national health survey data from England and Scotland.
- Further research in diverse populations is needed to investigate the applicability of different diet quality methodologies for examining CVD risk independent of genetic susceptibility.

INTRODUCTION

Cardiovascular disease (CVD) is one of the leading causes of morbidity and mortality worldwide.[1] As a multifactorial condition, CVD risk is attributable to a combination of genetic and behavioural influences.[2] With poor diet now a leading risk factor for non-communicable diseases,[3] further understanding of the role of diet on CVD risk is warranted.

The overall quality of diets is an emerging predictor of CVD events and mortality.[4 5] Diet quality indices, that score dietary intakes according to *a priori* knowledge,[6] have been used to investigate association between diet and CVD incidence and mortality.[4 5 7-11] These indices can capture different aspects of diet quality, for example being based on intakes for encouraged foods only (e.g. Recommended Food Score, RFS), a combination of foods and nutrients from dietary guidelines (e.g. Healthy Diet Indicator, HDI) or a dietary pattern identified as healthful (e.g. Mediterranean Diet Score, MDS). However, our understanding is limited by the use of contrasting diet quality methodologies and a paucity of studies comparing different indices in large prospective population-based cohorts. Comparison of contrasting diet quality indices will identify whether differences in these methodologies are important for understanding diet-disease associations and will inform the international standardisation of diet quality methodologies for assessing health outcomes.[5 7 12]

The role of diet and genetics on risk of CVD is an emerging area of research.[13 14] Prior to the accessibility to whole genome sequencing, most research focused on links between single nucleotide polymorphisms (SNPs) and CVD.[15-17] Recent research has shown that polygenic risk scores (PRS), that incorporate multiple SNPs, are a good indicator of risk for complex conditions, such as CVD,[14 18] although the extent to which they influence the association between diet quality and CVD risk is unclear. Further research is also needed to elucidate whether diet quality is a risk factor for CVD independent of genetic risk. Moreover, the longitudinal association between contrasting diet quality indices, genetic risk and different CVD subtypes is unknown. Thus, the aim of this study was to examine the prospective role of three diet quality indices (HDI, RFS and MDS) and a PRS on risk of stroke, myocardial infarction, CVD-related mortality and all-cause mortality. Findings will advance understanding of the applicability of diet quality indices for assessing CVD risk.

METHODS

Study design and participants

The UK Biobank is a population cohort of half a million individuals living in the United Kingdom that aimed to examine determinants of disease in middle-aged adults.[19] Persons aged 40 to 69 years were identified from National Health Service patient registers and invited to participate. Individuals were invited to one of 22 assessment centres across England, Scotland and Wales between 2006 and 2011. At each centre, participants completed a touchscreen questionnaire to collect information on demographic characteristics, lifestyle behaviours and general health. The Oxford WebQ, a web-based 24-h dietary assessment tool, was introduced in 2009 to collect information on dietary intake.[20] Physical measurements (e.g., height and weight) were taken and participants provided blood and urine samples. Participants were followed up via linkage to health records and death registries. The UK Biobank received ethical approval from the Research Ethics Committee (Reference 11/NW/0382). Electronic signed consent was obtained from all participants. Participants were excluded from the present analysis if they i) did not identify as White British, ii) were ineligible based on previous history of CVD before entering the study, pregnancy, implausible physical activity data and CVD events during the study prior to completion of last dietary questionnaire, iii) had missing data for outcomes, exposures and covariates/moderators and v) had less than two timepoints of dietary data between February 2011 - June 2012. Results are reported according to the STROBE-NUT checklist for cohort studies.[21]

Study measures

Dietary intake

The Oxford WebQ was used to collect information on the frequency of consumption of 206 foods and 32 beverages during the previous 24 hours.[20 22 23] The Oxford WebQ is a 24-hour dietary questionnaire that has been validated against a traditional interviewer-administered multiple-pass 24-hour dietary recall and biomarkers for protein, potassium, and total sugar intake and total energy expenditure estimated by accelerometery.[23] Energy and nutrient intakes were calculated by multiplying the frequency of consumption of each food or drink by the standard portion size and energy and nutrient composition of each item.[24 25] Participants recruited between April 2009 and September 2010 completed the Oxford WedQ using the touchscreen at the assessment centre. Repeat Oxford WebQs were

collected via four online cycles between February 2011 to June 2012: February 2011 to April 2011 (online cycle 1); June 2011 to September 2011 (online cycle 2); October 2011 to December 2011 (online cycle 3); April 2012 to June 2012 (online cycle 4). The total period of available dietary data from the Oxford WebQ was 38 months (Apr 2009 - Jun 2012). Email invitations were sent on different days of the week to capture variation in dietary intakes and participants were given 3 days to complete the questionnaire for cycles 1 and 2 and 14 days for cycles 3 and 4.

To establish a baseline dietary intake in the present analysis, we calculated a mean dietary intake based on the four online Oxford WebQ cycles only. This was because the time between the 1st and 4th online cycle measurements was 16 months (Feb 2011 - Jun 2012) and was considered a more credible timeframe for an average baseline than the 38 months available from all five Oxford WebQ measurements. This resulted in a minimal sample loss (<10%) while providing a shorter dietary exposure period and a more consistent approach to the use of the dietary data by using only the online cycles of the OxfordWebQ. To better capture usual intake, we calculated average nutrient intakes, food group intakes and diet quality scores for participants who had two or more valid measurements for the four online cycles of the Oxford WebQ.

Diet quality

Information on food and beverage intakes from the Oxford WebQ were used to calculate three diet quality indices: the Recommended Food Score (RFS), which is based on intakes of encouraged foods only,[26] and the Healthy Diet Indicator (HDI), which scores intakes of a combination of foods and nutrients from dietary guidelines,[27] and the Mediterranean Diet Score (MDS), representing dietary patterns identified as healthful.[28] These indices were selected as they represent three contrasting diet quality methodologies that have been applied internationally to assess diet-disease associations.[9 10 26 27 29 30]

The RFS is a food-based variety index designed to assess consumption of food groups encouraged in the dietary guidelines.[10] As detailed in Supplemental Table 1, food intakes were scored according to five food groups: fruits (7 items), vegetables (7 items), wholegrains (2 items), lean meat and alternatives (3 items) and reduced fat dairy products (2 items). Scoring was based on the RFS designed by Kant and Graubaud,[26] and has been

used elsewhere.[31 32] We summed intakes of food items within each group to create a total intake for each food group. Food groups were then assigned a score of 1 if they were consumed above the minimum amount threshold: 15 g/d for non-beverages and 30 g/d for beverages. Intakes below these thresholds were scored 0. Scores ranged between 0 and 21, with higher scores indicating a higher quality diet and a wider consumption of recommended foods.[33]

The HDI is a food- and nutrient-based index designed to reflect consumption of foods recommended for a healthy diet by the World Health Organisation.[34] The original HDI was developed and validated in 1997 based on the 1990 World Health Organisation's dietary recommendations for the prevention of chronic disease.[35] We adapted a 12-point Healthy Diet Score designed by Maynard et al.[27] to reflect adherence to the 2020 World Health Organisation healthy diet fact sheet.[34] As cholesterol intake is not part of the 2020 recommendations and information on its intake was not available in the UK Biobank, we used an 11-item score that included the following groups: saturated fat; poly-unsaturated fat; protein; total carbohydrates; dietary fibre; fruits and vegetables; pulses and nuts; total non-milk extrinsic sugars; fish; red meat and meat products; and calcium. Data on intake of non-milk extrinsic sugars was not available in the UK Biobank and so we adapted the HDI to score intakes of total sugars instead. Criteria for scoring was based on cut points detailed in Supplemental Table 2. We assigned intakes within the cut offs a score of 1 and those outside of the cut offs were assigned a score of 0. The total score ranged from 0 to 11, with a higher score reflecting a higher diet quality (Supplemental Table 2).

The MDS is a food- and nutrient-based score designed to reflect adherence to a Mediterranean style diet. The present study used the 9-item index developed and validated by Trichopoulou et al. as it is the first and most widely used version of the MDS.[36 37] Food and nutrient intakes were scored according to nine components: vegetables, legumes, fruits and nuts, cereals, fish and seafood, monounsaturated fats to saturated fats ratio, dairy products, meat and meat products and alcohol (Supplemental Table 3). As used by Trichopoulou et al.,[36] we used sex-specific median intakes as cut off points for intakes of each component. A score of 1 was assigned to participants whose intake of vegetables, legumes, fruits and nuts, cereals, fish and seafood and monounsaturated: saturated fats was above the median. A score of 1 was assigned to intake of dairy products, meat and meat

products below the median. For alcohol, a score of 1 was assigned for low to moderate intake (intake of no more than 2 times/day). A score of 0 was assigned for no alcohol intake or intake greater than 2 times per day.[38] Total MDS score ranged from 0 to 9, with higher scores reflecting better alignment to the Mediterranean diet.

Cardiovascular events and mortality

Mortality status and causes of death were determined by linkage of data with the UK National Death Index (NDI) using the most recent available data from the UK Biobank (September 2020). The accuracy of the NDI for identifying CVD deaths has been established previously.[39] CVD-related mortality was estimated from 2006 International Classification of Diseases 10th revision (ICD-10) codes in death certificates. CVD-related mortality was identified using ICD codes I05-I89. CVD events were recorded between enrolment (1999–2000) and the most recent inpatient hospital data available from the UK Biobank (September 2020). Incident MI (ST-Elevation Myocardial Infarction and Non-ST-Elevation Myocardial Infarction) and stroke (ischaemic, intracerebral haemorrhage, and subarachnoid haemorrhage) were available from algorithms provided by the UK Biobank.[40 41] Algorithms were produced to reliably identify incidence of selected illnesses through consideration of hospital and death register data. The adjudication of "algorithmically defined" outcomes for MI and stroke are detailed elsewhere.[40 41] A censoring data of 4 March 2020 was used for all outcomes. This date was chosen due to a spike in deaths from 5 March onwards, which is likely to correspond to increasing deaths due to COVID-19 recorded in the UK.[42]

Polygenic risk score

We used the March 2018 release of the imputed genetic data from UK Biobank (downloaded 11 November 2019). From the resulting dataset, we excluded those who self-reported ancestry other than white British, those who were missing more than 10% of the genetic data and those who were defined by UK Biobank as being heterozygosity outliers. Additionally, for every pair of who were individuals who were second cousins or closer (i.e. those with a kinship coefficient > 0.042) one was excluded at random. We used information on 300 single nucleotide polymorphisms (SNPs) known to be associated with coronary artery disease[43] to create a PRS for CVD for each individual.[44] Evidence indicates that a

genetic risk score estimated from these 300 SNPs is associated with traditional risk factors for CVD, such as type 2 diabetes and hypertension, contributes to the development of CVD-related conditions that have their origins in atherosclerosis, such as peripheral arterial disease and stroke, and is associated with premature mortality.[43] The PRS was estimated by generating the sum of the number of risk alleles present at each locus and weighting by the log of the odds for that locus[18] estimated from the list of 300 SNPs using the plink "–score" command – with no-mean-imputation flag. For participants included in the final study sample, PRS were transformed to standardised Z scores and were treated as a continuous variable in all modelling.

Demographic and health information

Information on demographics, medical history and health behaviours were collected using interview-administered questionnaires at recruitment and follow ups. Participant age at recruitment and sex were self-reported. No adjustments were made for discrepancies between self-reported sex and genetic sex. Education was assessed by asking "Which of the following qualifications do you have? (You can select more than one)," with the options college or university degree, A levels or equivalent, O levels or GCSEs or equivalent, CSEs, NVQ/HND/HNC, other professional qualifications (e.g., nursing or teaching). We operationalised this into 5 categories based on the highest level of education: i) college or university degree, ii) all professional qualification (NVQ/HND/HNC, other professional qualifications), iv) A levels or equivalent, v) O levels, GCSEs or equivalent or CSEs and v) none of the above or prefer not to answer. The Townsend deprivation index, a composite measure of deprivation based on unemployment, non-car ownership, non-home ownership, and household overcrowding,[45] was estimated from the preceding national census data, with each participant assigned a score corresponding to the postcode of their home dwelling and a negative value representing high socioeconomic status. We operationalised the Townsend deprivation index as quintiles.

Information on smoking (never, previous and current smoker), previous doctor diagnosis of any type of diabetes or a CVD event (yes, no) and use of medication (anti-hypertensive, lipid-lowering or exogenous hormones or diabetes; yes, no) were collected. We created a binary variable for family history (of father, mother and siblings) of CVD and related diseases

(yes, no). Body Mass Index (BMI) was calculated as weight/height^2 collected during the Assessment Centre visit. We created a binary variable to indicate overweight or obese according to standard World Health Organisation cut offs.[46] Physical activity was estimated using Metabolic Equivalents (METs), the ratio of a person's working metabolic rate relative to their resting metabolic rate. One MET was defined as the energy cost of sitting quietly and is equivalent to a caloric consumption of 1kcal/kg/hour. We used standard cut-offs to categorised participants as meeting physical activity guidelines of 150 min per week if their METs were ≥ 600 MET-min/week.[47]

Statistical analysis

Complete case analysis was used. We investigated missingness by comparing demographic characteristics of the excluded sample with the analytic sample. Descriptive analyses included mean (SD) for continuous variables and number (%) for categorical variables. We created sample-based tertiles of diet quality for RFS, HDI and MDS for descriptive purposes only. Unadjusted linear regression analyses were used to examine intakes of encouraged food groups and total energy and nutrient intakes across tertiles of diet quality indices.

We used multivariable Cox proportional hazard regression models to estimate hazard ratios (HR) and 95% Confidence Intervals (CI) of all-cause mortality, CVD mortality and risk of CVD events (MI and stroke) according to each diet quality index separately (RFS, HDI and MDS). We treated diet quality indices as continuous independent variables. CVD events and mortality were treated as time-to-event outcome/dependent variables. We estimated the duration of follow up as the time between the last day of dietary data and the first event of either an MI, stroke, mortality, or the censoring date (4 March 2020). In participants who had multiple events during the study period, the first event date was used. We adjusted the Cox regression analyses for covariates identified using a directed acyclic graph (Supplemental Figure 1). These included age (continuous), sex, deprivation (categorical), smoking status (categorical), physical activity (continuous), medication use (binary), family history of CVD (binary) and energy intake (continuous). The Cox proportional hazards models also included PRS as an independent variable and were additionally adjusted for the first 8 principal components of ancestry and genotyping batch.[14] We included an interaction term in the models to test for statistical interaction between each diet quality score and

PRS. Interactions were further inspected by conducting post-hoc estimation of the effects of diet quality indices on events at 'low' and 'high' PRS score of -1 and +1 which, given PRS was a standard score, represent minus and plus one SD of PRS. Data were analysed using Stata (version 16.0; StataCorp., College Station, TX, USA). To address possible reverse causation, sensitivity analyses excluded deaths and incident cases of MI and stroke within the first 2 years of follow up.

Patient and public involvement

The development of the research question or outcome measures was not informed by patients' priorities, experience, or preferences. No patients were involved in the design and conduct of the present study. There are no plans to disseminate the results to study participants.

RESULTS

Of the 502,536 participants who were recruited into the UK Biobank, 425,529 participants were excluded based on being not white British (n=92,907), having unusable genetic data (n=1,459), being ineligible (n=23,215) or missing dietary or covariate data (n=307,951; Supplemental Figure 2). Excluded participants were similar in age and sex to the included sample, with somewhat higher BMI and rates of smoking and deprivation (Supplemental Table 4). A total of 77,004 participants were included in the present analysis (Table 1). Mean age at recruitment was 56.2 (SD 7.8) years and 55% were female. Forty-eight per cent of participants had a colleague or university degree, most were experiencing low to mid deprivation (67%), had never smoked (58%) and had a family history of CVD (74%; Table 1). Fifty-nine per cent of the participants were overweight or obese and 85% met physical activity guidelines.

Mean RFS, HDI and MDS were 6.78 (SD 2.40), 3.57 (SD 1.26) and 5.31 (SD 1.04), respectively. Intake of fruits, vegetables, wholegrains and lean meat were higher with increasing tertile of RFS, HDI and MDS (Table 2). Intake of low-fat dairy were higher with increasing tertile of RFS and HDI and lower with increasing tertile of MDS. Intakes of total fat, saturated fat, carbohydrates and sugars were higher with increasing tertile of RFS, HDI and MDS. Intakes of protein were lower with increasing tertile of HDI and MDS, while intakes were higher with

increasing tertile of RFS. Intakes of PUFA were lower with increasing tertile of RFS, while intakes were higher with increasing tertile of HDI and MDS (Table 2).

Table 1. Baseline characteristics of participants in the UK Biobank

Characteristic	Overall N (%)	Males N (%)	Females N (%)
n	77,004	34,984 (45.4)	42,020 (54.5)
Age at recruitment (years), Mean ± SD	56.2 ± 7.8	57.0 ± 7.8	55.6 ± 7.7
Highest level of education			
College or University degree	36,709 (47.8)	17,271 (49.4)	19,438 (46.3)
A levels/AS levels or equivalent	10,580 (13.8)	4,390 (12.6)	6,190 (14.8)
O levels/GCSE/CSEs or equivalent	17,669 (23.0)	7,126 (20.4)	10,543 (25.1)
Professional qualifications	7,467 (9.7)	4,062 (11.6)	3,405 (8.1)
None/Prefer not to answer	4,454 (5.8)	2,082 (6.0)	2,372 (5.7)
Townsend Deprivation Index			
Least deprived	18,129 (23.5)	8,611 (24.6)	9,518 (22.7)
2nd least deprived	17,227 (22.4)	7,910 (22.6)	9,317 (22.2)
Medium deprivation	16,067 (20.9)	7,158 (20.5)	8,909 (21.2)
2nd most deprived	14,900 (19.3)	6,549 (18.7)	8,351 (19.9)
Most deprived	10,681 (13.9)	4,756 (13.6)	5,926 (14.1)
Smoking			
Never smoked	44,856 (58.3)	18,849 (53.9)	26,007 (61.9)
Ex-smoker	27,184 (35.3)	13,471 (38.5)	13,714 (32.6)
Current smoker	4,964 (6.5)	2,664 (7.6)	2,300 (5.5)
Body Mass Index (kg/m^2), Mean ± SD[1]	26.5 ± 4.4	27.1 ± 3.9	26.0 ± 4.7
Waist circumference (cm), Mean ± SD[1]	88.1 ± 13.0	95.2 ± 10.8	82.3 ± 11.6
Total PA (MET min), Mean ± SD	2477 ± 2326	2542 ± 2439	2423 ± 2227
Medication use[2]	16,573 (21.5)	9,713 (27.8)	6,860 (16.3)
Family history of CVD	57,211 (74.3)	25, 076 (71.7)	32,135 (76.5)
Energy Intake (kJ/day), Mean ± SD	8853 ± 2172	9574 ± 2253	8252 ± 1903

Townsend Deprivation Index is a composite measure of deprivation based on unemployment, non-car ownership, non-home ownership, and household overcrowding.

