

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

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

Journal:	<i>BMJ Open</i>
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

SCHOLARONE™
Manuscripts



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 Abbott¹, Steven J. Bowe², Joey Ward³, Catherine Milte¹, Sarah A. McNaughton¹

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

1
2
3 1 **ABSTRACT**
4

5
6 2 **Objectives:** To examine associations of three diet quality indices and a polygenic risk score
7
8 3 with incidence of all-cause mortality, CVD-related mortality, myocardial infarction (MI) and
9
10 4 stroke.

11
12 5 **Design:** Prospective cohort study.

13
14 6 **Setting:** The UK Biobank, UK.

15
16
17 7 **Participants:** 77,004 men and women (40-70 years) recruited between 2006 and 2010.

18
19 8 **Main outcome measures:** Cox proportional hazard ratios (HR) were used to estimate effects
20
21 9 of diet quality and genetic risk on risk of all-cause mortality, CVD-related mortality, MI and
22
23 10 stroke. Dietary intake, assessed using the Oxford WebQ, was used to calculate the
24
25 11 Recommended Food Score (RFS), Healthy Diet Indicator (HDI) and Mediterranean Diet Score
26
27 12 (MDS). A polygenic risk score was created from 300 single nucleotide polymorphisms
28
29 13 associated with CVD to examine moderation effects.

30
31 14 **Results:** New deaths due to CVD (n=364) and all-cause (n=2,409), and MI (n=1,141) and
32
33 15 stroke (n=748) events were identified during mean follow-ups of 7.9 and 7.8 years,
34
35 16 respectively. The adjusted HR associated with one-point higher RFS was 0.96 (0.94, 0.98) for
36
37 17 all-cause mortality, 0.94 (0.90, 0.98) for CVD-related mortality, 0.97 (95% CI: 0.95, 1.00) for
38
39 18 MI and 0.94 (95% CI: 0.91, 0.98) for stroke. The adjusted HR for all-cause mortality
40
41 19 associated with one-point higher HDI and MDS was 0.97 (0.93 to 0.99) and 0.95 (0.91 to
42
43 20 0.98), respectively. The adjusted HR associated with one-point higher MDS was 0.93 (95%
44
45 21 CI: 0.87, 1.00) for stroke. There was little evidence of associations between HDI and risk of
46
47 22 CVD-related mortality, MI or stroke. There was only evidence of an interaction between diet
48
49 23 quality and genetic risk score for MI.

50
51 24 **Conclusion:** Higher diet quality (RFS, HDI and MDS) predicted lower risk of all-cause
52
53 25 mortality, independent of genetic risk. Higher RFS was also associated with lower risk of
54
55 26 CVD-related mortality and MI. These findings demonstrate the benefit of following a healthy
56
57 27 diet, regardless of genetic risk.

58
59
60 28

29

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.

30 INTRODUCTION

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

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

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

60 METHODS

61 **Study design and participants**

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

82 **Study measures**

83 Dietary intake

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

1
2
3 93 collected via four online cycles between February 2011 to June 2012: February 2011 to April
4 94 2011 (online cycle 1); June 2011 to September 2011 (online cycle 2); October 2011 to
5 95 December 2011 (online cycle 3); April 2012 to June 2012 (online cycle 4). The total period of
6 96 available dietary data from the Oxford WebQ was 38 months (Apr 2009 - Jun 2012). Email
7 97 invitations were sent on different days of the week to capture variation in dietary intakes
8 98 and participants were given 3 days to complete the questionnaire for cycles 1 and 2 and 14
9 99 days for cycles 3 and 4.

10
11
12
13
14
15
16 100 To establish a baseline dietary intake in the present analysis, we calculated a mean dietary
17 101 intake based on the four online Oxford WebQ cycles only. This was because the time
18 102 between the 1st and 4th online cycle measurements was 16 months (Feb 2011 - Jun 2012)
19 103 and was considered a more credible timeframe for an average baseline than the 38 months
20 104 available from all five Oxford WebQ measurements. This resulted in a minimal sample loss
21 105 (<10%) while providing a shorter dietary exposure period and a more consistent approach to
22 106 the use of the dietary data by using only the online cycles of the OxfordWebQ. To better
23 107 capture usual intake, we calculated average nutrient intakes, food group intakes and diet
24 108 quality scores for participants who had two or more valid measurements for the four online
25 109 cycles of the Oxford WebQ.