1, Data on Body Mass Index and Waist Circumference were available in n=76,901 and n=76,950 respectively.

2, Medication use was restricted to lipid lowering or blood pressure.

Table 2. Baseline dietary intakes by tertile of diet quality index in the UK Biobank (n=77,004)[1]

	Recommended Food Score			Healthy Diet Indicator			Mediterranean Diet Score		
	T1	**T2**	**T3**	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**
Diet quality	4.2 (1.1)	6.8 (0.6)	9.5 (1.3)	2.3 (0.6)	3.7 (0.3)	5.1 (0.7)	4.4 (0.6)	5.5 (0.2)	6.5 (0.5)
Food group intake (g/day)									
Whole fruit	81.6 (76.2)	152.7 (98.5)	231.1 (133.6)	95.3 (78.8)	156.8 (106.9)	225.9 (138.7)	115.4 (98.7)	170.7 (124.0)	198.7 (132.0)
Vegetables	120.0 (89.6)	201.8 (113.0)	286.1 (149.7)	146.7 (99.0)	204.7 (127.4)	267.5 (157.7)	158.5 (114.3)	220.4 (136.3)	250.6 (150.0)
Wholegrains	32.8 (38.2)	51.5 (46.3)	65.8 (52.9)	36.2 (38.8)	50.9 (46.2)	65.8 (54.7)	37.3 (39.0)	54.5 (49.1)	64.7 (54.2)
Lean meats	51.1 (45.4)	66.0 (48.9)	79.9 (54.4)	59.9 (48.0)	66.5 (51.0)	71.6 (53.9)	57.3 (48.1)	71.2 (51.8)	73.4 (52.8)
Low-fat dairy	18.3 (37.3)	29.4 (43.7)	45.0 (51.9)	22.7 (40.6)	30.9 (44.7)	40.6 (51.2)	31.8 (48.5)	32.7 (46.7)	27.6 (41.0)
Total EI (kJ/day)	8572 (2224)	8798 (2080)	9204 (2158)	8244 (2032)	9022 (2112)	9474 (2195)	8785 (2234)	8862 (2106)	8946 (2121)
Nutrient intake (% energy)									
Total fat	34.5 (6.1)	33.1 (5.7)	31.8 (5.7)	34.4 (5.9)	33.2 (5.7)	31.4 (5.8)	34.2 (5.8)	32.9 (5.7)	31.7 (5.9)
Saturated fat	13.6 (3.1)	12.7 (2.8)	11.8 (2.8)	13.7 (3.0)	12.7 (2.8)	11.4 (2.7)	13.8 (2.9)	12.7 (2.7=8)	11.4 (2.7)
PUFA	6.18 (2.0)	6.08 (1.9)	6.06 (1.9)	5.87 (2.0)	6.16 (1.9)	6.37 (1.9)	5.91 (1.9)	6.16 (1.9)	6.37 (2.0)
Carbohydrate	45.8 (7.7)	47.9 (7.1)	50.0 (6.8)	44.8 (7.1)	48.2 (6.7)	51.6 (6.7)	46.7 (7.5)	48.0 (7.3)	49.5 (7.1)
Total sugars	20.1 (6.1)	22.6 (5.6)	25.6 (5.7)	20.4 (5.8)	22.9 (5.7)	25.5 (6.0)	21.9 (6.3)	22.8 (6.2)	23.9 (6.0)
Protein	15.4 (3.1)	15.8 (2.9)	16.2 (2.9)	16.4 (3.2)	15.6 (2.9)	15.2 (2.7)	16.0 (3.1)	15.8 (2.9)	15.4 (2.9)

EI, Energy intake

1, Unadjusted linear regression analyses were used to examine linear trend across tertiles of diet quality index; $p<0.001$ for all associations.

During a mean follow-up of 7.8 years (a total of 600,193 person-years), we observed 1,141 new MI events and 748 new stroke events. During a mean follow-up of 7.9 years (a total of 604,431 person-years), we observed 364 deaths due to CVD and 2,409 all-cause deaths. Of these, the majority of MI (72%) and stroke (60%) events, and CVD-related (72%) and all-cause (59%) deaths were in males.

RFS and risk of all-cause mortality, CVD-related mortality, MI and stroke

The adjusted HR associated with a one-point higher RFS was 0.96 (95% CI: 0.94, 0.98) for all-cause mortality, 0.94 (95% CI: 0.90, 0.98) for CVD-related mortality, 0.97 (95% CI: 0.95, 1.00) for MI and 0.94 (95% CI: 0.91, 0.98) for stroke (Table 3). When stratified by sex, associations were comparable in men, while there was only evidence of an association between RFS and all-cause mortality and stroke in females. The adjusted HR associated with a one-point higher PRS was 1.33 (95% CI: 1.25, 1.41) for MI; when stratified by sex, there was evidence of a stronger association in males. When an interaction term was added to the models, there was no evidence (at the $p<0.05$ level) of interaction between RFS and PRS for any outcomes (p-interaction=0.40 [all-cause mortality], p-interaction=0.77 [CVD-related mortality], p-interaction=0.17 [MI], and p-interaction=0.10 [stroke]). Effect sizes were consistent when deaths and incident cases of MI and stroke within the first 2 years of follow up were excluded (data not shown).

HDI and risk of all-cause mortality, CVD-related mortality, MI and stroke

The adjusted HR associated with a one-point higher HDI was 0.97 (95% CI: 0.93, 0.99) for all-cause mortality. There was little evidence of associations between HDI and risk of CVD-related mortality, MI, or stroke (Table 4). When stratified by sex, there was evidence of an association between HDI and all-cause mortality in males only. The adjusted HR associated with a one-point higher PRS was 1.33 (95% CI: 1.25, 1.41) for MI, which when stratified by sex, there was evidence of a stronger association in males. When an interaction term was added to the models, there was no evidence of interaction between HDI and PRS for other outcomes (p-interaction=0.66 [all-cause mortality], p-interaction=0.86 [CVD-related mortality] and p-interaction=0.17 [stroke]). There was some evidence of interaction

between HDI and PRS for MI events (p-interaction=0.049). While there was no evidence of an effect of HDI on MI for participants with low (-1 SD) PRS (HR=1.02 [95% CI: 0.95, 1.10], p=0.61), there was some evidence of an association between higher HDI and reduced risk of MI events for those with high (+1 SD) PRS (HR=0.93 [95% CI: 0.88, 0.99], p=0.017). Effect sizes were consistent when deaths and incident cases of MI and stroke within the first 2 years of follow up were excluded (data not shown).

MDS and risk of all-cause mortality, CVD-related mortality, MI and stroke

The adjusted HR associated with a one-point higher MDS was 0.95 (95% CI: 0.91, 0.98) for all-cause mortality and 0.93 (95% CI: 0.87, 1.00) for stroke (Table 5). There was limited evidence of associations between MDS and risk of CVD-related mortality and MI. When stratified by sex, there was evidence of an association between MDS and all-cause mortality and MI in males only. The adjusted HR associated with a one-point higher PRS was 1.33 (95% CI: 1.25, 1.41) for MI; when stratified by sex, there was evidence of a stronger association in males. When an interaction term was added to the models, there was no evidence of interaction between MDS and PRS for other outcomes (p-interaction=0.58 [all-cause mortality], p-interaction=0.72 [CVD-related mortality] and p-interaction=0.12 [stroke]). There was evidence of interaction between MDS and PRS for MI events (p-interaction=0.026). While there was no evidence of an effect of MDS on MI for those with low (-1 SD) PRS (HR=1.03 [95% CI: 0.94, 1.12], p=0.56) there was strong evidence of an association between higher MDS and reduced risk of MI events for those with high (+1 SD) PRS (HR=0.91 [95% CI: 0.85, 0.97], p=0.004). Effect sizes were consistent when deaths and incident cases of MI and stroke within the first 2 years of follow up were excluded (data not shown).

Table 3. Cox-proportional hazard ratios and 95% CI for risk of all-cause mortality, CVD-related mortality and CVD events with increasing Recommended Food Score (RFS) and polygenic risk score in participants from the UK Biobank

	Overall (n=77,004)				Males (n=34,984)				Females (n=42,020)			
	No of Events	**Hazard Ratio**	**95% CI**	**P-value**	**No of Events**	**Hazard Ratio**	**95% CI**	**P-value**	**No of Events**	**Hazard Ratio**	**95% CI**	**P-value**
All-cause mortality	2,409				1,416				993			
RFS		0.96	(0.94, 0.98)	<0.001		0.95	(0.93, 0.97)	<0.001		0.98	(0.95, 1.00)	0.08
Polygenic risk score		1.00	(0.96, 1.04)	0.93		1.02	(0.97, 1.07)	0.53		0.98	(0.92, 1.05)	0.57
CVD-related mortality	364				263				101			
RFS		0.94	(0.90,0.98)	0.007		0.93	(0.88, 0.98)	0.011		0.96	(0.88, 1.05)	0.34
Polygenic risk score		1.08	(0.98, 1.20)	0.13		1.10	(0.98, 1.25)	0.11		1.03	(0.85, 1.25)	0.79
Myocardial Infarction	1,141				822				319			
RFS		0.97	(0.95, 1.00)	0.048		0.97	(0.94, 1.00)	0.045		0.99	(0.94, 1.04)	0.57
Polygenic risk score		1.33	(1.25, 1.41)	<0.001		1.40	(1.31, 1.50)	<0.001		1.16	(1.04, 1.29)	0.008
Stroke	748				447				301			
RFS		0.94	(0.91, 0.98)	0.001		0.95	(0.91, 0.99)	0.018		0.94	(0.89, 0.99)	0.012
Polygenic risk score		1.02	(0.94, 1.10)	0.68		0.97	(0.89, 1.07)	0.56		1.08	(0.96, 1.21)	0.29

CVD, cardiovascular disease; Models were adjusted for age (continuous), sex (when not used to stratify), deprivation (categorical), smoking status (categorical), physical activity (continuous), medication use (binary), family history of CVD (binary), energy intake (continuous) and the first 8 principal components of ancestry and genotyping batch. All models include main effects of diet quality and polygenic risk score but not interaction terms.

Table 4. Cox-proportional hazard ratios and 95% CI for risk of all-cause mortality, CVD-related mortality and CVD events with increasing Healthy Diet Indicator (HDI) and polygenic risk score in participants from the UK Biobank

	Overall (n=77,004)				Males (n=34,984)				Females (n=42,020)			
	No of Events	Hazard Ratio	95% CI	P-value	No of Events	Hazard Ratio	95% CI	P-value	No of Events	Hazard Ratio	95% CI	P-value
All-cause mortality	2,409				1,416				993			
HDI		0.97	(0.93, 0.99)	0.041		0.95	(0.91, 0.99)	0.039		0.98	(0.93, 1.03)	0.46
Polygenic risk score		1.00	(0.96, 1.04)	0.92		1.02	(0.96, 1.07)	0.54		0.98	(0.92, 1.05)	0.58
CVD-related mortality	364				263				101			
HDI		0.99	(0.90, 1.08)	0.76		0.94	(0.85, 1.05)	0.28		1.10	(0.94, 1.30)	0.23
Polygenic risk score		1.08	(0.98, 1.20)	0.13		1.10	(0.98, 1.25)	0.11		1.03	(0.85, 1.25)	0.79
Myocardial Infarction	1,141				822				319			
HDI		0.96	(0.92, 1.01)	0.12		0.95	(0.89, 1.00)	0.06		1.00	(0.92, 1.10)	0.93
Polygenic risk score		1.33	(1.25, 1.41)	<0.001		1.40	(1.31, 1.50)	<0.001		1.16	(1.04, 1.29)	0.008
Stroke	748				447				301			
HDI		0.97	(0.91, 1.03)	0.25		0.96	(0.89, 1.04)	0.31		0.98	(0.89, 1.07)	0.63
Polygenic risk score		1.02	(0.94, 1.09)	0.68		0.97	(0.89, 1.07)	0.56		1.08	(0.96, 1.21)	0.19

CVD, cardiovascular disease; Models were adjusted for age (continuous), sex (when not used to stratify), deprivation (categorical), smoking status (categorical), physical activity (continuous), medication use (binary), family history of CVD (binary), energy intake (continuous) and the first 8 principal components of ancestry and genotyping batch. All models include main effects of diet quality and polygenic risk score but not interaction terms.

Table 5. Cox-proportional hazard ratios and 95% CI for risk of all-cause mortality, CVD-related mortality and CVD events with increasing Mediterranean Diet Score (MDS) and polygenic risk score in participants from the UK Biobank

	Overall (n=77,004)				Males (n=34,984)				Females (n=42,020)			
	No of Events	**Hazard Ratio**	**95% CI**	**P-value**	**No of Events**	**Hazard Ratio**	**95% CI**	**P-value**	**No of Events**	**Hazard Ratio**	**95% CI**	**P-value**
All-cause mortality	2,409				1,416				993			
MDS		0.95	(0.91, 0.98)	0.005		0.93	(0.88, 0.98)	0.004		0.97	(0.91, 1.03)	0.33
Polygenic risk score		1.00	(0.96, 1.04)	0.91		1.02	(0.97, 1.07)	0.51		0.98	(0.92, 1.05)	0.58
CVD-related mortality	364				263				101			
MDS		0.97	(0.88, 1.08)	0.60		0.94	(0.84, 1.06)	0.32		1.07	(0.88, 1.29)	0.49
Polygenic risk score		1.08	(0.98, 1.20)	0.12		1.10	(0.98, 1.25)	0.11		1.03	(0.85, 1.25)	0.79
Myocardial Infarction	1,141				822				319			
MDS		0.95	(0.90, 1.00)	0.06		0.94	(0.88, 1.00)	0.049		0.97	(0.88, 1.08)	0.64
Polygenic risk score		1.33	(1.25, 1.41)	<0.001		1.40	(1.31, 1.50)	<0.001		1.16	(1.04, 1.29)	0.008
Stroke	748				447				301			
MDS		0.93	(0.87, 1.00)	0.037		0.93	(0.85, 1.01)	0.10		0.93	(0.83, 1.04)	0.21
Polygenic risk score		1.02	(0.95, 1.09)	0.67		0.97	(0.89, 1.07)	0.57		1.08	(0.96, 1.21)	0.19

CVD, cardiovascular disease; Models were adjusted for age (continuous), sex (when not used to stratify), deprivation (categorical), smoking status (categorical), physical activity (continuous), medication use (binary), family history of CVD (binary), energy intake (continuous) and the first 8 principal components of ancestry and genotyping batch. All models include main effects of diet quality and polygenic risk score but not interaction terms.

DISCUSSION

This prospective population-based cohort study aimed to examine the association of three diet quality indices (RFS, HDI, MDS) and a genetic risk score with incidence of CVD and mortality. Our main findings were that higher RFS, HDI and MDS were associated with lower risk of mortality, regardless of genetic CVD risk. However, only the RFS showed evidence of lower risk of CVD-related mortality, MI and stroke, suggesting the applicability of the diet quality indices may depend on the health outcome in question. We also identified that increasing genetic risk of CVD was associated with MI only, which was strongest in males, suggesting that the underlying genetics of both MI and death may follow different pathways in males and females. Interaction analyses suggested that following a healthy diet may be of particular importance for reducing risk of MI in individuals with high genetic risk of CVD. Nonetheless, our findings demonstrate the benefit of following a healthy diet independent of genetic risk.

Our findings for reduced risk of all-cause mortality with higher diet quality are consistent with previous research on the MDS,[5 9 48 49] HDI[48 50] and RFS.[51] Moreover, a comparison of 10 diet quality indices in over 450,000 European adults showed that all indices examined were inversely associated with 10-year risk of all-cause mortality.[50] In the present study, the predictive role of diet quality on risk of all-cause mortality remained after adjusting for major non-modifiable determinants of all-cause mortality, including age, sex and family history of CVD. This highlights the importance of modifiable risk factors for death, regardless of whether the diet quality index is based on intakes of encouraged foods (i.e. RFS), foods and nutrients from dietary guidelines (i.e. HDI) or a dietary pattern identified as healthful (i.e. MDS). Moreover, the common elements across all three indices is the inclusion of food-based components, such as fruit and vegetables and lean meat and alternatives, rather than nutrients, affirming the value of food-based dietary guidelines in preventative healthcare rather a reductionist nutrient-based approach.[52]

Evidence for an association between diet quality indices and CVD risk is mixed.[5 7 11 14 32 53-55] Confirming our findings, large-scale studies in the UK population have shown independent associations between healthy diets and lifestyles and low genetic risk in reducing risk of CVD, with mixed results for interactions.[14 55] Only one study to date has used an overall diet quality index,[11] with comparable results to the present study, highlighting the potential to

include plant-based diet quality components when assessing diet-disease associations.[54 56 57] However, given the predominately white and highly-educated participants in the UK Biobank, further research in diverse populations is needed to investigate the applicability of these diet quality methodologies for examining CVD risk independent of genetic risk.[7]

Our stronger associations between diet quality and genetic CVD risk in males confirm previous research.[14 58] Although this may be partly explained by the high prevalence of diabetes and unhealthy behaviours in men,[59] it may also be due the lower number of events and deaths in women compared with men in the present study. Nonetheless, it is likely that the biological and behavioural pathways in which risk factors exert their effects on CVD risk are different between men and women.[58]

Strengths and limitations

Our main strength was the large sample size and inclusion of genetic data. This enabled investigation of a genetic risk score created 300 SNPs known to be associated with CVD, more than any previous publications in the UK Biobank.[14 55] While the PRS used was specific to coronary disease, it has been used to identify predispositions to a wide variety of CVD and non-CVDs, as well as premature mortality, given these may develop in parallel with coronary disease for the same genetic origins. The dietary questionnaire has been previously validated and included sufficient detail to allow us to create three contrasting diet quality indices. There are a number of limitations that should be acknowledged. While the dietary assessment method is a short-term measure of intake, our use of up to four instances of dietary assessments provided an estimate of longer-term intake. Although the present analysis is likely to be subject to self-selection bias associated with the number of participants who completed the dietary assessment and the low response rate, associations between demographic and behavioural risk factors and mortality in the UK Biobank have been shown to be comparable to those from national health survey data from England and Scotland.[60] Whilst we adjusted analyses based on a range of confounders identified using a directed acyclic graph, we cannot discount the possibility of residual or unmeasured confounding.

Conclusion

This prospective population-based cohort study provided evidence that higher diet quality (RFS, HDI and MDS) was associated with lower risk of all-cause mortality, regardless of genetic CVD risk. Diet quality, when estimated using the RFS only, was associated with lower risk of CVD-related mortality and MI, independent of genetic CVD risk. The diet quality indices investigated in this study have common food-based scoring components, providing further evidence for the best practice design and implementation of food-based diet quality indices for assessing health outcomes. Further research in diverse populations is needed to investigate the applicability of different diet quality methodologies for examining CVD risk independent of genetic susceptibility.

Authors' contributions: KML, GA, SB, JW, CM and SAM designed the analysis. KML, SB, JW and GA conducted the statistical analysis. KML drafted the manuscript. All authors contributed to the critical review of the manuscript and approved the final version of the manuscript.

Funding: KML is supported by a National Health and Medical Research Council Emerging Leadership Fellowship (APP1173803). JW is funded by the Lister Prize Fellowship (173096). The funding source had no role in the design or conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; KML is a consultant for HeadUp Labs; no other financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: UK Biobank received ethical approval from the research ethics committee (reference 13/NW/0382). All participants provided informed consent to participate. An ethics exemption was granted by Deakin University Human Research Ethics Committee (Reference number 2019_293).

Data availability statement: The genetic and phenotypic UK Biobank data are available on application to the UK Biobank. This research has been conducted using the UK Biobank Resource under Application 34894.

Transparency: The lead author (KML) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted and that discrepancies from the study as planned have been explained

REFERENCES

1. Roth GA, Abate D, Abate KH, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2018;392(10159):1736-88. doi: https://doi.org/10.1016/S0140-6736(18)32203-7
2. Joseph P, Leong D, McKee M, et al. Reducing the Global Burden of Cardiovascular Disease, Part 1: The Epidemiology and Risk Factors. *Circ Res* 2017;121(6):677-94. doi: 10.1161/circresaha.117.308903 [published Online First: 2017/09/02]
3. Afshin A, Sur PJ, Fay KA, et al. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2019;393(10184):1958-72. doi: 10.1016/S0140-6736(19)30041-8
4. Sotos-Prieto M, Bhupathiraju SN, Mattei J, et al. Changes in Diet Quality Scores and Risk of Cardiovascular Disease Among US Men and Women. 2015;132(23):2212-19. doi: doi:10.1161/CIRCULATIONAHA.115.017158
5. Miller V, Webb P, Micha R, et al. Defining diet quality: a synthesis of dietary quality metrics and their validity for the double burden of malnutrition. *The Lancet Planetary Health* 2020;4(8):e352-e70. doi: 10.1016/S2542-5196(20)30162-5
6. Tapsell LC, Neale EP, Satija A, et al. Foods, Nutrients, and Dietary Patterns: Interconnections and Implications for Dietary Guidelines. *Advances in Nutrition: An International Review Journal* 2016;7(3):445-54. doi: 10.3945/an.115.011718
7. Aljuraiban GS, Gibson R, Oude Griep LM, et al. Perspective: The Application of A Priori Diet Quality Scores to Cardiovascular Disease Risk—A Critical Evaluation of Current Scoring Systems. *Adv Nutr* 2019;11(1):10-24. doi: 10.1093/advances/nmz059 %J Advances in Nutrition
8. Aigner A, Becher H, Jacobs S, et al. Low diet quality and the risk of stroke mortality: the multiethnic cohort study. *Eur J Clin Nutr* 2018;72(7):1035-45. doi: 10.1038/s41430-018-0103-4
9. Sotos-Prieto M, Bhupathiraju SN, Mattei J, et al. Association of Changes in Diet Quality with Total and Cause-Specific Mortality. *N Engl J Med* 2017;377(2):143-53. doi: 10.1056/NEJMoa1613502
10. Kant AK, Schatzkin A, Graubard BI, et al. A prospective study of diet quality and mortality in women. *JAMA* 2000;283(16):2109-15. doi: 10.1001/jama.283.16.2109
11. Heianza Y, Zhou T, Sun D, et al. Genetic susceptibility, plant-based dietary patterns, and risk of cardiovascular disease. *The American Journal of Clinical Nutrition* 2020 doi: 10.1093/ajcn/nqaa107
12. Liese AD, Krebs-Smith SM, Subar AF, et al. The Dietary Patterns Methods Project: synthesis of findings across cohorts and relevance to dietary guidance. *The Journal of nutrition* 2015;145(3):393-402. doi: 10.3945/jn.114.205336 [published Online First: 2015/01/21]
13. Voruganti VS. Nutritional Genomics of Cardiovascular Disease. *Curr Genet Med Rep* 2018;6(2):98-106. doi: 10.1007/s40142-018-0143-z [published Online First: 2018/04/30]
14. Rutten-Jacobs LC, Larsson SC, Malik R, et al. Genetic risk, incident stroke, and the benefits of adhering to a healthy lifestyle: cohort study of 306 473 UK Biobank participants. *BMJ : British Medical Journal* 2018;363:k4168. doi: 10.1136/bmj.k4168 %J BMJ
15. Ahmad T, Chasman DI, Mora S, et al. The Fat-Mass and Obesity-Associated (FTO) gene, physical activity, and risk of incident cardiovascular events in white women. *American Heart Journal* 2010;160(6):1163-69. doi: https://doi.org/10.1016/j.ahj.2010.08.002
16. Song Y, Stampfer MJ, Liu S. Meta-Analysis: Apolipoprotein E Genotypes and Risk for Coronary Heart Disease. *Annals of Internal Medicine* 2004;141(2):137-47. doi: 10.7326/0003-4819-141-2-200407200-00013 %J Annals of Internal Medicine
17. Cambien F, Tiret L. Genetics of Cardiovascular Diseases. 2007;116(15):1714-24. doi: doi:10.1161/CIRCULATIONAHA.106.661751

18. Sugrue LP, Desikan RS. What Are Polygenic Scores and Why Are They Important? *JAMA* 2019;321(18):1820-21. doi: 10.1001/jama.2019.3893 %J JAMA
19. Sudlow C, Gallacher J, Allen N, 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. doi: 10.1371/journal.pmed.1001779
20. Bradbury KE, Young HJ, Guo W, et al. Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire. *J Nutr Sci* 2018;7:e6-e6. doi: 10.1017/jns.2017.66
21. Lachat C, Hawwash D, Ocké MC, et al. Strengthening the Reporting of Observational Studies in Epidemiology—Nutritional Epidemiology (STROBE-nut): An Extension of the STROBE Statement. *PLoS Med* 2016;13(6):e1002036. doi: 10.1371/journal.pmed.1002036
22. Liu B, Young H, Crowe FL, 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. doi: 10.1017/S1368980011000942 [published Online First: 2011/06/01]
23. Greenwood DC, Hardie LJ, Frost GS, et al. Validation of the Oxford WebQ Online 24-Hour Dietary Questionnaire Using Biomarkers. *Am J Epidemiol* 2019;188(10):1858-67. doi: 10.1093/aje/kwz165 %J American Journal of Epidemiology
24. Food Standards Agency. Food Portion Sizes. London: The Stationary Office2005.
25. Royal Society of Chemistry; Ministry of Agriculture FaF. McCance and Widdowson's The Composition of Foods. 6th ed. Cambridge, United Kingdom: Royal Society of Chemistry; 2002.
26. Kant AK, Graubard BI. A Comparison of Three Dietary Pattern Indexes for Predicting Biomarkers of Diet and Disease. *J Am Coll Nutr* 2005;24(4):294-303. doi: 10.1080/07315724.2005.10719477
27. Maynard M, Gunnell D, Ness AR, et al. What influences diet in early old age? Prospective and cross-sectional analyses of the Boyd Orr cohort. *Eur J Public Health* 2005;16(3):315-23. doi: 10.1093/eurpub/cki167
28. Panagiotakos DB, Pitsavos C, Arvaniti F, et al. Adherence to the Mediterranean food pattern predicts the prevalence of hypertension, hypercholesterolemia, diabetes and obesity, among healthy adults; the accuracy of the MedDietScore. *Preventive Medicine* 2007;44(4):335-40. doi: https://doi.org/10.1016/j.ypmed.2006.12.009
29. Waijers PMCM, Feskens EJM, Ocké MC. A critical review of predefined diet quality scores. *British Journal of Nutrition* 2007;97(02):219-31. doi: doi:10.1017/S0007114507250421
30. Kanauchi M, Kanauchi K. The World Health Organization's Healthy Diet Indicator and its associated factors: A cross-sectional study in central Kinki, Japan. *Prev Med Rep* 2018;12:198-202. doi: https://doi.org/10.1016/j.pmedr.2018.09.011
31. McNaughton SA, Bates CJ, Mishra GD. Diet quality is associated with all-cause mortality in adults aged 65 years and older. *J Nutr* 2012;142(2):320-5. doi: 10.3945/jn.111.148692 [published Online First: 2011/12/23]
32. Livingstone KM, McNaughton SA. Diet quality is associated with obesity and hypertension in Australian adults: a cross sectional study. *BMC Public Health* 2016;16(1):1037. doi: 10.1186/s12889-016-3714-5
33. McCullough ML, Feskanich D, Stampfer MJ, et al. Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. *Am J Clin Nutr* 2002;76(6):1261-71. doi: 10.1093/ajcn/76.6.1261 [published Online First: 2002/11/27]
34. World Health Organization. Healthy diet WHO Fact Sheet, No 394 (2015) 2020 [Available from: http://www.who.int/mediacentre/factsheets/fs394/en/ accessed 9 March 2020.
35. Huijbregts P, Feskens E, Räsänen L, et al. Dietary pattern and 20 year mortality in elderly men in Finland, Italy, and the Netherlands: longitudinal cohort study. 1997;315(7099):13-17. doi: 10.1136/bmj.315.7099.13 %J BMJ

36. Trichopoulou A, Orfanos P, Norat T, et al. Modified Mediterranean diet and survival: EPIC-elderly prospective cohort study. 2005;330(7498):991. doi: 10.1136/bmj.38415.644155.8F %J BMJ
37. Trichopoulou A, Costacou T, Bamia C, et al. Adherence to a Mediterranean Diet and Survival in a Greek Population. 2003;348(26):2599-608. doi: 10.1056/NEJMoa025039
38. Bach-Faig A, Berry EM, Lairon D, et al. Mediterranean diet pyramid today. Science and cultural updates. *Public Health Nutr* 2011;14(12A):2274-84. doi: 10.1017/S1368980011002515 [published Online First: 2011/12/13]
39. Magliano D, Liew D, Pater H, et al. Accuracy of the Australian National Death Index: comparison with adjudicated fatal outcomes among Australian participants in the Long-term Intervention with Pravastatin in Ischaemic Disease (LIPID) study. 2003;27(6):649-53. doi: 10.1111/j.1467-842X.2003.tb00615.x
40. Schnier C, Bush K, Nolan J, et al. Definitions of Stroke for UK Biobank Phase 1 Outcomes Adjudication Documentation prepared by: On behalf of UK Biobank Outcome Adjudication Group. 2017. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_mi.pdf (accessed 16 September 2020).
41. Schnier C, Bush K, Nolan J, et al. Definitions of Acute Myocardial Infarction and Main Myocardial Infarction Pathological Types UK Biobank Phase 1 Outcomes Adjudication Documentation prepared by: On behalf of UK Biobank Outcome Adjudication Group. 2017. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_mi.pdf (accessed 16 September 2020).
42. National Health Service. COVID-19 Daily Deaths. 2020. https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-daily-deaths/ (accessed 16 September 2020).
43. Ntalla I, Kanoni S, Zeng L, et al. Genetic Risk Score for Coronary Disease Identifies Predispositions to Cardiovascular and Noncardiovascular Diseases. *J Am Coll Cardiol* 2019;73(23):2932-42. doi: https://doi.org/10.1016/j.jacc.2019.03.512
44. Dudbridge F. Polygenic Epidemiology. 2016;40(4):268-72. doi: 10.1002/gepi.21966
45. Towsend P, Phillimore P, Beattie A. Health and Deprivation. *Nurs Stand* 1988;2(17):34. doi: 10.7748/ns.2.17.34.s66 [published Online First: 1988/01/30]
46. World Health Organization. BMI classification 2015 [updated 21 July 2015. Available from: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html accessed 3rd December 2015.
47. The IPAQ Group. IPAQ scoring protocol - International Physical Activity Questionnaire. Available at: https://sites.google.com/site/theipaq/scoring-protocol. Accessed 22 May 2020. [
48. Neelakantan N, Koh W-P, Yuan J-M, et al. Diet-Quality Indexes Are Associated with a Lower Risk of Cardiovascular, Respiratory, and All-Cause Mortality among Chinese Adults. *The Journal of Nutrition* 2018;148(8):1323-32. doi: 10.1093/jn/nxy094 %J The Journal of Nutrition
49. Stefler D, Malyutina S, Kubinova R, et al. Mediterranean diet score and total and cardiovascular mortality in Eastern Europe: the HAPIEE study. *Eur J Nutr* 2017;56(1):421-29. doi: 10.1007/s00394-015-1092-x
50. Lassale C, Gunter MJ, Romaguera D, et al. Diet Quality Scores and Prediction of All-Cause, Cardiovascular and Cancer Mortality in a Pan-European Cohort Study. *PLoS One* 2016;11(7):e0159025. doi: 10.1371/journal.pone.0159025
51. Kaluza J, Håkansson N, Brzozowska A, et al. Diet quality and mortality: a population-based prospective study of men. *Eur J Clin Nutr* 2009;63(4):451-57. doi: 10.1038/sj.ejcn.1602968
52. Schulze MB, Martínez-González MA, Fung TT, et al. Food based dietary patterns and chronic disease prevention. *BMJ* 2018;361 doi: 10.1136/bmj.k2396
53. Lakkur S, Judd SE. Diet and Stroke. 2015;46(7):2007-11. doi: doi:10.1161/STROKEAHA.114.006306
54. Chiavaroli L, Viguiliouk E, Nishi SK, et al. DASH Dietary Pattern and Cardiometabolic Outcomes: An Umbrella Review of Systematic Reviews and Meta-Analyses. *Nutrients* 2019;11(2) doi: 10.3390/nu11020338 [published Online First: 2019/02/16]

55. Said MA, Verweij N, van der Harst P. Associations of Combined Genetic and Lifestyle Risks With Incident Cardiovascular Disease and Diabetes in the UK Biobank Study. *JAMA Cardiol* 2018;3(8):693-702. doi: 10.1001/jamacardio.2018.1717 [published Online First: 2018/06/30]
56. Baden MY, Liu G, Satija A, et al. Changes in Plant-Based Diet Quality and Total and Cause-Specific Mortality. *Circulation* 2019;140(12):979-91. doi: doi:10.1161/CIRCULATIONAHA.119.041014
57. Huang T, Yang B, Zheng J, et al. Cardiovascular Disease Mortality and Cancer Incidence in Vegetarians: A Meta-Analysis and Systematic Review. *Ann Nutr Metab* 2012;60(4):233-40. doi: 10.1159/000337301
58. Den Ruijter HM, Haitjema S, Asselbergs FW, et al. Sex matters to the heart: A special issue dedicated to the impact of sex related differences of cardiovascular diseases. *Atherosclerosis* 2015;241(1):205-07. doi: 10.1016/j.atherosclerosis.2015.05.003
59. Andersen KK, Andersen ZJ, Olsen TS. Age- and Gender-Specific Prevalence of Cardiovascular Risk Factors in 40 102 Patients With First-Ever Ischemic Stroke. *Stroke* 2010;41(12):2768-74. doi: doi:10.1161/STROKEAHA.110.595785
60. Batty GD, Gale CR, Kivimäki M, et al. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *BMJ* 2020;368:m131. doi: 10.1136/bmj.m131

Supplemental Table 1. Components and scoring methods of the Recommended Food Score (RFS)

Dietary Indictor	Indicator food groups[1]	Criteria for scoring
1. Fruits	1. Pome fruit (apples, pears) 2. Berry fruit (berry) 3. Citrus fruit (orange, satsuma, grapefruit) 4. Stone fruit (nectarine, peach, plum, cherry, prune) 5. Tropical and subtropical fruit (banana, pineapple, mango) 6. Other fruit (other fruit, grape, melon, dried fruit, stewed fruit) 7. Fruit juice (orange juice, grapefruit juice, pure fruit/vegetable juice)	Indicator food groups were assigned a score of 1 if they were consumed above the minimum threshold of 15 g/d. Fruit juice was assigned a score of 1 if it was consumed above the minimum threshold of 30 g/d.
2. Vegetables	1. Green (lettuce, spinach, sprouts, watercress, cucumber, celery, courgette) and brassica vegetables (cabbage, cauliflower, broccoli) 2. Legumes (pulses, broad bean) 3. Carrot and root vegetables (carrot, turnip/swede, beetroot parsnip, onion, garlic, leek) 4. Starchy vegetables (boiled/baked potatoes (*butter/margarine added to potatoes, butternut squash), mashed potato, sweet potato, sweetcorn) 5. Tomato and tomato products (fresh tomato, tinned tomato) 6. Peas and beans (green bean, pea) 7. Other vegetables (other vegetables, mushroom, sweet pepper, side salad, olives)	Indicator food groups were assigned a score of 1 if they were consumed above the minimum threshold of 15 g/d.
3. Whole grains	1. Wholegrain (whole-wheat cereal, sliced bread (wholemeal), baguette (wholemeal), bap (wholemeal), bread roll (wholemeal)) 2. High fibre cereals (porridge, muesli, oat crunch, bran cereal) and wholegrain pasta and brown rice	Indicator food groups were assigned a score of 1 if they were consumed above the minimum threshold of 15 g/d.
4. Lean meats and alternatives	1. Poultry 2. Fish (tinned tuna, oily fish, white fish, prawns, lobster/crab, shellfish) 3. Alternatives (whole egg, omelette, egg in sandwich, other egg, seed (e.g. unsalted peanuts, unsalted nuts, types of spreads/sauces consumed (Peanut butter) seeds), tofu, quorn)	Indicator food groups were assigned a score of 1 if they were consumed above the minimum threshold of 15 g/d.
5. Low-fat dairy	1. 2%, 1% or skim milk (type of milk consumed (semi skimmed, skimmed, goat/sheep milk, powdered milk, cholesterol lowering)) 2. Low fat cheese and yogurt (Low fat hard cheese, low fat cheese spread, cottage cheese, yogurt (low fat yogurt consumer), goat's cheese)	Milk was assigned a score of 1 if it was consumed above the minimum threshold of 30 g/d. Indicator food groups were assigned a score of 1 if they were consumed above the minimum threshold of 15 g/d.