110 111 Diet quality

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

119 The RFS is a food-based variety index designed to assess consumption of food groups
120 encouraged in the dietary guidelines.¹⁰ As detailed in Supplemental Table 1, food intakes
121 were scored according to five food groups: fruits (7 items), vegetables (7 items),
122 wholegrains (2 items), lean meat and alternatives (3 items) and reduced fat dairy products
123 (2 items). Scoring was based on the RFS designed by Kant and Graubaud,²⁶ and has been

1
2
3 124 used elsewhere.^{31 32} We summed intakes of food items within each group to create a total
4
5 125 intake for each food group. Food groups were then assigned a score of 1 if they were
6
7 126 consumed above the minimum amount threshold: 15 g/d for non-beverages and 30 g/d for
8
9 127 beverages. Intakes below these thresholds were scored 0. Scores ranged between 0 and 21,
10
11 128 with higher scores indicating a higher quality diet and a wider consumption of
12
13 129 recommended foods.³³

14
15 130 The HDI is a food- and nutrient-based index designed to reflect consumption of foods
16
17 131 recommended for a healthy diet by the World Health Organisation.³⁴ The original HDI was
18
19 132 developed and validated in 1997 based on the 1990 World Health Organisation's dietary
20
21 133 recommendations for the prevention of chronic disease.³⁵ We adapted a 12-point Healthy
22
23 134 Diet Score designed by Maynard et al.²⁷ to reflect adherence to the 2020 World Health
24
25 135 Organisation healthy diet fact sheet.³⁴ As cholesterol intake is not part of the 2020
26
27 136 recommendations and information on its intake was not available in the UK Biobank, we
28
29 137 used an 11-item score that included the following groups: saturated fat; poly-unsaturated
30
31 138 fat; protein; total carbohydrates; dietary fibre; fruits and vegetables; pulses and nuts; total
32
33 139 non-milk extrinsic sugars; fish; red meat and meat products; and calcium. Data on intake of
34
35 140 non-milk extrinsic sugars was not available in the UK Biobank and so we adapted the HDI to
36
37 141 score intakes of total sugars instead. Criteria for scoring was based on cut points detailed in
38
39 142 Supplemental Table 2. We assigned intakes within the cut offs a score of 1 and those outside
40
41 143 of the cut offs were assigned a score of 0. The total score ranged from 0 to 11, with a higher
42
43 144 score reflecting a higher diet quality (Supplemental Table 2).

44
45 145 The MDS is a food- and nutrient-based score designed to reflect adherence to a
46
47 146 Mediterranean style diet. The present study used the 9-item index developed and validated
48
49 147 by Trichopoulou et al. as it is the first and most widely used version of the MDS.^{36 37} Food
50
51 148 and nutrient intakes were scored according to nine components: vegetables, legumes, fruits
52
53 149 and nuts, cereals, fish and seafood, monounsaturated fats to saturated fats ratio, dairy
54
55 150 products, meat and meat products and alcohol (Supplemental Table 3). As used by
56
57 151 Trichopoulou et al.,³⁶ we used sex-specific median intakes as cut off points for intakes of
58
59 152 each component. A score of 1 was assigned to participants whose intake of vegetables,
60
153 legumes, fruits and nuts, cereals, fish and seafood and monounsaturated: saturated fats was
154
above the median. A score of 1 was assigned to intake of dairy products, meat and meat

1
2
3 155 products below the median. For alcohol, a score of 1 was assigned for low to moderate
4
5 156 intake (intake of no more than 2 times/day). A score of 0 was assigned for no alcohol intake
6
7 157 or intake greater than 2 times per day.³⁸ Total MDS score ranged from 0 to 9, with higher
8
9 158 scores reflecting better alignment to the Mediterranean diet.

10
11 159

12 13 160 Cardiovascular events and mortality

14
15 161 Mortality status and causes of death were determined by linkage of data with the UK
16
17 162 National Death Index (NDI) using the most recent available data from the UK Biobank
18
19 163 (September 2020). The accuracy of the NDI for identifying CVD deaths has been established
20
21 164 previously.³⁹ CVD-related mortality was estimated from 2006 International Classification of
22
23 165 Diseases 10th revision (ICD-10) codes in death certificates. CVD-related mortality was
24
25 166 identified using ICD codes I05-I89. CVD events were recorded between enrolment (1999–
26
27 167 2000) and the most recent inpatient hospital data available from the UK Biobank
28
29 168 (September 2020). Incident MI (ST-Elevation Myocardial Infarction and Non-ST-Elevation
30
31 169 Myocardial Infarction) and stroke (ischaemic, intracerebral haemorrhage, and subarachnoid
32
33 170 haemorrhage) were available from algorithms provided by the UK Biobank.^{40 41} Algorithms
34
35 171 were produced to reliably identify incidence of selected illnesses through consideration of
36
37 172 hospital and death register data. The adjudication of “algorithmically defined” outcomes for
38
39 173 MI and stroke are detailed elsewhere.^{40 41} A censoring data of 4 March 2020 was used for all
40
41 174 outcomes. This date was chosen due to a spike in deaths from 5 March onwards, which is
42
43 175 likely to correspond to increasing deaths due to COVID-19 recorded in the UK.⁴²

44
45 176

46 47 177 Polygenic risk score

48
49 178 We used the March 2018 release of the imputed genetic data from UK Biobank
50
51 179 (downloaded 11 November 2019). From the resulting dataset, we excluded those who self-
52
53 180 reported ancestry other than white British, those who were missing more than 10% of the
54
55 181 genetic data and those who were defined by UK Biobank as being heterozygosity outliers.
56
57 182 Additionally, for every pair of who were individuals who were second cousins or closer (i.e.
58
59 183 those with a kinship coefficient > 0.042) one was excluded at random. We used information
60
184 on 300 single nucleotide polymorphisms (SNPs) known to be associated with coronary
185
185 185 artery disease⁴³ to create a PRS for CVD for each individual.⁴⁴ Evidence indicates that a

1
2
3 186 genetic risk score estimated from these 300 SNPs is associated with traditional risk factors
4
5 187 for CVD, such as type 2 diabetes and hypertension, contributes to the development of CVD-
6
7 188 related conditions that have their origins in atherosclerosis, such as peripheral arterial
8
9 189 disease and stroke, and is associated with premature mortality.⁴³ The PRS was estimated by
10
11 190 generating the sum of the number of risk alleles present at each locus and weighting by the
12
13 191 log of the odds for that locus¹⁸ estimated from the list of 300 SNPs using the plink “-score”
14
15 192 command – with no-mean-imputation flag. For participants included in the final study
16
17 193 sample, PRS were transformed to standardised Z scores and were treated as a continuous
18
19 194 variable in all modelling.
19
20 195

21 196 Demographic and health information

22
23 197 Information on demographics, medical history and health behaviours were collected using
24
25 198 interview-administered questionnaires at recruitment and follow ups. Participant age at
26
27 199 recruitment and sex were self-reported. No adjustments were made for discrepancies
28
29 200 between self-reported sex and genetic sex. Education was assessed by asking “Which of the
30
31 201 following qualifications do you have? (You can select more than one),” with the options
32
33 202 college or university degree, A levels or equivalent, O levels or GCSEs or equivalent, CSEs,
34
35 203 NVQ/HND/HNC, other professional qualifications (e.g., nursing or teaching). We
36
37 204 operationalised this into 5 categories based on the highest level of education: i) college or
38
39 205 university degree, ii) all professional qualification (NVQ/HND/HNC, other professional
40
41 206 qualifications), iv) A levels or equivalent, v) O levels, GCSEs or equivalent or CSEs and v)
42
43 207 none of the above or prefer not to answer. The Townsend deprivation index, a composite
44
45 208 measure of deprivation based on unemployment, non-car ownership, non-home ownership,
46
47 209 and household overcrowding,⁴⁵ was estimated from the preceding national census data,
48
49 210 with each participant assigned a score corresponding to the postcode of their home
50
51 211 dwelling and a negative value representing high socioeconomic status. We operationalised
52
53 212 the Townsend deprivation index as quintiles.