1. As available in the UK Biobank

Supplemental Table 2. Components and scoring methods of the Healthy Diet Indicator (HDI)

Dietary Indictor	Indicator foods[1]	Criteria for scoring
1. Saturated fatty acids	Saturated fat	>10% energy intake=0 0-10% energy intake=1
2. Polyunsaturated fatty acids	Polyunsaturated fat	<6 or >10% energy intake=0 6-10% energy intake=1
3. Protein	Protein	<10 or >15% energy intake=0 10-15% energy intake=1
4. Total carbohydrates	Carbohydrates	<50% or >70% energy intake=0 50-70% energy intake=1
5. Dietary fibre	Englyst dietary fibre	<18 or >32 g/day=0 18-32 g/day =1
6. Fruits and vegetables	Mixed vegetable, vegetable pieces, avocado, beetroot, broccoli, butternut squash, cabbage/kale, carrot, cauliflower, celery, courgette, cucumber, garlic, leek, lettuce, mushroom, onion, olives, parsnip, pea, side salad, sweet pepper, spinach, sprouts, sweetcorn, fresh tomato, tinned tomato, green bean, turnip/swede, watercress, other vegetables, homemade soup (vegetables) Stewed fruit, prune, dried fruit, mixed fruit, apple, banana, berry, cherry, grapefruit, grape, mango, melon, orange, satsuma, peach/nectarine intake, pear, pineapple, plum, other fruit	<400 g/day=0 ≥400 g/day=1
7. Pulses and nuts	Baked bean, pulses, broad bean Salted peanuts, unsalted peanuts, salted nuts, unsalted nuts, seeds, types of spreads/sauces consumed (Peanut butter)	<30 g/day=0 ≥30 g/day=1
8. Total non-milk extrinsic sugars	Total sugars	>10 % energy intake=0 0-10 % energy intake=1
9. Fish	Tinned tuna, oily fish, white fish, prawns, lobster/crab, shellfish, other fish Homemade soup, ingredients in homemade soup (fish)	<32 g/day=0 ≥32 g/day=1
10. Red meat and meat products	Beef, pork, lamb, other meat Poultry intake (skin removed from poultry (no); fat removed from poultry(no)) Homemade soup, ingredients in homemade soup (meat) Sausage, bacon, ham, liver	>90 g/day=0 ≤90 g/day=1
11. Calcium	Calcium	<700 mg/day=0 ≥700 mg/day=1

1. As available in the UK Biobank

Supplemental Table 3. Components and scoring methods of the Mediterranean Diet Score (MDS)

Dietary Indictor	Indicator foods[1]	Criteria for scoring
1. Vegetables (excluding potatoes, legumes or fruit juice)	Mixed vegetable, vegetable pieces, avocado, beetroot, broccoli, butternut squash, cabbage/kale, carrot, cauliflower, celery, courgette, cucumber, garlic, leek, lettuce, mushroom, onion, olives, parsnip, pea, side salad, sweet pepper, spinach Sprouts, sweetcorn, fresh tomato, tinned tomato, green bean, turnip/swede, watercress, other vegetables, homemade soup (vegetables)	Sex-specific median intakes used as cut points. Intakes (for indictors 1-6) above median score 1 and intakes below the median score 0.
2. Legumes	Baked bean, pulses, broad bean, homemade soup (pulses)	
3. Fruit and nuts	Stewed fruit, prune, dried fruit, mixed fruit, apple, banana, berry, cherry, grapefruit, grape, mango, melon, orange, satsuma, peach/nectarine, pear intake, pineapple, plum, other fruit Orange juice, grapefruit juice, pure fruit/vegetable juice Unsalted peanuts, unsalted nuts, types of spreads/sauces consumed (Peanut butter), seeds	
4. Cereals	Porridge, muesli, oat crunch, plain cereal, bran cereal, whole-wheat cereal, other cereal Bread consumed, sliced bread (mixed; wholemeal; seeded; other), baguette (mixed; wholemeal; seeded; other), bap (mixed; wholemeal; seeded; other), bread roll (mixed; wholemeal; seeded; other), other bread White pasta, wholemeal pasta, white rice, brown rice,, couscous, other grain Homemade soup, ingredients in homemade soup (pasta)	
5. Fish and seafood	Tinned tuna, oily fish, white fish, prawns, lobster/crab, shellfish, other fish Homemade soup, ingredients in homemade soup (fish)	
6. Monounsaturated/ saturated fats ratio	Monounsaturated fats, saturated fats	
7. Dairy products	Milk, milk added to cereal Low fat hard cheese, low fat cheese spread, cottage cheese Yogurt (low fat yogurt consumer; full fat yogurt consumer) Goat's cheese, hard cheese, soft cheese, blue cheese, cheese spread, feta, mozzarella, other cheese Dairy smoothie, latte, added milk to instant coffee, added milk to filtered coffee, added milk to espresso, added milk to other coffee type, added milk to standard tea, added milk to rooibos tea, cappuccino	Sex-specific median intakes used as cut points. Intakes (for indictors 7-8) below median score 1 and intakes below the median score 0.
8. Meat and meat products	Beef, pork, lamb, other meat Whole egg, omelette, eggs in sandwiches, scotch egg, other egg Homemade soup, ingredients in homemade soup (meat), sausage, bacon, ham	

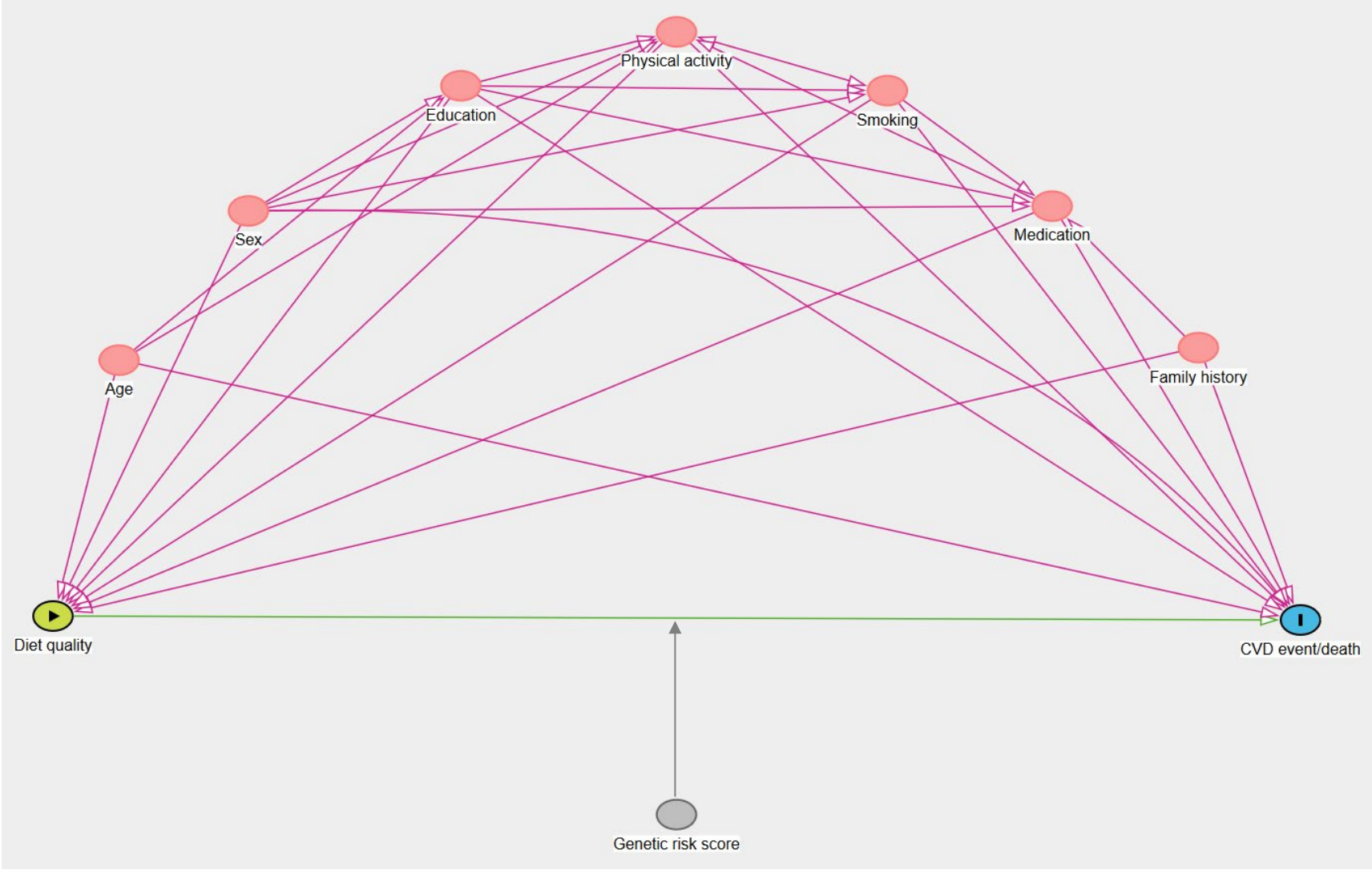
9. Alcohol	Red wine, rose wine, white wine Beer/cider Fortified wine, spirits intake, other alcohol	No more than 2 drinks/day = 1; Never drink or over 2 drinks/day = 0.

1. As available in the UK Biobank

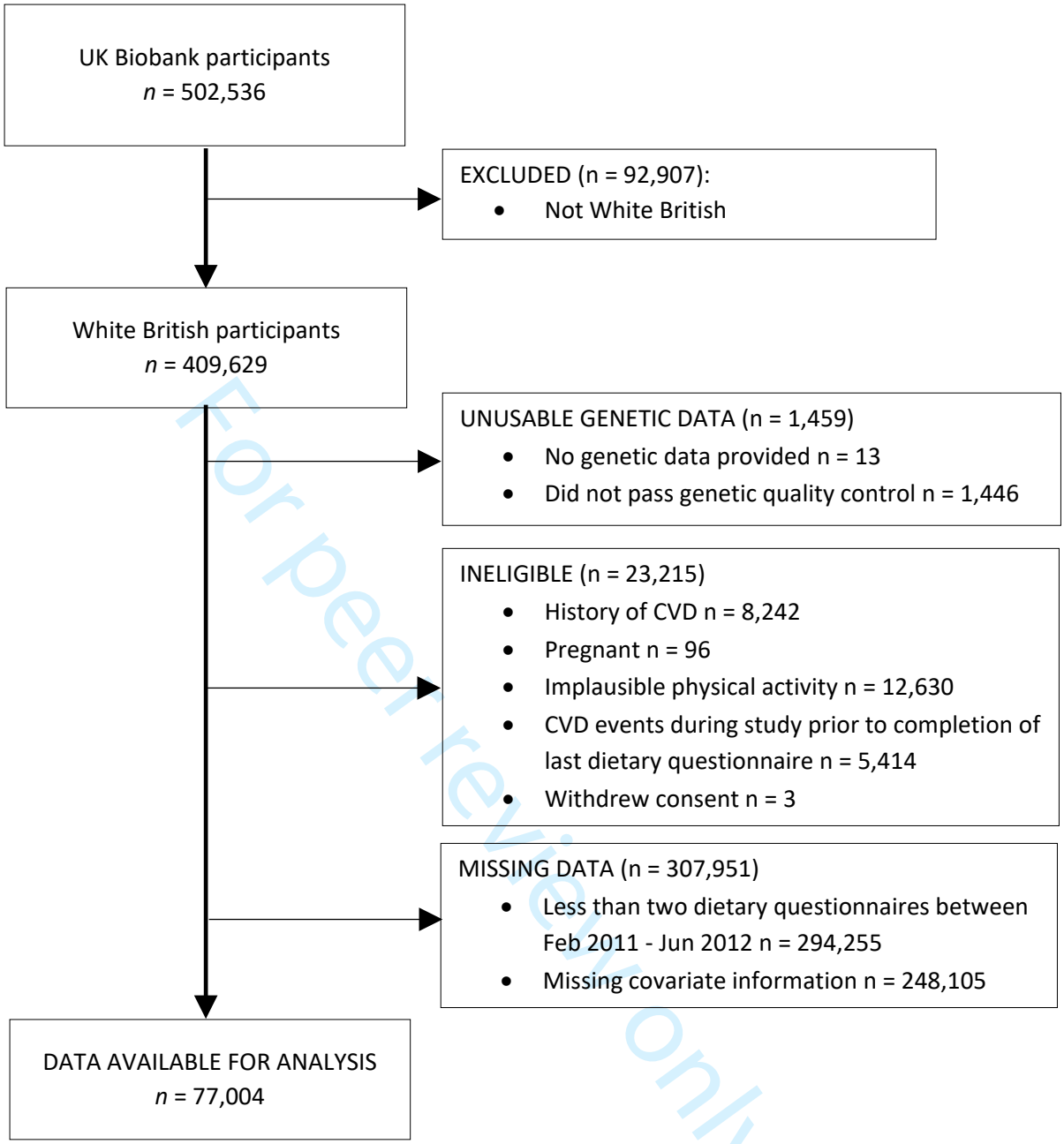
Supplemental Table 4. Comparison of participant characteristics between the excluded and analytic sample

Characteristic	Excluded N (%)	Analytic N (%)
N[1]	425,532	77,004
Sex female	231,382 (54.4)	42,020 (54.6)
Age at recruitment (years), Mean ± SD	56.6 ± 8.2	56.2 ± 7.8
Townsend Deprivation Index		
Least deprived	82,535 (19.4)	18,129 (23.5)
2nd least deprived	82,878 (19.5)	17,227 (22.4)
Medium deprivation	84,323 (19.8)	16,067 (20.9)
2nd most deprived	85,475 (20.1)	14,900 (19.4)
Most deprived	89,698 (21.1)	10,681 (13.9)
Smoking		
Never smoked	228,689 (54.1)	44,856 (58.3)
Ex-smoker	145,891 (34.5)	27,184 (35.3)
Current smoker	48,016 (11.4)	4,964 (6.4)
Body Mass Index (kg/m^2), Mean ± SD	27.6 ± 4.9	26.5 ± 4.4

Townsend Deprivation Index is a composite measure of deprivation based on unemployment, non-car ownership, non-home ownership, and household overcrowding.
1, In the excluded sample, data on Townsend Deprivation Index and smoking were available in n=424,909 and n= 422,596, respectively. Data on Body Mass Index were available in n= 422,530 and n=76,901 in the excluded and analytic sample, respectively.



Supplemental Figure 1. Directed acyclic graph showing relationship between the exposure (diet quality) and outcome (CVD events/death). Confounders are represented by red dots. The moderator (polygenic risk score) is represent by a grey dot.



Supplemental Figure 2. Flow diagram of participants in the UK Biobank

Reviewer only Supplemental Table. STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	5
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-9
Bias	9	Describe any efforts to address potential sources of bias	10
Study size	10	Explain how the study size was arrived at	10
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6-10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	10-11
		(*b*) Describe any methods used to examine subgroups and interactions	10-11
		(*c*) Explain how missing data were addressed	10-11
		(*d*) If applicable, explain how loss to follow-up was addressed	10-11
		(*e*) Describe any sensitivity analyses	N/A
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	11
		(b) Give reasons for non-participation at each stage	11
		(c) Consider use of a flow diagram	11
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	11
		(b) Indicate number of participants with missing data for each variable of interest	11
		(c) Summarise follow-up time (eg, average and total amount)	11
Outcome data	15*	Report numbers of outcome events or summary measures over time	11
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	11-18
		(*b*) Report category boundaries when continuous variables were categorized	11-18
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	11-18
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11-18
Discussion			
Key results	18	Summarise key results with reference to study objectives	19
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	20
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	19-21
Generalisability	21	Discuss the generalisability (external validity) of the study results	19-21
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	21-22

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in

conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

Diet quality indices, genetic risk and risk of cardiovascular disease and mortality: a longitudinal analysis of 77,004 UK Biobank participants

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045362.R1
Article Type:	Original research
Date Submitted by the Author:	17-Dec-2020
Complete List of Authors:	Livingstone, Katherine; Deakin University, Institute for Physical Activity and Nutrition Abbott, Gavin; Deakin University, Institute for Physical Activity and Nutrition Bowe, Steven; Deakin University - Geelong Waterfront Campus, Deakin Biostatistics Unit Ward, Joey; University of Glasgow, Institute of Health and Wellbeing Milte, CM ; Deakin University, Institute for Physical Activity and Nutrition McNaughton, Sarah; Deakin University - Geelong Waterfront Campus, Institute for Physical Activity and Nutrition
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Nutrition and metabolism, Public health
Keywords:	NUTRITION & DIETETICS, EPIDEMIOLOGY, PUBLIC HEALTH, Coronary heart disease < CARDIOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

TITLE

Diet quality indices, genetic risk and risk of cardiovascular disease and mortality: a longitudinal analysis of 77,004 UK Biobank participants

AUTHORS

Katherine M. Livingstone[1]*, Gavin Abbott[1], Steven J. Bowe[2], Joey Ward[3], Catherine Milte[1], Sarah A. McNaughton[1]

AFFILIATIONS

1, Deakin University, Institute for Physical Activity and Nutrition Research, School of Exercise and Nutrition Sciences, Geelong, VIC 3220, Australia, k.livingstone@deakin.edu.au, gavin.abbott@deakin.edu.au, catherine.milte@deakin.edu.au sarah.mcnaughton@deakin.edu.au

2, Deakin Biostatistics Unit, Deakin University, Geelong, VIC 3220, Australia, s.bowe@deakin.edu.au

3, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK, Joey.Ward@glasgow.ac.uk

CONTACT INFO:

Dr Katherine Mary Livingstone,

Deakin University, Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Melbourne, Melbourne Burwood Campus, 221 Burwood Highway, Victoria 3125, Australia

Tel: +61 3 92445416 Email: k.livingstone@deakin.edu.au

WORD COUNT: 4925

ABBREVIATIONS: Body Mass Index, BMI; Cardiovascular Disease, CVD; Healthy Diet Indicator, HDI; Myocardial Infraction, MI; Mediterranean Diet Score, MDS; Polygenic Risk Score, PRS; Recommended Food Score, RFS; Single Nucleotide Polymorphism, SNP

ABSTRACT

Objectives: To examine associations of three diet quality indices and a polygenic risk score with incidence of all-cause mortality, cardiovascular disease (CVD) mortality, myocardial infarction (MI) and stroke.

Design: Prospective cohort study.

Setting: UK Biobank, UK.

Participants: 77,004 men and women (40-70 years) recruited between 2006 and 2010.

Main outcome measures: A polygenic risk score was created from 300 single nucleotide polymorphisms (SNP) associated with CVD. Cox proportional hazard ratios (HR) were used to estimate independent effects of diet quality and genetic risk on all-cause mortality, CVD mortality, MI and stroke risk. Dietary intake (Oxford WebQ) was used to calculate Recommended Food Score (RFS), Healthy Diet Indicator (HDI) and Mediterranean Diet Score (MDS).

Results: New all-cause (n=2,409) and CVD (n=364) deaths, and MI (n=1,141) and stroke (n=748) events were identified during mean follow-ups of 7.9 and 7.8 years, respectively. The adjusted HR associated with one-point higher RFS for all-cause mortality was 0.96 (95% CI: 0.94 to 0.98), CVD mortality was 0.94 (95% CI: 0.90 to 0.98), MI was 0.97 (95% CI: 0.95 to 1.00) and stroke was 0.94 (95% CI: 0.91 to 0.98). The adjusted HR for all-cause mortality associated with one-point higher HDI and MDS was 0.97 (95% CI: 0.93 to 0.99) and 0.95 (95% CI: 0.91 to 0.98), respectively. The adjusted HR associated with one-point higher MDS for stroke was 0.93 (95% CI: 0.87 to 1.00). There was little evidence of associations between HDI and risk of CVD mortality, MI or stroke. There was evidence of an interaction between diet quality and genetic risk score for MI.

Conclusion: Higher diet quality predicted lower risk of all-cause mortality, independent of genetic risk. Higher RFS was also associated with lower risk of CVD mortality and MI. These findings demonstrate the benefit of following a healthy diet, regardless of genetic risk.

Strengths and limitations of this study

- This large prospective population-based cohort included repeat dietary assessments, using a validated questionnaire, and hospital register data on CVD incidence and mortality.
- The creation of three contrasting diet quality indices informs the best practice design and implementation of food-based diet quality indices for assessing diet-disease relationships.
- A polygenic genetic risk score was created for each participant using 300 SNPs known to be associated with CVD and all-cause mortality.
- The present analysis is likely to be subject to self-selection bias associated with the number of participants who completed the dietary assessment and the low response rate.
- Findings are not generalisable to non-Caucasian populations, thus future research in diverse populations is needed to investigate the applicability of different diet quality methodologies for examining CVD risk independent of genetic susceptibility.

INTRODUCTION

Cardiovascular disease (CVD) is one of the leading causes of morbidity and mortality worldwide.[1] As a multifactorial condition, CVD risk is attributable to a combination of genetic and behavioural influences.[2] With poor diet now a leading risk factor for non-communicable diseases,[3] further understanding of the role of diet on CVD risk is warranted.

The overall quality of diets is an emerging predictor of CVD events and mortality.[4 5] Diet quality indices, that score dietary intakes according to *a priori* knowledge,[6] have been used to investigate association between diet and CVD incidence and mortality.[4 5 7-11] These indices can capture different aspects of diet quality, for example being based on intakes for encouraged foods only (e.g. Recommended Food Score, RFS), a combination of foods and nutrients from dietary guidelines (e.g. Healthy Diet Indicator, HDI) or a dietary pattern identified as healthful (e.g. Mediterranean Diet Score, MDS). However, our understanding is limited by the use of contrasting diet quality methodologies and a paucity of studies comparing different indices in large prospective population-based cohorts. Comparison of contrasting diet quality indices will identify whether differences in these methodologies are important for understanding diet-disease associations and will inform the international standardisation of diet quality methodologies for assessing health outcomes.[5 7 12]

The role of diet and genetics on risk of CVD is an emerging area of research.[13 14] Prior to the accessibility to whole genome sequencing, most research focused on links between single nucleotide polymorphisms (SNPs) and CVD.[15-17] Recent research has shown that polygenic risk scores (PRS), that incorporate multiple SNPs, are a good indicator of risk for complex conditions, such as CVD,[14 18] although the extent to which they influence the association between diet quality and CVD risk is unclear. Further research is also needed to elucidate whether diet quality is a risk factor for CVD independent of genetic risk. Moreover, the longitudinal association between contrasting diet quality indices, genetic risk and different CVD subtypes is unknown. Thus, the aim of this study was to examine the prospective role of three diet quality indices (HDI, RFS and MDS) and a PRS on risk of stroke, myocardial infarction, CVD mortality and all-cause mortality. Findings will advance understanding of the applicability of diet quality indices for assessing CVD risk.

METHODS

Study design and participants

The UK Biobank is a population cohort of half a million individuals living in the United Kingdom that aimed to examine determinants of disease in middle-aged adults.[19] Persons aged 40 to 69 years were identified from National Health Service patient registers and invited to participate. Individuals were invited to one of 22 assessment centres across England, Scotland and Wales between 2006 and 2011. At each centre, participants completed a touchscreen questionnaire to collect information on demographic characteristics, lifestyle behaviours and general health. The Oxford WebQ, a web-based 24-h dietary assessment tool, was introduced in 2009 to collect information on dietary intake.[20] Physical measurements (e.g., height and weight) were taken and participants provided blood and urine samples. Participants were followed up via linkage to health records and death registries. The UK Biobank received ethical approval from the Research Ethics Committee (Reference 11/NW/0382). Electronic signed consent was obtained from all participants. Participants were excluded from the present analysis if they i) did not identify as White British, ii) were ineligible based on previous history of CVD before entering the study, pregnancy, implausible physical activity data and CVD events during the study prior to completion of last dietary questionnaire, iii) had missing data for outcomes, exposures and covariates/moderators and v) had less than two timepoints of dietary data between February 2011 - June 2012. Results are reported according to the STROBE-NUT checklist for cohort studies.[21]

Study measures

Dietary intake

The Oxford WebQ was used to collect information on the frequency of consumption of 206 foods and 32 beverages during the previous 24 hours.[20 22 23] The Oxford WebQ is a 24-hour dietary questionnaire that has been validated against a traditional interviewer-administered multiple-pass 24-hour dietary recall and biomarkers for protein, potassium, and total sugar intake and total energy expenditure estimated by accelerometery.[23] Energy and nutrient intakes were calculated by multiplying the frequency of consumption of each food or drink by the standard portion size and energy and nutrient composition of each item.[24 25] Participants recruited between April 2009 and September 2010 completed the Oxford WedQ using the touchscreen at the assessment centre. Repeat Oxford WebQs were

collected via four online cycles between February 2011 to June 2012: February 2011 to April 2011 (online cycle 1); June 2011 to September 2011 (online cycle 2); October 2011 to December 2011 (online cycle 3); April 2012 to June 2012 (online cycle 4). The total period of available dietary data from the Oxford WebQ was 38 months (Apr 2009 - Jun 2012). Email invitations were sent on different days of the week to capture variation in dietary intakes and participants were given 3 days to complete the questionnaire for cycles 1 and 2 and 14 days for cycles 3 and 4.

To establish a baseline dietary intake in the present analysis, we calculated a mean dietary intake based on the four online Oxford WebQ cycles only. This was because the time between the 1st and 4th online cycle measurements was 16 months (Feb 2011 - Jun 2012) and was considered a more credible timeframe for an average baseline than the 38 months available from all five Oxford WebQ measurements. This resulted in a minimal sample loss (<10%) while providing a shorter dietary exposure period and a more consistent approach to the use of the dietary data by using only the online cycles of the OxfordWebQ. To better capture usual intake, we calculated average nutrient intakes, food group intakes and diet quality scores for participants who had two or more valid measurements for the four online cycles of the Oxford WebQ.

Diet quality

Information on food and beverage intakes from the Oxford WebQ were used to calculate three diet quality indices: the Recommended Food Score (RFS), which is based on intakes of encouraged foods only,[26] and the Healthy Diet Indicator (HDI), which scores intakes of a combination of foods and nutrients from dietary guidelines,[27] and the Mediterranean Diet Score (MDS), representing dietary patterns identified as healthful.[28] These indices were selected as they represent three contrasting diet quality methodologies that have been applied internationally to assess diet-disease associations.[9 10 26 27 29 30]

The RFS is a food-based variety index designed to assess consumption of food groups encouraged in the dietary guidelines.[10] As detailed in Supplemental Table 1, food intakes were scored according to five food groups: fruits (7 items), vegetables (7 items), wholegrains (2 items), lean meat and alternatives (3 items) and reduced fat dairy products (2 items). Scoring was based on the RFS designed by Kant and Graubaud,[26] and has been

used elsewhere.[31 32] We summed intakes of food items within each group to create a total intake for each food group. Food groups were then assigned a score of 1 if they were consumed above the minimum amount threshold: 15 g/d for non-beverages and 30 g/d for beverages. Intakes below these thresholds were scored 0. Scores ranged between 0 and 21, with higher scores indicating a higher quality diet and a wider consumption of recommended foods.[33]

The HDI is a food- and nutrient-based index designed to reflect consumption of foods recommended for a healthy diet by the World Health Organisation.[34] The original HDI was developed and validated in 1997 based on the 1990 World Health Organisation's dietary recommendations for the prevention of chronic disease.[35] We adapted a 12-point Healthy Diet Score designed by Maynard et al.[27] to reflect adherence to the 2020 World Health Organisation healthy diet fact sheet.[34] As cholesterol intake is not part of the 2020 recommendations and information on its intake was not available in the UK Biobank, we used an 11-item score that included the following groups: saturated fat; poly-unsaturated fat; protein; total carbohydrates; dietary fibre; fruits and vegetables; pulses and nuts; total non-milk extrinsic sugars; fish; red meat and meat products; and calcium. Data on intake of non-milk extrinsic sugars was not available in the UK Biobank and so we adapted the HDI to score intakes of total sugars instead. Criteria for scoring was based on cut points detailed in Supplemental Table 2. We assigned intakes within the cut offs a score of 1 and those outside of the cut offs were assigned a score of 0. The total score ranged from 0 to 11, with a higher score reflecting a higher diet quality (Supplemental Table 2).

The MDS is a food- and nutrient-based score designed to reflect adherence to a Mediterranean style diet. The present study used the 9-item index developed and validated by Trichopoulou et al. as it is the first and most widely used version of the MDS.[36 37] Food and nutrient intakes were scored according to nine components: vegetables, legumes, fruits and nuts, cereals, fish and seafood, monounsaturated fats to saturated fats ratio, dairy products, meat and meat products and alcohol (Supplemental Table 3). As used by Trichopoulou et al.,[36] we used sex-specific median intakes as cut off points for intakes of each component. A score of 1 was assigned to participants whose intake of vegetables, legumes, fruits and nuts, cereals, fish and seafood and monounsaturated: saturated fats was above the median. A score of 1 was assigned to intake of dairy products, meat and meat

products below the median. For alcohol, a score of 1 was assigned for low to moderate intake (intake of no more than 2 times/day). A score of 0 was assigned for no alcohol intake or intake greater than 2 times per day.[38] Total MDS score ranged from 0 to 9, with higher scores reflecting better alignment to the Mediterranean diet.

Cardiovascular events and mortality

Mortality status and causes of death were determined by linkage of data with the UK National Death Index (NDI) using the most recent available data from the UK Biobank (September 2020). The accuracy of the NDI for identifying CVD deaths has been established previously in Australia.[39] CVD mortality was estimated from 2006 International Classification of Diseases 10th revision (ICD-10) codes in death certificates. CVD mortality was identified using ICD codes I05-I89. CVD events were recorded between enrolment (1999–2000) and the most recent inpatient hospital data available from the UK Biobank (September 2020). Incident MI (ST-Elevation Myocardial Infarction and Non-ST-Elevation Myocardial Infarction) and stroke (ischaemic, intracerebral haemorrhage, and subarachnoid haemorrhage) were available from algorithms provided by the UK Biobank.[40 41] Algorithms were produced to reliably identify incidence of selected illnesses through consideration of hospital and death register data. The adjudication of "algorithmically defined" outcomes for MI and stroke are detailed elsewhere.[40 41] A censoring data of 4 March 2020 was used for all outcomes. This date was chosen due to a spike in deaths from 5 March onwards, which is likely to correspond to increasing deaths due to COVID-19 recorded in the UK.[42]

Polygenic risk score

We used the March 2018 release of the imputed genetic data from UK Biobank (downloaded 11 November 2019). From the resulting dataset, we excluded those who self-reported ancestry other than white British, those who were missing more than 10% of the genetic data and those who were defined by UK Biobank as being heterozygosity outliers. Additionally, for every pair of who were individuals who were second cousins or closer (i.e. those with a kinship coefficient > 0.042) one was excluded at random. We used information on 300 single nucleotide polymorphisms (SNPs) known to be associated with coronary artery disease[43] to create a PRS for CVD for each individual.[44] Evidence indicates that a

genetic risk score estimated from these 300 SNPs is associated with traditional risk factors for CVD, such as type 2 diabetes and hypertension, contributes to the development of CVD conditions that have their origins in atherosclerosis, such as peripheral arterial disease and stroke, and is associated with premature mortality.[43] The PRS was estimated using PLINK, an open-source tool for genomic research,[45] by generating the sum of the number of risk alleles present at each locus and weighting by the log of the odds for that locus[18] estimated from the list of 300 SNPs using the PLINK "–score" command – with no-mean-imputation flag. PRS were available for all participants included in the final study sample, where PRS were transformed to standardised Z scores and were treated as a continuous variable in all modelling.

Demographic and health information

Information on demographics, medical history and health behaviours were collected using interview-administered questionnaires at recruitment and follow ups. Participant age at recruitment and sex were self-reported. No adjustments were made for discrepancies between self-reported sex and genetic sex. Education was assessed by asking "Which of the following qualifications do you have? (You can select more than one)," with the options college or university degree, A levels or equivalent, O levels or GCSEs or equivalent, CSEs, NVQ/HND/HNC, other professional qualifications (e.g., nursing or teaching). We operationalised this into 5 categories based on the highest level of education: i) college or university degree, ii) all professional qualification (NVQ/HND/HNC, other professional qualifications), iii) A levels or equivalent, iv) O levels, GCSEs or equivalent or CSEs and v) none of the above or prefer not to answer. The Townsend deprivation index, a composite measure of deprivation based on unemployment, non-car ownership, non-home ownership, and household overcrowding,[46] was estimated from the preceding national census data, with each participant assigned a score corresponding to the postcode of their home dwelling and a negative value representing high socioeconomic status. We operationalised the Townsend deprivation index as quintiles.

Information on smoking (never, previous and current smoker), previous doctor diagnosis of any type of diabetes or a CVD event (yes, no) and use of medication (anti-hypertensive, lipid-lowering or exogenous hormones or diabetes; yes, no) were collected. We created a

binary variable for family history (of father, mother and siblings) of CVD and related diseases (yes, no). Body Mass Index (BMI) was calculated as weight/height^2 collected during the Assessment Centre visit. We created a binary variable to indicate overweight or obese according to standard World Health Organisation cut offs.[47] Physical activity was estimated using Metabolic Equivalents (METs), the ratio of a person's working metabolic rate relative to their resting metabolic rate. One MET was defined as the energy cost of sitting quietly and is equivalent to a caloric consumption of 1kcal/kg/hour. We used standard cut-offs to categorised participants as meeting physical activity guidelines of 150 min per week if their METs were ≥ 600 MET-min/week.[48]

Statistical analysis

Complete case analysis was used. We investigated missingness by comparing demographic characteristics of the excluded sample with the analytic sample. Descriptive analyses included mean (SD) for continuous variables and number (%) for categorical variables. We created sample-based tertiles of diet quality for RFS, HDI and MDS for descriptive purposes only. Unadjusted linear regression analyses were used to examine intakes of encouraged food groups and total energy and nutrient intakes across tertiles of diet quality indices. This descriptive analysis aimed to show that diet quality scores reflect differences in underlying food and nutrient intakes, thus assisting with interpretation and translation into actual food intakes.

We used multivariable Cox proportional hazard regression models to estimate hazard ratios (HR) and 95% Confidence Intervals (CI) of all-cause mortality, CVD mortality and risk of CVD events (MI and stroke) according to each diet quality index separately (RFS, HDI and MDS). We treated diet quality indices as continuous independent variables. CVD events and mortality were treated as time-to-event outcome/dependent variables. We estimated the duration of follow up as the time between the last day of dietary data and the first event of either an MI, stroke, mortality, or the censoring date (4 March 2020). In participants who had multiple events during the study period, the first event date was used. We adjusted the Cox regression analyses for covariates identified using a directed acyclic graph (Supplemental Figure 1). These included age (continuous), sex, deprivation (categorical), smoking status (categorical), physical activity (continuous), medication use (binary), family

history of CVD (binary) and energy intake (continuous). The role of sex by diet quality and by PRS interactions were further tested by adding an interaction term to each model. Consistent with recommendations for sex differences in cardiovascular associations,[49] analyses were presented stratified by sex of whether there were any apparent sex differences. The Cox proportional hazards models also included PRS as an independent variable and were additionally adjusted for the first 8 principal components of ancestry and genotyping batch.[14] We included an interaction term in the models to test for statistical interaction between each diet quality score and PRS. Interactions were further inspected by conducting post-hoc estimation of the effects of diet quality indices on events at 'low' and 'high' PRS score of -1 and +1 which, given PRS was a standard score, represent minus and plus one SD of PRS. Data were analysed using Stata (version 16.0; StataCorp., College Station, TX, USA). To address possible reverse causation, sensitivity analyses excluded deaths and incident cases of MI and stroke within the first 2 years of follow up.

Patient and public involvement

The development of the research question or outcome measures was not informed by patients' priorities, experience, or preferences. No patients were involved in the design and conduct of the present study. There are no plans to disseminate the results to study participants.

RESULTS

Of the 502,536 participants who were recruited into the UK Biobank, 425,529 participants were excluded based on being not white British (n=92,907), having unusable genetic data (n=1,459), being ineligible (n=23,215) or missing dietary or covariate data (n=307,951; Supplemental Figure 2). Excluded participants were similar in age and sex to the included sample, with somewhat higher BMI and rates of smoking and deprivation (Supplemental Table 4). A total of 77,004 participants were included in the present analysis (Table 1). Mean age at recruitment was 56.2 (SD 7.8) years and 55% were female. Forty-eight per cent of participants had a colleague or university degree, most were experiencing low to mid deprivation (67%), had never smoked (58%) and had a family history of CVD (74%; Table 1).

Fifty-nine per cent of the participants were overweight or obese and 85% met physical activity guidelines.

Mean RFS, HDI and MDS were 6.78 (SD 2.40), 3.57 (SD 1.26) and 5.31 (SD 1.04), respectively. Intake of fruits, vegetables, wholegrains and lean meat were higher with increasing tertile of RFS, HDI and MDS (Table 2). Intake of low-fat dairy were higher with increasing tertile of RFS and HDI and lower with increasing tertile of MDS. Intakes of total fat, saturated fat, carbohydrates and sugars were higher with increasing tertile of RFS, HDI and MDS. Intakes of protein were lower with increasing tertile of HDI and MDS, while intakes were higher with increasing tertile of RFS. Intakes of PUFA were lower with increasing tertile of RFS, while intakes were higher with increasing tertile of HDI and MDS (Table 2).