53 213 Information on smoking (never, previous and current smoker), previous doctor diagnosis of
54
55 214 any type of diabetes or a CVD event (yes, no) and use of medication (anti-hypertensive,
56
57 215 lipid-lowering or exogenous hormones or diabetes; yes, no) were collected. We created a
58
59 216 binary variable for family history (of father, mother and siblings) of CVD and related diseases
60

1
2
3 217 (yes, no). Body Mass Index (BMI) was calculated as weight/height² collected during the
4
5 218 Assessment Centre visit. We created a binary variable to indicate overweight or obese
6
7 219 according to standard World Health Organisation cut offs.⁴⁶ Physical activity was estimated
8
9 220 using Metabolic Equivalents (METs), the ratio of a person's working metabolic rate relative
10
11 221 to their resting metabolic rate. One MET was defined as the energy cost of sitting quietly
12
13 222 and is equivalent to a caloric consumption of 1kcal/kg/hour. We used standard cut-offs to
14
15 223 categorised participants as meeting physical activity guidelines of 150 min per week if their
16
17 224 METs were ≥ 600 MET-min/week.⁴⁷
18
19 225

20 226 **Statistical analysis**

21 227 Complete case analysis was used. We investigated missingness by comparing demographic
22
23 228 characteristics of the excluded sample with the analytic sample. Descriptive analyses
24
25 229 included mean (SD) for continuous variables and number (%) for categorical variables. We
26
27 230 created sample-based tertiles of diet quality for RFS, HDI and MDS for descriptive purposes
28
29 231 only. Unadjusted linear regression analyses were used to examine intakes of encouraged
30
31 232 food groups and total energy and nutrient intakes across tertiles of diet quality indices.
32
33 233 We used multivariable Cox proportional hazard regression models to estimate hazard ratios
34
35 234 (HR) and 95% Confidence Intervals (CI) of all-cause mortality, CVD mortality and risk of CVD
36
37 235 events (MI and stroke) according to each diet quality index separately (RFS, HDI and MDS).
38
39 236 We treated diet quality indices as continuous independent variables. CVD events and
40
41 237 mortality were treated as time-to-event outcome/dependent variables. We estimated the
42
43 238 duration of follow up as the time between the last day of dietary data and the first event of
44
45 239 either an MI, stroke, mortality, or the censoring date (4 March 2020). In participants who
46
47 240 had multiple events during the study period, the first event date was used. We adjusted the
48
49 241 Cox regression analyses for covariates identified using a directed acyclic graph
50
51 242 (Supplemental Figure 1). These included age (continuous), sex, deprivation (categorical),
52
53 243 smoking status (categorical), physical activity (continuous), medication use (binary), family
54
55 244 history of CVD (binary) and energy intake (continuous). The Cox proportional hazards
56
57 245 models also included PRS as an independent variable and were additionally adjusted for the
58
59 246 first 8 principal components of ancestry and genotyping batch.¹⁴ We included an interaction
60
247 term in the models to test for statistical interaction between each diet quality score and

1
2
3 248 PRS. Interactions were further inspected by conducting post-hoc estimation of the effects of
4
5 249 diet quality indices on events at 'low' and 'high' PRS score of -1 and +1 which, given PRS was
6
7 250 a standard score, represent minus and plus one SD of PRS. Data were analysed using Stata
8
9 251 (version 16.0; StataCorp., College Station, TX, USA). To address possible reverse causation,
10
11 252 sensitivity analyses excluded deaths and incident cases of MI and stroke within the first 2
12
13 253 years of follow up.

14 254

15 16 255 **Patient and public involvement**

17
18 256 The development of the research question or outcome measures was not informed by
19
20 257 patients' priorities, experience, or preferences. No patients were involved in the design and
21
22 258 conduct of the present study. There are no plans to disseminate the results to study
23
24 259 participants.

25 260

26 27 261 **RESULTS**

28
29 262 Of the 502,536 participants who were recruited into the UK Biobank, 425,529 participants
30
31 263 were excluded based on being not white British (n=92,907), having unusable genetic data
32
33 264 (n=1,459), being ineligible (n=23,215) or missing dietary or covariate data (n=307,951;
34
35 265 Supplemental Figure 2). Excluded participants were similar in age and sex to the included
36
37 266 sample, with somewhat higher BMI and rates of smoking and deprivation (Supplemental
38
39 267 Table 4). A total of 77,004 participants were included in the present analysis (Table 1). Mean
40
41 268 age at recruitment was 56.2 (SD 7.8) years and 55% were female. Forty-eight per cent of
42
43 269 participants had a colleague or university degree, most were experiencing low to mid
44
45 270 deprivation (67%), had never smoked (58%) and had a family history of CVD (74%; Table 1).
46
47 271 Fifty-nine per cent of the participants were overweight or obese and 85% met physical
48
49 272 activity guidelines.