Table 1. Baseline characteristics of participants in the UK Biobank

Characteristic	Overall N (%)	Males N (%)	Females N (%)
n	77,004	34,984 (45.4)	42,020 (54.5)
Age at recruitment (years), Mean ± SD	56.2 ± 7.8	57.0 ± 7.8	55.6 ± 7.7
Highest level of education			
College or University degree	36,709 (47.8)	17,271 (49.4)	19,438 (46.3)
A levels/AS levels or equivalent	10,580 (13.8)	4,390 (12.6)	6,190 (14.8)
O levels/GCSE/CSEs or equivalent	17,669 (23.0)	7,126 (20.4)	10,543 (25.1)
Professional qualifications	7,467 (9.7)	4,062 (11.6)	3,405 (8.1)
None/Prefer not to answer	4,454 (5.8)	2,082 (6.0)	2,372 (5.7)
Townsend Deprivation Index			
Least deprived	18,129 (23.5)	8,611 (24.6)	9,518 (22.7)
2nd least deprived	17,227 (22.4)	7,910 (22.6)	9,317 (22.2)
Medium deprivation	16,067 (20.9)	7,158 (20.5)	8,909 (21.2)
2nd most deprived	14,900 (19.3)	6,549 (18.7)	8,351 (19.9)
Most deprived	10,681 (13.9)	4,756 (13.6)	5,926 (14.1)
Smoking			
Never smoked	44,856 (58.3)	18,849 (53.9)	26,007 (61.9)
Ex-smoker	27,184 (35.3)	13,471 (38.5)	13,714 (32.6)
Current smoker	4,964 (6.5)	2,664 (7.6)	2,300 (5.5)
Body Mass Index (kg/m^2), Mean ± SD[1]	26.5 ± 4.4	27.1 ± 3.9	26.0 ± 4.7
Waist circumference (cm), Mean ± SD[1]	88.1 ± 13.0	95.2 ± 10.8	82.3 ± 11.6
Total PA (MET min), Mean ± SD	2477 ± 2326	2542 ± 2439	2423 ± 2227
Medication use[2]	16,573 (21.5)	9,713 (27.8)	6,860 (16.3)
Family history of CVD	57,211 (74.3)	25, 076 (71.7)	32,135 (76.5)
Energy Intake (kJ/day), Mean ± SD	8853 ± 2172	9574 ± 2253	8252 ± 1903

A levels/AS levels, Advanced levels/Advanced Subsidiary levels; O levels/GCSE/CSEs, Ordinary levels/General Certificate of Secondary Education/General Certificate of Education; Professional qualifications include NVQ (National Vocational Qualification)/HND (Higher National Diploma)/HNC (Higher National Certificate), other professional qualifications; Townsend Deprivation Index is a composite measure of deprivation based on unemployment, non-car ownership, non-home ownership, and household overcrowding.

1, Data on Body Mass Index and Waist Circumference were available in n=76,901 and n=76,950 respectively. 2, Medication use was restricted to lipid lowering or blood pressure.

Table 2. Baseline dietary intakes by tertile of diet quality index in the UK Biobank (n=77,004)[1]

	Recommended Food Score			Healthy Diet Indicator			Mediterranean Diet Score		
	T1	**T2**	**T3**	**T1**	**T2**	**T3**	**T1**	**T2**	**T3**
Diet quality	4.2 (1.1)	6.8 (0.6)	9.5 (1.3)	2.3 (0.6)	3.7 (0.3)	5.1 (0.7)	4.4 (0.6)	5.5 (0.2)	6.5 (0.5)
Food group intake (g/day)									
Whole fruit	81.6 (76.2)	152.7 (98.5)	231.1 (133.6)	95.3 (78.8)	156.8 (106.9)	225.9 (138.7)	115.4 (98.7)	170.7 (124.0)	198.7 (132.0)
Vegetables	120.0 (89.6)	201.8 (113.0)	286.1 (149.7)	146.7 (99.0)	204.7 (127.4)	267.5 (157.7)	158.5 (114.3)	220.4 (136.3)	250.6 (150.0)
Wholegrains	32.8 (38.2)	51.5 (46.3)	65.8 (52.9)	36.2 (38.8)	50.9 (46.2)	65.8 (54.7)	37.3 (39.0)	54.5 (49.1)	64.7 (54.2)
Lean meats	51.1 (45.4)	66.0 (48.9)	79.9 (54.4)	59.9 (48.0)	66.5 (51.0)	71.6 (53.9)	57.3 (48.1)	71.2 (51.8)	73.4 (52.8)
Low-fat dairy	18.3 (37.3)	29.4 (43.7)	45.0 (51.9)	22.7 (40.6)	30.9 (44.7)	40.6 (51.2)	31.8 (48.5)	32.7 (46.7)	27.6 (41.0)
Total EI (kJ/day)	8572 (2224)	8798 (2080)	9204 (2158)	8244 (2032)	9022 (2112)	9474 (2195)	8785 (2234)	8862 (2106)	8946 (2121)
Nutrient intake (% energy)									
Total fat	34.5 (6.1)	33.1 (5.7)	31.8 (5.7)	34.4 (5.9)	33.2 (5.7)	31.4 (5.8)	34.2 (5.8)	32.9 (5.7)	31.7 (5.9)
Saturated fat	13.6 (3.1)	12.7 (2.8)	11.8 (2.8)	13.7 (3.0)	12.7 (2.8)	11.4 (2.7)	13.8 (2.9)	12.7 (2.7=8)	11.4 (2.7)
PUFA	6.18 (2.0)	6.08 (1.9)	6.06 (1.9)	5.87 (2.0)	6.16 (1.9)	6.37 (1.9)	5.91 (1.9)	6.16 (1.9)	6.37 (2.0)
Carbohydrate	45.8 (7.7)	47.9 (7.1)	50.0 (6.8)	44.8 (7.1)	48.2 (6.7)	51.6 (6.7)	46.7 (7.5)	48.0 (7.3)	49.5 (7.1)
Total sugars	20.1 (6.1)	22.6 (5.6)	25.6 (5.7)	20.4 (5.8)	22.9 (5.7)	25.5 (6.0)	21.9 (6.3)	22.8 (6.2)	23.9 (6.0)
Protein	15.4 (3.1)	15.8 (2.9)	16.2 (2.9)	16.4 (3.2)	15.6 (2.9)	15.2 (2.7)	16.0 (3.1)	15.8 (2.9)	15.4 (2.9)

EI, Energy intake

1, Unadjusted linear regression analyses were used to examine linear trend across tertiles of diet quality index; p<0.001 for all associations. The total possible scores for the diet quality indices were: RFS between 0 to 21, HDI between 0 to 11 and MDS between 0 to 9.

During a mean follow-up of 7.8 years (a total of 600,193 person-years), we observed 1,141 new MI events and 748 new stroke events. During a mean follow-up of 7.9 years (a total of 604,431 person-years), we observed 364 deaths due to CVD and 2,409 all-cause deaths. Of these, the majority of MI (72%) and stroke (60%) events, and CVD (72%) and all-cause (59%) deaths were in males.

RFS and risk of all-cause mortality, CVD mortality, MI and stroke

The adjusted HR associated with a one-point higher RFS for all-cause mortality was 0.96 (95% CI: 0.94, 0.98), for CVD mortality was 0.94 (95% CI: 0.90, 0.98), for MI was 0.97 (95% CI: 0.95, 1.00) and for stroke was 0.94 (95% CI: 0.91, 0.98) (Table 3). There was limited evidence (all p-values > 0.1) of sex by diet interactions. When stratified by sex, associations were comparable in men, while there was only evidence of an association between RFS and all-cause mortality and stroke in females. The adjusted HR associated with a one-point higher PRS for MI was 1.33 (95% CI: 1.25, 1.41); when stratified by sex, there was evidence of a stronger association in males. When an interaction term for PRS was added to the models, there was no evidence (at the p<0.05 level) of interaction between RFS and PRS for any outcomes (p-interaction=0.40 [all-cause mortality], p-interaction=0.77 [CVD mortality], p-interaction=0.17 [MI], and p-interaction=0.10 [stroke]). The interaction of sex by PRS showed evidence that the effect of higher PRS on higher risk of MI was more pronounced for males (HR=1.21 [95% CI: 1.06, 1.37], p=.004). Effect sizes were consistent when deaths and incident cases of MI and stroke within the first 2 years of follow up were excluded (data not shown).

HDI and risk of all-cause mortality, CVD mortality, MI and stroke

The adjusted HR associated with a one-point higher HDI for all-cause mortality was 0.97 (95% CI: 0.93, 0.99). There was little evidence of associations between HDI and risk of CVD mortality, MI, or stroke (Table 4). There was limited evidence (all p-values > 0.1) of sex by diet interactions. When stratified by sex, there was evidence of an association between HDI and all-cause mortality in males only. The adjusted HR associated with a one-point higher PRS for MI was 1.33 (95% CI: 1.25, 1.41), which when stratified by sex, there was evidence

of a stronger association in males. When an interaction term for PRS was added to the models, there was no evidence of interaction between HDI and PRS for other outcomes (p-interaction=0.66 [all-cause mortality], p-interaction=0.86 [CVD mortality] and p-interaction=0.17 [stroke]). There was some evidence of interaction between HDI and PRS for MI events (p-interaction=0.049). While there was no evidence of an effect of HDI on MI for participants with low (-1 SD) PRS (HR=1.02 [95% CI: 0.95, 1.10], p=0.61), there was some evidence of an association between higher HDI and reduced risk of MI events for those with high (+1 SD) PRS (HR=0.93 [95% CI: 0.88, 0.99], p=0.017). The interaction of sex by PRS showed evidence that the effect of higher PRS on higher risk of MI was more pronounced for males (HR=1.21 [95% CI: 1.06, 1.37], p=.004). Effect sizes were consistent when deaths and incident cases of MI and stroke within the first 2 years of follow up were excluded (data not shown).

MDS and risk of all-cause mortality, CVD mortality, MI and stroke

The adjusted HR associated with a one-point higher MDS for all-cause mortality was 0.95 (95% CI: 0.91, 0.98) and for stroke was 0.93 (95% CI: 0.87, 1.00) (Table 5). There was limited evidence of associations between MDS and risk of CVD mortality and MI. There was limited evidence (all p-values > 0.1) of sex by diet interactions. When stratified by sex, there was evidence of an association between MDS and all-cause mortality and MI in males only. The adjusted HR associated with a one-point higher PRS for MI was 1.33 (95% CI: 1.25, 1.41); when stratified by sex, there was evidence of a stronger association in males. When an interaction term for PRS was added to the models, there was no evidence of interaction between MDS and PRS for other outcomes (p-interaction=0.58 [all-cause mortality], p-interaction=0.72 [CVD mortality] and p-interaction=0.12 [stroke]). There was evidence of interaction between MDS and PRS for MI events (p-interaction=0.026). While there was no evidence of an effect of MDS on MI for those with low (-1 SD) PRS (HR=1.03 [95% CI: 0.94, 1.12], p=0.56) there was strong evidence of an association between higher MDS and reduced risk of MI events for those with high (+1 SD) PRS (HR=0.91 [95% CI: 0.85, 0.97], p=0.004). The interaction of sex by PRS showed evidence that the effect of higher PRS on higher risk of MI was more pronounced for males (HR=1.21 [95% CI: 1.06, 1.37], p=.004).

Effect sizes were consistent when deaths and incident cases of MI and stroke within the first 2 years of follow up were excluded (data not shown).

Table 3. Cox-proportional hazard ratios and 95% CI for risk of all-cause mortality, CVD mortality and CVD events with increasing Recommended Food Score (RFS) and polygenic risk score in participants from the UK Biobank

	Overall (n=77,004)				Males (n=34,984)				Females (n=42,020)			
	No of Events	Hazard Ratio	95% CI	P-value	No of Events	Hazard Ratio	95% CI	P-value	No of Events	Hazard Ratio	95% CI	P-value
All-cause mortality	2,409				1,416				993			
RFS		0.96	(0.94, 0.98)	<0.001		0.95	(0.93, 0.97)	<0.001		0.98	(0.95, 1.00)	0.08
Polygenic risk score		1.00	(0.96, 1.04)	0.93		1.02	(0.97, 1.07)	0.53		0.98	(0.92, 1.05)	0.57
CVD mortality	364				263				101			
RFS		0.94	(0.90,0.98)	0.007		0.93	(0.88, 0.98)	0.011		0.96	(0.88, 1.05)	0.34
Polygenic risk score		1.08	(0.98, 1.20)	0.13		1.10	(0.98, 1.25)	0.11		1.03	(0.85, 1.25)	0.79
Myocardial Infarction	1,141				822				319			
RFS		0.97	(0.95, 1.00)	0.048		0.97	(0.94, 1.00)	0.045		0.99	(0.94, 1.04)	0.57
Polygenic risk score		1.33	(1.25, 1.41)	<0.001		1.40	(1.31, 1.50)	<0.001		1.16	(1.04, 1.29)	0.008
Stroke	748				447				301			
RFS		0.94	(0.91, 0.98)	0.001		0.95	(0.91, 0.99)	0.018		0.94	(0.89, 0.99)	0.012
Polygenic risk score		1.02	(0.94, 1.10)	0.68		0.97	(0.89, 1.07)	0.56		1.08	(0.96, 1.21)	0.29

CVD, cardiovascular disease; Models were adjusted for age (continuous), sex (when not used to stratify), deprivation (categorical), smoking status (categorical), physical activity (continuous), medication use (binary), family history of CVD (binary), energy intake (continuous) and the first 8 principal components of ancestry and genotyping batch. All models include main effects of diet quality and polygenic risk score but not interaction terms.

Table 4. Cox-proportional hazard ratios and 95% CI for risk of all-cause mortality, CVD mortality and CVD events with increasing Healthy Diet Indicator (HDI) and polygenic risk score in participants from the UK Biobank

	Overall (n=77,004)				Males (n=34,984)				Females (n=42,020)			
	No of Events	Hazard Ratio	95% CI	P-value	No of Events	Hazard Ratio	95% CI	P-value	No of Events	Hazard Ratio	95% CI	P-value
All-cause mortality	2,409				1,416				993			
HDI		0.97	(0.93, 0.99)	0.041		0.95	(0.91, 0.99)	0.039		0.98	(0.93, 1.03)	0.46
Polygenic risk score		1.00	(0.96, 1.04)	0.92		1.02	(0.96, 1.07)	0.54		0.98	(0.92, 1.05)	0.58
CVD mortality	364				263				101			
HDI		0.99	(0.90, 1.08)	0.76		0.94	(0.85, 1.05)	0.28		1.10	(0.94, 1.30)	0.23
Polygenic risk score		1.08	(0.98, 1.20)	0.13		1.10	(0.98, 1.25)	0.11		1.03	(0.85, 1.25)	0.79
Myocardial Infarction	1,141				822				319			
HDI		0.96	(0.92, 1.01)	0.12		0.95	(0.89, 1.00)	0.06		1.00	(0.92, 1.10)	0.93
Polygenic risk score		1.33	(1.25, 1.41)	<0.001		1.40	(1.31, 1.50)	<0.001		1.16	(1.04, 1.29)	0.008
Stroke	748				447				301			
HDI		0.97	(0.91, 1.03)	0.25		0.96	(0.89, 1.04)	0.31		0.98	(0.89, 1.07)	0.63
Polygenic risk score		1.02	(0.94, 1.09)	0.68		0.97	(0.89, 1.07)	0.56		1.08	(0.96, 1.21)	0.19

CVD, cardiovascular disease; Models were adjusted for age (continuous), sex (when not used to stratify), deprivation (categorical), smoking status (categorical), physical activity (continuous), medication use (binary), family history of CVD (binary), energy intake (continuous) and the first 8 principal components of ancestry and genotyping batch. All models include main effects of diet quality and polygenic risk score but not interaction terms.

Table 5. Cox-proportional hazard ratios and 95% CI for risk of all-cause mortality, CVD mortality and CVD events with increasing Mediterranean Diet Score (MDS) and polygenic risk score in participants from the UK Biobank

	Overall (n=77,004)				Males (n=34,984)				Females (n=42,020)			
	No of Events	Hazard Ratio	95% CI	P-value	No of Events	Hazard Ratio	95% CI	P-value	No of Events	Hazard Ratio	95% CI	P-value
All-cause mortality	2,409				1,416				993			
MDS		0.95	(0.91, 0.98)	0.005		0.93	(0.88, 0.98)	0.004		0.97	(0.91, 1.03)	0.33
Polygenic risk score		1.00	(0.96, 1.04)	0.91		1.02	(0.97, 1.07)	0.51		0.98	(0.92, 1.05)	0.58
CVD mortality	364				263				101			
MDS		0.97	(0.88, 1.08)	0.60		0.94	(0.84, 1.06)	0.32		1.07	(0.88, 1.29)	0.49
Polygenic risk score		1.08	(0.98, 1.20)	0.12		1.10	(0.98, 1.25)	0.11		1.03	(0.85, 1.25)	0.79
Myocardial Infarction	1,141				822				319			
MDS		0.95	(0.90, 1.00)	0.06		0.94	(0.88, 1.00)	0.049		0.97	(0.88, 1.08)	0.64
Polygenic risk score		1.33	(1.25, 1.41)	<0.001		1.40	(1.31, 1.50)	<0.001		1.16	(1.04, 1.29)	0.008
Stroke	748				447				301			
MDS		0.93	(0.87, 1.00)	0.037		0.93	(0.85, 1.01)	0.10		0.93	(0.83, 1.04)	0.21
Polygenic risk score		1.02	(0.95, 1.09)	0.67		0.97	(0.89, 1.07)	0.57		1.08	(0.96, 1.21)	0.19

CVD, cardiovascular disease; Models were adjusted for age (continuous), sex (when not used to stratify), deprivation (categorical), smoking status (categorical), physical activity (continuous), medication use (binary), family history of CVD (binary), energy intake (continuous) and the first 8 principal components of ancestry and genotyping batch. All models include main effects of diet quality and polygenic risk score but not interaction terms.

DISCUSSION

This prospective population-based cohort study aimed to examine the association of three diet quality indices (RFS, HDI, MDS) and a genetic risk score with incidence of CVD and mortality. Our main findings were that higher RFS, HDI and MDS were associated with lower risk of mortality, regardless of genetic CVD risk. However, only the RFS showed evidence of lower risk of CVD mortality, MI and stroke, suggesting the applicability of the diet quality indices may depend on the health outcome in question. We also identified that increasing genetic risk of CVD was associated with MI only. There was some evidence suggesting that the underlying genetics of both MI and death may follow different pathways in males and females. Interaction analyses suggested that following a healthy diet may be of particular importance for reducing risk of MI in individuals with high genetic risk of CVD. Nonetheless, our findings demonstrate the benefit of following a healthy diet independent of genetic risk.