50 273 Mean RFS, HDI and MDS were 6.78 (SD 2.40), 3.57 (SD 1.26) and 5.31 (SD 1.04), respectively.
51
52 274 Intake of fruits, vegetables, wholegrains and lean meat were higher with increasing tertile of
53
54 275 RFS, HDI and MDS (Table 2). Intake of low-fat dairy were higher with increasing tertile of RFS
55
56 276 and HDI and lower with increasing tertile of MDS. Intakes of total fat, saturated fat,
57
58 277 carbohydrates and sugars were higher with increasing tertile of RFS, HDI and MDS. Intakes
59
60 278 of protein were lower with increasing tertile of HDI and MDS, while intakes were higher with

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

For peer review only

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 \pm SD	56.2 \pm 7.8	57.0 \pm 7.8	55.6 \pm 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 \pm SD ¹	26.5 \pm 4.4	27.1 \pm 3.9	26.0 \pm 4.7
Waist circumference (cm), Mean \pm SD ¹	88.1 \pm 13.0	95.2 \pm 10.8	82.3 \pm 11.6
Total PA (MET min), Mean \pm SD	2477 \pm 2326	2542 \pm 2439	2423 \pm 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 \pm SD	8853 \pm 2172	9574 \pm 2253	8252 \pm 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.

1
2
3 281 During a mean follow-up of 7.8 years (a total of 600,193 person-years), we observed 1,141
4
5 282 new MI events and 748 new stroke events. During a mean follow-up of 7.9 years (a total of
6
7 283 604,431 person-years), we observed 364 deaths due to CVD and 2,409 all-cause deaths. Of
8
9 284 these, the majority of MI (72%) and stroke (60%) events, and CVD-related (72%) and all-
10 285 cause (59%) deaths were in males.
11
12

13 286

15 287 **RFS and risk of all-cause mortality, CVD-related mortality, MI and stroke**

17
18 288 The adjusted HR associated with a one-point higher RFS was 0.96 (95% CI: 0.94, 0.98) for
19
20 289 all-cause mortality, 0.94 (95% CI: 0.90, 0.98) for CVD-related mortality, 0.97 (95% CI: 0.95,
21
22 290 1.00) for MI and 0.94 (95% CI: 0.91, 0.98) for stroke (Table 3). When stratified by sex,
23
24 291 associations were comparable in men, while there was only evidence of an association
25
26 292 between RFS and all-cause mortality and stroke in females. The adjusted HR associated with
27
28 293 a one-point higher PRS was 1.33 (95% CI: 1.25, 1.41) for MI; when stratified by sex, there
29
30 294 was evidence of a stronger association in males. When an interaction term was added to the
31
32 295 models, there was no evidence (at the $p < 0.05$ level) of interaction between RFS and PRS for
33
34 296 any outcomes (p -interaction=0.40 [all-cause mortality], p -interaction=0.77 [CVD-related
35
36 297 mortality], p -interaction=0.17 [MI], and p -interaction=0.10 [stroke]). Effect sizes were
37
38 298 consistent when deaths and incident cases of MI and stroke within the first 2 years of follow
39
40 299 up were excluded (data not shown).
41

42 300

43 301 **HDI and risk of all-cause mortality, CVD-related mortality, MI and stroke**

45 302 The adjusted HR associated with a one-point higher HDI was 0.97 (95% CI: 0.93, 0.99) for all-
46
47 303 cause mortality. There was little evidence of associations between HDI and risk of CVD-
48
49 304 related mortality, MI, or stroke (Table 4). When stratified by sex, there was evidence of an
50
51 305 association between HDI and all-cause mortality in males only. The adjusted HR associated
52
53 306 with a one-point higher PRS was 1.33 (95% CI: 1.25, 1.41) for MI, which when stratified by
54
55 307 sex, there was evidence of a stronger association in males. When an interaction term was
56
57 308 added to the models, there was no evidence of interaction between HDI and PRS for other
58
59 309 outcomes (p -interaction=0.66 [all-cause mortality], p -interaction=0.86 [CVD-related
60
310 mortality] and p -interaction=0.17 [stroke]). There was some evidence of interaction