Our findings for reduced risk of all-cause mortality with higher diet quality are consistent with previous research on the MDS,[5 9 50 51] HDI[50 52] and RFS.[53] Moreover, a comparison of 10 diet quality indices in over 450,000 European adults showed that all indices examined were inversely associated with 10-year risk of all-cause mortality.[52] In the present study, the predictive role of diet quality on risk of all-cause mortality remained after adjusting for major non-modifiable determinants of all-cause mortality, including age, sex and family history of CVD. This highlights the importance of modifiable risk factors for death, regardless of whether the diet quality index is based on intakes of encouraged foods (i.e. RFS), foods and nutrients from dietary guidelines (i.e. HDI) or a dietary pattern identified as healthful (i.e. MDS). Moreover, the common elements across all three indices is the inclusion of food-based components, such as fruit and vegetables and lean meat and alternatives, rather than nutrients, affirming the value of food-based dietary guidelines in preventative healthcare rather a reductionist nutrient-based approach.[54]

Evidence for an association between diet quality indices and CVD risk is mixed.[5 7 11 14 32 55-57] Confirming our findings, large-scale studies in the UK population have shown independent associations between healthy diets and lifestyles and low genetic risk in reducing risk of CVD, with mixed results for interactions.[14 57] Only one study to date has used an overall diet quality index,[11] with comparable results to the present study, highlighting the potential to include plant-based diet quality components when assessing diet-disease associations.[56 58 59]

Future diet-disease research should extend this to better understand the role of specific plant and animal foods as part of overall dietary patterns. Moreover, the UK Biobank participants are predominately white and highly-educated and genetic risk profiles and their associations with risk of CVD may be different in more diverse populations. Thus, further research in diverse populations is needed to investigate the applicability of these diet quality methodologies for examining CVD risk independent of genetic risk.[7]

Although there was limited evidence of interactions for sex by diet quality index, our stratified results showed large effect sizes for associations between diet quality and genetic CVD risk in males. This warrants further investigation as previous research shows stronger associations in males than females.[14 60] The role of sex may be partly explained by the high prevalence of diabetes and unhealthy behaviours in men,[61] and in this study may be due the lower number of events and deaths in women compared with men in the present study. Nonetheless, it is possible that the biological and behavioural pathways in which risk factors exert their effects on CVD risk are different between men and women.[60]

Strengths and limitations

Our main strength was the large sample size and inclusion of genetic data. This enabled investigation of a genetic risk score created 300 SNPs known to be associated with CVD, more than any previous publications in the UK Biobank.[14 57] While the PRS used was specific to coronary disease, it has been used to identify predispositions to a wide variety of CVD and non-CVDs, as well as premature mortality, given these may develop in parallel with coronary disease for the same genetic origins. The dietary questionnaire has been previously validated and included sufficient detail to allow us to create three contrasting diet quality indices. There are a number of limitations that should be acknowledged. While the dietary assessment method is a short-term measure of intake, our use of up to four instances of dietary assessments provided an estimate of longer-term intake. Although the present analysis is likely to be subject to self-selection bias associated with the number of participants who completed the dietary assessment and the low response rate, associations between demographic and behavioural risk factors and mortality in the UK Biobank have been shown to be comparable to those from national health survey data from England and Scotland.[62] Whilst we adjusted analyses based on a range of confounders identified using a

directed acyclic graph, we cannot discount the possibility of residual or unmeasured confounding.

Conclusion

This prospective population-based cohort study provided evidence that higher diet quality (RFS, HDI and MDS) was associated with lower risk of all-cause mortality, regardless of genetic CVD risk. Diet quality, when estimated using the RFS only, was associated with lower risk of CVD mortality and MI, independent of genetic CVD risk. The diet quality indices investigated in this study have common food-based scoring components, providing further evidence for the best practice design and implementation of food-based diet quality indices for assessing health outcomes. Further research in diverse populations is needed to investigate the applicability of different diet quality methodologies for examining CVD risk independent of genetic susceptibility.

Authors' contributions: KML, GA, SB, JW, CM and SAM designed the analysis. KML, SB, JW and GA conducted the statistical analysis. KML drafted the manuscript. All authors contributed to the critical review of the manuscript and approved the final version of the manuscript.

Funding: KML is supported by a National Health and Medical Research Council Emerging Leadership Fellowship (APP1173803). JW is funded by the Lister Prize Fellowship (173096). The funding source had no role in the design or conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; KML is a consultant for HeadUp Labs; no other financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: UK Biobank received ethical approval from the research ethics committee (reference 13/NW/0382). All participants provided informed consent to participate. An

ethics exemption was granted by Deakin University Human Research Ethics Committee (Reference number 2019_293).

Data availability statement: The genetic and phenotypic UK Biobank data are available on application to the UK Biobank. This research has been conducted using the UK Biobank Resource under Application 34894.

Transparency: The lead author (KML) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted and that discrepancies from the study as planned have been explained

REFERENCES

1. Roth GA, Abate D, Abate KH, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2018;392(10159):1736-88. doi: https://doi.org/10.1016/S0140-6736(18)32203-7
2. Joseph P, Leong D, McKee M, et al. Reducing the Global Burden of Cardiovascular Disease, Part 1: The Epidemiology and Risk Factors. *Circ Res* 2017;121(6):677-94. doi: 10.1161/circresaha.117.308903 [published Online First: 2017/09/02]
3. Afshin A, Sur PJ, Fay KA, et al. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2019;393(10184):1958-72. doi: 10.1016/S0140-6736(19)30041-8
4. Sotos-Prieto M, Bhupathiraju SN, Mattei J, et al. Changes in Diet Quality Scores and Risk of Cardiovascular Disease Among US Men and Women. 2015;132(23):2212-19. doi: doi:10.1161/CIRCULATIONAHA.115.017158
5. Miller V, Webb P, Micha R, et al. Defining diet quality: a synthesis of dietary quality metrics and their validity for the double burden of malnutrition. *The Lancet Planetary Health* 2020;4(8):e352-e70. doi: 10.1016/S2542-5196(20)30162-5
6. Tapsell LC, Neale EP, Satija A, et al. Foods, Nutrients, and Dietary Patterns: Interconnections and Implications for Dietary Guidelines. *Advances in Nutrition: An International Review Journal* 2016;7(3):445-54. doi: 10.3945/an.115.011718
7. Aljuraiban GS, Gibson R, Oude Griep LM, et al. Perspective: The Application of A Priori Diet Quality Scores to Cardiovascular Disease Risk—A Critical Evaluation of Current Scoring Systems. *Adv Nutr* 2019;11(1):10-24. doi: 10.1093/advances/nmz059 %J Advances in Nutrition
8. Aigner A, Becher H, Jacobs S, et al. Low diet quality and the risk of stroke mortality: the multiethnic cohort study. *Eur J Clin Nutr* 2018;72(7):1035-45. doi: 10.1038/s41430-018-0103-4
9. Sotos-Prieto M, Bhupathiraju SN, Mattei J, et al. Association of Changes in Diet Quality with Total and Cause-Specific Mortality. *N Engl J Med* 2017;377(2):143-53. doi: 10.1056/NEJMoa1613502
10. Kant AK, Schatzkin A, Graubard BI, et al. A prospective study of diet quality and mortality in women. *JAMA* 2000;283(16):2109-15. doi: 10.1001/jama.283.16.2109
11. Heianza Y, Zhou T, Sun D, et al. Genetic susceptibility, plant-based dietary patterns, and risk of cardiovascular disease. *The American Journal of Clinical Nutrition* 2020 doi: 10.1093/ajcn/nqaa107
12. Liese AD, Krebs-Smith SM, Subar AF, et al. The Dietary Patterns Methods Project: synthesis of findings across cohorts and relevance to dietary guidance. *The Journal of nutrition* 2015;145(3):393-402. doi: 10.3945/jn.114.205336 [published Online First: 2015/01/21]
13. Voruganti VS. Nutritional Genomics of Cardiovascular Disease. *Curr Genet Med Rep* 2018;6(2):98-106. doi: 10.1007/s40142-018-0143-z [published Online First: 2018/04/30]
14. Rutten-Jacobs LC, Larsson SC, Malik R, et al. Genetic risk, incident stroke, and the benefits of adhering to a healthy lifestyle: cohort study of 306 473 UK Biobank participants. *BMJ : British Medical Journal* 2018;363:k4168. doi: 10.1136/bmj.k4168 %J BMJ
15. Ahmad T, Chasman DI, Mora S, et al. The Fat-Mass and Obesity-Associated (FTO) gene, physical activity, and risk of incident cardiovascular events in white women. *American Heart Journal* 2010;160(6):1163-69. doi: https://doi.org/10.1016/j.ahj.2010.08.002
16. Song Y, Stampfer MJ, Liu S. Meta-Analysis: Apolipoprotein E Genotypes and Risk for Coronary Heart Disease. *Annals of Internal Medicine* 2004;141(2):137-47. doi: 10.7326/0003-4819-141-2-200407200-00013 %J Annals of Internal Medicine
17. Cambien F, Tiret L. Genetics of Cardiovascular Diseases. 2007;116(15):1714-24. doi: doi:10.1161/CIRCULATIONAHA.106.661751

18. Sugrue LP, Desikan RS. What Are Polygenic Scores and Why Are They Important? *JAMA* 2019;321(18):1820-21. doi: 10.1001/jama.2019.3893 %J JAMA
19. Sudlow C, Gallacher J, Allen N, 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. doi: 10.1371/journal.pmed.1001779
20. Bradbury KE, Young HJ, Guo W, et al. Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire. *J Nutr Sci* 2018;7:e6-e6. doi: 10.1017/jns.2017.66
21. Lachat C, Hawwash D, Ocké MC, et al. Strengthening the Reporting of Observational Studies in Epidemiology—Nutritional Epidemiology (STROBE-nut): An Extension of the STROBE Statement. *PLoS Med* 2016;13(6):e1002036. doi: 10.1371/journal.pmed.1002036
22. Liu B, Young H, Crowe FL, 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. doi: 10.1017/S1368980011000942 [published Online First: 2011/06/01]
23. Greenwood DC, Hardie LJ, Frost GS, et al. Validation of the Oxford WebQ Online 24-Hour Dietary Questionnaire Using Biomarkers. *Am J Epidemiol* 2019;188(10):1858-67. doi: 10.1093/aje/kwz165 %J American Journal of Epidemiology
24. Food Standards Agency. Food Portion Sizes. London: The Stationary Office2005.
25. Royal Society of Chemistry; Ministry of Agriculture FaF. McCance and Widdowson's The Composition of Foods. 6th ed. Cambridge, United Kingdom: Royal Society of Chemistry; 2002.
26. Kant AK, Graubard BI. A Comparison of Three Dietary Pattern Indexes for Predicting Biomarkers of Diet and Disease. *J Am Coll Nutr* 2005;24(4):294-303. doi: 10.1080/07315724.2005.10719477
27. Maynard M, Gunnell D, Ness AR, et al. What influences diet in early old age? Prospective and cross-sectional analyses of the Boyd Orr cohort. *Eur J Public Health* 2005;16(3):315-23. doi: 10.1093/eurpub/cki167
28. Panagiotakos DB, Pitsavos C, Arvaniti F, et al. Adherence to the Mediterranean food pattern predicts the prevalence of hypertension, hypercholesterolemia, diabetes and obesity, among healthy adults; the accuracy of the MedDietScore. *Preventive Medicine* 2007;44(4):335-40. doi: https://doi.org/10.1016/j.ypmed.2006.12.009
29. Waijers PMCM, Feskens EJM, Ocké MC. A critical review of predefined diet quality scores. *British Journal of Nutrition* 2007;97(02):219-31. doi: doi:10.1017/S0007114507250421
30. Kanauchi M, Kanauchi K. The World Health Organization's Healthy Diet Indicator and its associated factors: A cross-sectional study in central Kinki, Japan. *Prev Med Rep* 2018;12:198-202. doi: https://doi.org/10.1016/j.pmedr.2018.09.011
31. McNaughton SA, Bates CJ, Mishra GD. Diet quality is associated with all-cause mortality in adults aged 65 years and older. *J Nutr* 2012;142(2):320-5. doi: 10.3945/jn.111.148692 [published Online First: 2011/12/23]
32. Livingstone KM, McNaughton SA. Diet quality is associated with obesity and hypertension in Australian adults: a cross sectional study. *BMC Public Health* 2016;16(1):1037. doi: 10.1186/s12889-016-3714-5
33. McCullough ML, Feskanich D, Stampfer MJ, et al. Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. *Am J Clin Nutr* 2002;76(6):1261-71. doi: 10.1093/ajcn/76.6.1261 [published Online First: 2002/11/27]
34. World Health Organization. Healthy diet WHO Fact Sheet, No 394 (2015) 2020 [Available from: http://www.who.int/mediacentre/factsheets/fs394/en/ accessed 9 March 2020.
35. Huijbregts P, Feskens E, Räsänen L, et al. Dietary pattern and 20 year mortality in elderly men in Finland, Italy, and the Netherlands: longitudinal cohort study. 1997;315(7099):13-17. doi: 10.1136/bmj.315.7099.13 %J BMJ

36. Trichopoulou A, Orfanos P, Norat T, et al. Modified Mediterranean diet and survival: EPIC-elderly prospective cohort study. 2005;330(7498):991. doi: 10.1136/bmj.38415.644155.8F %J BMJ
37. Trichopoulou A, Costacou T, Bamia C, et al. Adherence to a Mediterranean Diet and Survival in a Greek Population. 2003;348(26):2599-608. doi: 10.1056/NEJMoa025039
38. Bach-Faig A, Berry EM, Lairon D, et al. Mediterranean diet pyramid today. Science and cultural updates. *Public Health Nutr* 2011;14(12A):2274-84. doi: 10.1017/S1368980011002515 [published Online First: 2011/12/13]
39. Magliano D, Liew D, Pater H, et al. Accuracy of the Australian National Death Index: comparison with adjudicated fatal outcomes among Australian participants in the Long-term Intervention with Pravastatin in Ischaemic Disease (LIPID) study. 2003;27(6):649-53. doi: 10.1111/j.1467-842X.2003.tb00615.x
40. Schnier C, Bush K, Nolan J, et al. Definitions of Stroke for UK Biobank Phase 1 Outcomes Adjudication Documentation prepared by: On behalf of UK Biobank Outcome Adjudication Group. 2017. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_mi.pdf (accessed 16 September 2020).
41. Schnier C, Bush K, Nolan J, et al. Definitions of Acute Myocardial Infarction and Main Myocardial Infarction Pathological Types UK Biobank Phase 1 Outcomes Adjudication Documentation prepared by: On behalf of UK Biobank Outcome Adjudication Group. 2017. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_mi.pdf (accessed 16 September 2020).
42. National Health Service. COVID-19 Daily Deaths. 2020. https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-daily-deaths/ (accessed 16 September 2020).
43. Ntalla I, Kanoni S, Zeng L, et al. Genetic Risk Score for Coronary Disease Identifies Predispositions to Cardiovascular and Noncardiovascular Diseases. *J Am Coll Cardiol* 2019;73(23):2932-42. doi: https://doi.org/10.1016/j.jacc.2019.03.512
44. Dudbridge F. Polygenic Epidemiology. 2016;40(4):268-72. doi: 10.1002/gepi.21966
45. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81(3):559-75. doi: 10.1086/519795 [published Online First: 2007/08/19]
46. Towsend P, Phillimore P, Beattie A. Health and Deprivation. *Nurs Stand* 1988;2(17):34. doi: 10.7748/ns.2.17.34.s66 [published Online First: 1988/01/30]
47. World Health Organization. BMI classification 2015 [updated 21 July 2015. Available from: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html accessed 3rd December 2015.
48. The IPAQ Group. IPAQ scoring protocol - International Physical Activity Questionnaire. Available at: https://sites.google.com/site/theipaq/scoring-protocol. Accessed 22 May 2020. [
49. Woodward M. Rationale and tutorial for analysing and reporting sex differences in cardiovascular associations. *Heart* 2019;105(22):1701-08. doi: 10.1136/heartjnl-2019-315299
50. Neelakantan N, Koh W-P, Yuan J-M, et al. Diet-Quality Indexes Are Associated with a Lower Risk of Cardiovascular, Respiratory, and All-Cause Mortality among Chinese Adults. *The Journal of Nutrition* 2018;148(8):1323-32. doi: 10.1093/jn/nxy094 %J The Journal of Nutrition
51. Stefler D, Malyutina S, Kubinova R, et al. Mediterranean diet score and total and cardiovascular mortality in Eastern Europe: the HAPIEE study. *Eur J Nutr* 2017;56(1):421-29. doi: 10.1007/s00394-015-1092-x
52. Lassale C, Gunter MJ, Romaguera D, et al. Diet Quality Scores and Prediction of All-Cause, Cardiovascular and Cancer Mortality in a Pan-European Cohort Study. *PLoS One* 2016;11(7):e0159025. doi: 10.1371/journal.pone.0159025
53. Kaluza J, Håkansson N, Brzozowska A, et al. Diet quality and mortality: a population-based prospective study of men. *Eur J Clin Nutr* 2009;63(4):451-57. doi: 10.1038/sj.ejcn.1602968
54. Schulze MB, Martínez-González MA, Fung TT, et al. Food based dietary patterns and chronic disease prevention. *BMJ* 2018;361 doi: 10.1136/bmj.k2396

55. Lakkur S, Judd SE. Diet and Stroke. 2015;46(7):2007-11. doi: doi:10.1161/STROKEAHA.114.006306
56. Chiavaroli L, Viguiliouk E, Nishi SK, et al. DASH Dietary Pattern and Cardiometabolic Outcomes: An Umbrella Review of Systematic Reviews and Meta-Analyses. *Nutrients* 2019;11(2) doi: 10.3390/nu11020338 [published Online First: 2019/02/16]
57. Said MA, Verweij N, van der Harst P. Associations of Combined Genetic and Lifestyle Risks With Incident Cardiovascular Disease and Diabetes in the UK Biobank Study. *JAMA Cardiol* 2018;3(8):693-702. doi: 10.1001/jamacardio.2018.1717 [published Online First: 2018/06/30]
58. Baden MY, Liu G, Satija A, et al. Changes in Plant-Based Diet Quality and Total and Cause-Specific Mortality. *Circulation* 2019;140(12):979-91. doi: doi:10.1161/CIRCULATIONAHA.119.041014
59. Huang T, Yang B, Zheng J, et al. Cardiovascular Disease Mortality and Cancer Incidence in Vegetarians: A Meta-Analysis and Systematic Review. *Ann Nutr Metab* 2012;60(4):233-40. doi: 10.1159/000337301
60. Den Ruijter HM, Haitjema S, Asselbergs FW, et al. Sex matters to the heart: A special issue dedicated to the impact of sex related differences of cardiovascular diseases. *Atherosclerosis* 2015;241(1):205-07. doi: 10.1016/j.atherosclerosis.2015.05.003
61. Andersen KK, Andersen ZJ, Olsen TS. Age- and Gender-Specific Prevalence of Cardiovascular Risk Factors in 40 102 Patients With First-Ever Ischemic Stroke. *Stroke* 2010;41(12):2768-74. doi: doi:10.1161/STROKEAHA.110.595785
62. Batty GD, Gale CR, Kivimäki M, et al. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *BMJ* 2020;368:m131. doi: 10.1136/bmj.m131

Supplemental Table 1. Components and scoring methods of the Recommended Food Score (RFS)

Dietary Indictor	Indicator food groups[1]	Criteria for scoring
1. Fruits	1. Pome fruit (apples, pears) 2. Berry fruit (berry) 3. Citrus fruit (orange, satsuma, grapefruit) 4. Stone fruit (nectarine, peach, plum, cherry, prune) 5. Tropical and subtropical fruit (banana, pineapple, mango) 6. Other fruit (other fruit, grape, melon, dried fruit, stewed fruit)	Indicator food groups were assigned a score of 1 if they were consumed above the minimum threshold of 15 g/d.
	7. Fruit juice (orange juice, grapefruit juice, pure fruit/vegetable juice)	Fruit juice was assigned a score of 1 if it was consumed above the minimum threshold of 30 g/d.
2. Vegetables	1. Green (lettuce, spinach, sprouts, watercress, cucumber, celery, courgette) and brassica vegetables (cabbage, cauliflower, broccoli) 2. Legumes (pulses, broad bean) 3. Carrot and root vegetables (carrot, turnip/swede, beetroot parsnip, onion, garlic, leek) 4. Starchy vegetables (boiled/baked potatoes [*butter/margarine added to potatoes, butternut squash], mashed potato, sweet potato, sweetcorn) 5. Tomato and tomato products (fresh tomato, tinned tomato) 6. Peas and beans (green bean, pea) 7. Other vegetables (other vegetables, mushroom, sweet pepper, side salad, olives)	Indicator food groups were assigned a score of 1 if they were consumed above the minimum threshold of 15 g/d.
3. Whole grains	1. Wholegrain (whole-wheat cereal, sliced bread (wholemeal), baguette (wholemeal), bap (wholemeal), bread roll [wholemeal]) 2. High fibre cereals (porridge, muesli, oat crunch, bran cereal) and wholegrain pasta and brown rice	Indicator food groups were assigned a score of 1 if they were consumed above the minimum threshold of 15 g/d.
4. Lean meats and alternatives	1. Poultry 2. Fish (tinned tuna, oily fish, white fish, prawns, lobster/crab, shellfish) 3. Alternatives (whole egg, omelette, egg in sandwich, other egg, seed (e.g. unsalted peanuts, unsalted nuts, types of spreads/sauces consumed [peanut butter] seeds), tofu, Quorn)	Indicator food groups were assigned a score of 1 if they were consumed above the minimum threshold of 15 g/d.
5. Low-fat dairy	1. 2%, 1% or skim milk (type of milk consumed (semi skimmed, skimmed, goat/sheep milk, powdered milk, cholesterol lowering))	Milk was assigned a score of 1 if it was consumed above the minimum threshold of 30 g/d.
	2. Low fat cheese and yogurt (Low fat hard cheese, low fat cheese spread, cottage cheese, yogurt [low fat yogurt consumer], goat's cheese)	Indicator food groups were assigned a score of 1 if they were consumed above the minimum threshold of 15 g/d.

1. As available in the UK Biobank

Supplemental Table 2. Components and scoring methods of the Healthy Diet Indicator (HDI)

Dietary Indictor	Indicator foods[1]	Criteria for scoring
1. Saturated fatty acids	Saturated fat	>10% energy intake=0 0-10% energy intake=1
2. Polyunsaturated fatty acids	Polyunsaturated fat	<6 or >10% energy intake=0 6-10% energy intake=1
3. Protein	Protein	<10 or >15% energy intake=0 10-15% energy intake=1
4. Total carbohydrates	Carbohydrates	<50% or >70% energy intake=0 50-70% energy intake=1
5. Dietary fibre	Englyst dietary fibre	<18 or >32 g/day=0 18-32 g/day =1
6. Fruits and vegetables	Mixed vegetable, vegetable pieces, avocado, beetroot, broccoli, butternut squash, cabbage/kale, carrot, cauliflower, celery, courgette, cucumber, garlic, leek, lettuce, mushroom, onion, olives, parsnip, pea, side salad, sweet pepper, spinach, sprouts, sweetcorn, fresh tomato, tinned tomato, green bean, turnip/swede, watercress, other vegetables, homemade soup (vegetables) Stewed fruit, prune, dried fruit, mixed fruit, apple, banana, berry, cherry, grapefruit, grape, mango, melon, orange, satsuma, peach/nectarine intake, pear, pineapple, plum, other fruit	<400 g/day=0 ≥400 g/day=1
7. Pulses and nuts	Baked bean, pulses, broad bean Salted peanuts, unsalted peanuts, salted nuts, unsalted nuts, seeds, types of spreads/sauces consumed (peanut butter)	<30 g/day=0 ≥30 g/day=1
8. Total non-milk extrinsic sugars	Total sugars	>10 % energy intake=0 0-10 % energy intake=1
9. Fish	Tinned tuna, oily fish, white fish, prawns, lobster/crab, shellfish, other fish Homemade soup, ingredients in homemade soup (fish)	<32 g/day=0 ≥32 g/day=1
10. Red meat and meat products	Beef, pork, lamb, other meat Poultry intake (skin removed from poultry (no); fat removed from poultry(no)) Homemade soup, ingredients in homemade soup (meat) Sausage, bacon, ham, liver	>90 g/day=0 ≤90 g/day=1
11. Calcium	Calcium	<700 mg/day=0 ≥700 mg/day=1

1. As available in the UK Biobank

Supplemental Table 3. Components and scoring methods of the Mediterranean Diet Score (MDS)

Dietary Indictor	Indicator foods[1]	Criteria for scoring
1. Vegetables (excluding potatoes, legumes or fruit juice)	Mixed vegetable, vegetable pieces, avocado, beetroot, broccoli, butternut squash, cabbage/kale, carrot, cauliflower, celery, courgette, cucumber, garlic, leek, lettuce, mushroom, onion, olives, parsnip, pea, side salad, sweet pepper, spinach Sprouts, sweetcorn, fresh tomato, tinned tomato, green bean, turnip/swede, watercress, other vegetables, homemade soup (vegetables)	Sex-specific median intakes used as cut points. Intakes (for indictors 1-6) above median score 1 and intakes below the median score 0.
2. Legumes	Baked bean, pulses, broad bean, homemade soup (pulses)	
3. Fruit and nuts	Stewed fruit, prune, dried fruit, mixed fruit, apple, banana, berry, cherry, grapefruit, grape, mango, melon, orange, satsuma, peach/nectarine, pear intake, pineapple, plum, other fruit Orange juice, grapefruit juice, pure fruit/vegetable juice Unsalted peanuts, unsalted nuts, types of spreads/sauces consumed (Peanut butter), seeds	
4. Cereals	Porridge, muesli, oat crunch, plain cereal, bran cereal, whole-wheat cereal, other cereal Bread consumed, sliced bread (mixed; wholemeal; seeded; other), baguette (mixed; wholemeal; seeded; other), bap (mixed; wholemeal; seeded; other), bread roll (mixed; wholemeal; seeded; other), other bread White pasta, wholemeal pasta, white rice, brown rice, couscous, other grain Homemade soup, ingredients in homemade soup (pasta)	
5. Fish and seafood	Tinned tuna, oily fish, white fish, prawns, lobster/crab, shellfish, other fish Homemade soup, ingredients in homemade soup (fish)	
6. Monounsaturated/ saturated fats ratio	Monounsaturated fats, saturated fats	
7. Dairy products	Milk, milk added to cereal Low fat hard cheese, low fat cheese spread, cottage cheese Yogurt (low fat yogurt consumer; full fat yogurt consumer) Goat's cheese, hard cheese, soft cheese, blue cheese, cheese spread, feta, mozzarella, other cheese Dairy smoothie, latte, added milk to instant coffee, added milk to filtered coffee, added milk to espresso, added milk to other coffee type, added milk to standard tea, added milk to rooibos tea, cappuccino	Sex-specific median intakes used as cut points. Intakes (for indictors 7-8) below median score 1 and intakes below the median score 0.
8. Meat and meat products	Beef, pork, lamb, other meat Whole egg, omelette, eggs in sandwiches, scotch egg, other egg Homemade soup, ingredients in homemade soup (meat), sausage, bacon, ham	

9. Alcohol	Red wine, rose wine, white wine Beer/cider Fortified wine, spirits intake, other alcohol	No more than 2 drinks/day = 1; Never drink or over 2 drinks/day = 0.

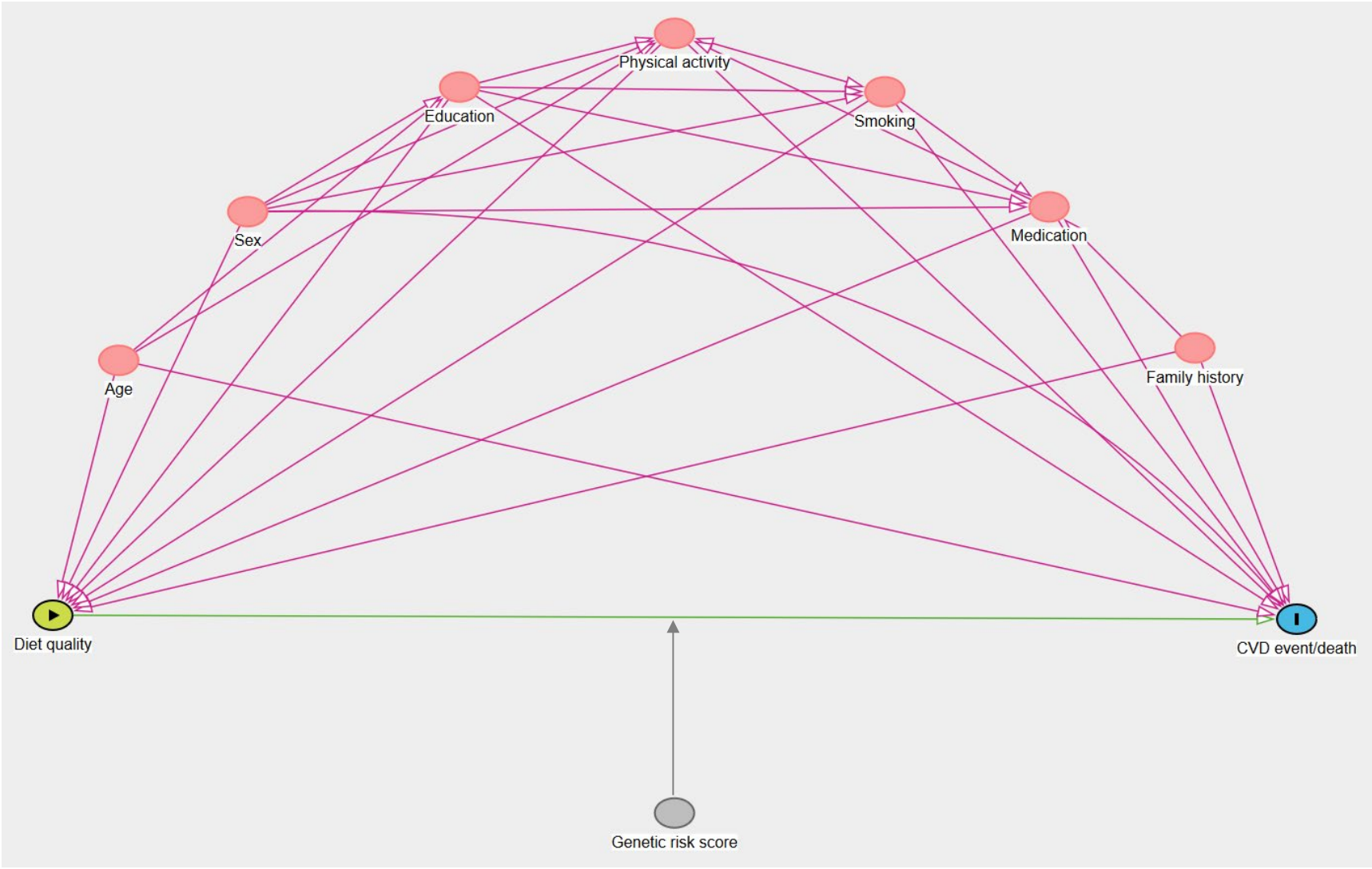
1. As available in the UK Biobank

Supplemental Table 4. Comparison of participant characteristics between the excluded and analytic sample

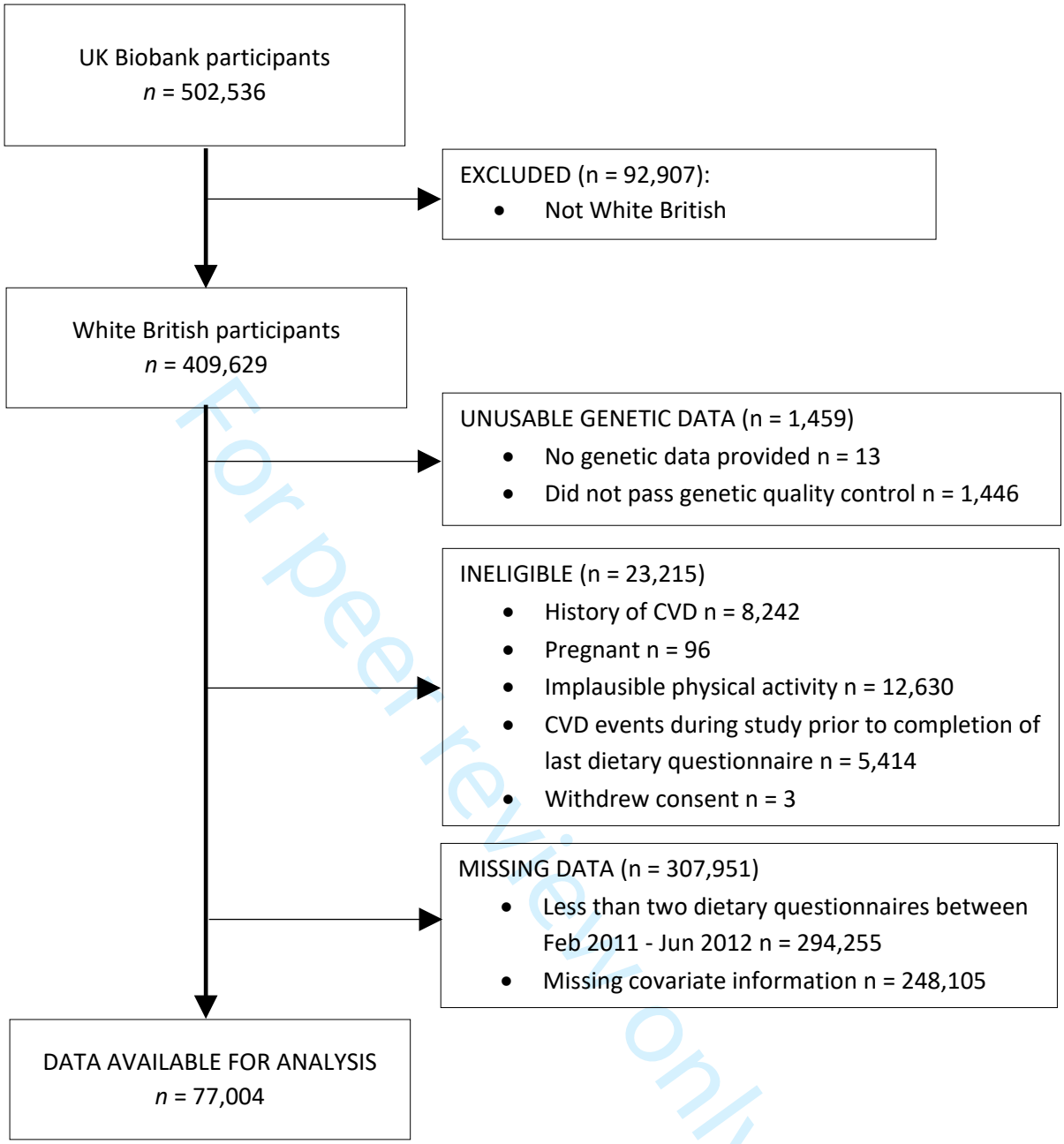
Characteristic	Excluded N (%)	Analytic N (%)
N[1]	425,532	77,004
Sex female	231,382 (54.4)	42,020 (54.6)
Age at recruitment (years), Mean ± SD	56.6 ± 8.2	56.2 ± 7.8
Townsend Deprivation Index		
Least deprived	82,535 (19.4)	18,129 (23.5)
2nd least deprived	82,878 (19.5)	17,227 (22.4)
Medium deprivation	84,323 (19.8)	16,067 (20.9)
2nd most deprived	85,475 (20.1)	14,900 (19.4)
Most deprived	89,698 (21.1)	10,681 (13.9)
Smoking		
Never smoked	228,689 (54.1)	44,856 (58.3)
Ex-smoker	145,891 (34.5)	27,184 (35.3)
Current smoker	48,016 (11.4)	4,964 (6.4)
Body Mass Index (kg/m^2), Mean ± SD	27.6 ± 4.9	26.5 ± 4.4

Townsend Deprivation Index is a composite measure of deprivation based on unemployment, non-car ownership, non-home ownership, and household overcrowding.

1, In the excluded sample, data on Townsend Deprivation Index and smoking were available in n=424,909 and n= 422,596, respectively. Data on Body Mass Index were available in n= 422,530 and n=76,901 in the excluded and analytic sample, respectively.



Supplemental Figure 1. Directed acyclic graph showing relationship between the exposure (diet quality) and outcome (CVD events/death). Confounders are represented by red dots. The moderator (polygenic risk score) is represent by a grey dot.



Supplemental Figure 2. Flow diagram of participants in the UK Biobank

Reviewer only Supplemental Table. STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	5
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-9
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-9
Bias	9	Describe any efforts to address potential sources of bias	10
Study size	10	Explain how the study size was arrived at	10
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6-10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	10-11
		(*b*) Describe any methods used to examine subgroups and interactions	10-11
		(*c*) Explain how missing data were addressed	10-11
		(*d*) If applicable, explain how loss to follow-up was addressed	10-11
		(*e*) Describe any sensitivity analyses	N/A
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	11
		(b) Give reasons for non-participation at each stage	11
		(c) Consider use of a flow diagram	11
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	11
		(b) Indicate number of participants with missing data for each variable of interest	11
		(c) Summarise follow-up time (eg, average and total amount)	11
Outcome data	15*	Report numbers of outcome events or summary measures over time	11
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	11-18
		(*b*) Report category boundaries when continuous variables were categorized	11-18
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	11-18
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11-18
Discussion			
Key results	18	Summarise key results with reference to study objectives	19
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	20
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	19-21
Generalisability	21	Discuss the generalisability (external validity) of the study results	19-21
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	21-22

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in

conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.