1
2
3 311 between HDI and PRS for MI events (p-interaction=0.049). While there was no evidence of
4
5 312 an effect of HDI on MI for participants with low (-1 SD) PRS (HR=1.02 [95% CI: 0.95, 1.10],
6
7 313 p=0.61), there was some evidence of an association between higher HDI and reduced risk of
8
9 314 MI events for those with high (+1 SD) PRS (HR=0.93 [95% CI: 0.88, 0.99], p=0.017). Effect
10
11 315 sizes were consistent when deaths and incident cases of MI and stroke within the first 2
12
13 316 years of follow up were excluded (data not shown).

14
15 317

17 318 **MDS and risk of all-cause mortality, CVD-related mortality, MI and stroke**

19
20 319 The adjusted HR associated with a one-point higher MDS was 0.95 (95% CI: 0.91, 0.98) for
21
22 320 all-cause mortality and 0.93 (95% CI: 0.87, 1.00) for stroke (Table 5). There was limited
23
24 321 evidence of associations between MDS and risk of CVD-related mortality and MI. When
25
26 322 stratified by sex, there was evidence of an association between MDS and all-cause mortality
27
28 323 and MI in males only. The adjusted HR associated with a one-point higher PRS was 1.33
29
30 324 (95% CI: 1.25, 1.41) for MI; when stratified by sex, there was evidence of a stronger
31
32 325 association in males. When an interaction term was added to the models, there was no
33
34 326 evidence of interaction between MDS and PRS for other outcomes (p-interaction=0.58 [all-
35
36 327 cause mortality], p-interaction=0.72 [CVD-related mortality] and p-interaction=0.12
37
38 328 [stroke]). There was evidence of interaction between MDS and PRS for MI events (p-
39
40 329 interaction=0.026). While there was no evidence of an effect of MDS on MI for those with
41
42 330 low (-1 SD) PRS (HR=1.03 [95% CI: 0.94, 1.12], p=0.56) there was strong evidence of an
43
44 331 association between higher MDS and reduced risk of MI events for those with high (+1 SD)
45
46 332 PRS (HR=0.91 [95% CI: 0.85, 0.97], p=0.004). Effect sizes were consistent when deaths and
47
48 333 incident cases of MI and stroke within the first 2 years of follow up were excluded (data not
49
50 334 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.

335 DISCUSSION

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

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

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

1
2
3 366 include plant-based diet quality components when assessing diet-disease associations.^{54 56 57}

4
5 367 However, given the predominately white and highly-educated participants in the UK

6
7 368 Biobank, further research in diverse populations is needed to investigate the applicability of

8
9 369 these diet quality methodologies for examining CVD risk independent of genetic risk.⁷

10
11 370 Our stronger associations between diet quality and genetic CVD risk in males confirm

12
13 371 previous research.^{14 58} Although this may be partly explained by the high prevalence of

14
15 372 diabetes and unhealthy behaviours in men,⁵⁹ it may also be due the lower number of events

16
17 373 and deaths in women compared with men in the present study. Nonetheless, it is likely that

18
19 374 the biological and behavioural pathways in which risk factors exert their effects on CVD risk

20
21 375 are different between men and women.⁵⁸

22 23 376 **Strengths and limitations**

24
25 377 Our main strength was the large sample size and inclusion of genetic data. This enabled

26
27 378 investigation of a genetic risk score created 300 SNPs known to be associated with CVD,

28
29 379 more than any previous publications in the UK Biobank.^{14 55} While the PRS used was specific

30
31 380 to coronary disease, it has been used to identify predispositions to a wide variety of CVD

32
33 381 and non-CVDs, as well as premature mortality, given these may develop in parallel with

34
35 382 coronary disease for the same genetic origins. The dietary questionnaire has been

36
37 383 previously validated and included sufficient detail to allow us to create three contrasting

38
39 384 diet quality indices. There are a number of limitations that should be acknowledged. While

40
41 385 the dietary assessment method is a short-term measure of intake, our use of up to four

42
43 386 instances of dietary assessments provided an estimate of longer-term intake. Although the

44
45 387 present analysis is likely to be subject to self-selection bias associated with the number of

46
47 388 participants who completed the dietary assessment and the low response rate, associations

48
49 389 between demographic and behavioural risk factors and mortality in the UK Biobank have

50
51 390 been shown to be comparable to those from national health survey data from England and

52
53 391 Scotland.⁶⁰ Whilst we adjusted analyses based on a range of confounders identified using a

54
55 392 directed acyclic graph, we cannot discount the possibility of residual or unmeasured

56
57 393 confounding.

58 59 394 **Conclusion**

60

1
2
3 395 This prospective population-based cohort study provided evidence that higher diet quality
4 (RFS, HDI and MDS) was associated with lower risk of all-cause mortality, regardless of
5 396 (RFS, HDI and MDS) was associated with lower risk of all-cause mortality, regardless of
6 genetic CVD risk. Diet quality, when estimated using the RFS only, was associated with lower
7 397 genetic CVD risk. Diet quality, when estimated using the RFS only, was associated with lower
8 risk of CVD-related mortality and MI, independent of genetic CVD risk. The diet quality
9 398 risk of CVD-related mortality and MI, independent of genetic CVD risk. The diet quality
10 indices investigated in this study have common food-based scoring components, providing
11 399 indices investigated in this study have common food-based scoring components, providing
12 further evidence for the best practice design and implementation of food-based diet quality
13 400 further evidence for the best practice design and implementation of food-based diet quality
14 indices for assessing health outcomes. Further research in diverse populations is needed to
15 401 indices for assessing health outcomes. Further research in diverse populations is needed to
16 investigate the applicability of different diet quality methodologies for examining CVD risk
17 402 investigate the applicability of different diet quality methodologies for examining CVD risk
18 independent of genetic susceptibility.
19 403 independent of genetic susceptibility.

404

22
23 405 **Authors' contributions:** KML, GA, SB, JW, CM and SAM designed the analysis. KML, SB, JW
24 and GA conducted the statistical analysis. KML drafted the manuscript. All authors
25 406 and GA conducted the statistical analysis. KML drafted the manuscript. All authors
26 contributed to the critical review of the manuscript and approved the final version of the
27 407 contributed to the critical review of the manuscript and approved the final version of the
28 408 manuscript.

30
31 409 **Funding:** KML is supported by a National Health and Medical Research Council Emerging
32 Leadership Fellowship (APP1173803). JW is funded by the Lister Prize Fellowship (173096).
33 410 Leadership Fellowship (APP1173803). JW is funded by the Lister Prize Fellowship (173096).
34 The funding source had no role in the design or conduct of the study; collection,
35 411 The funding source had no role in the design or conduct of the study; collection,
36 management, analysis, and interpretation of the data; or preparation, review, or approval of
37 412 management, analysis, and interpretation of the data; or preparation, review, or approval of
38 the manuscript.
39 413 the manuscript.

40 414 **Competing interests:** All authors have completed the ICMJE uniform disclosure form at
41 www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the
42 415 www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the
43 submitted work; KML is a consultant for HeadUp Labs; no other financial relationships with
44 416 submitted work; KML is a consultant for HeadUp Labs; no other financial relationships with
45 any organisations that might have an interest in the submitted work in the previous three
46 417 any organisations that might have an interest in the submitted work in the previous three
47 418 years; no other relationships or activities that could appear to have influenced the
48 submitted work.
49 419 submitted work.

50
51 420 **Ethical approval:** UK Biobank received ethical approval from the research ethics committee
52 (reference 13/NW/0382). All participants provided informed consent to participate. An
53 421 (reference 13/NW/0382). All participants provided informed consent to participate. An
54 ethics exemption was granted by Deakin University Human Research Ethics Committee
55 422 ethics exemption was granted by Deakin University Human Research Ethics Committee
56 (Reference number 2019_293).
57 423 (Reference number 2019_293).
58
59
60

1
2
3 424 **Data availability statement:** The genetic and phenotypic UK Biobank data are available on
4 application to the UK Biobank. This research has been conducted using the UK Biobank
5 425
6 Resource under Application 34894.
7 426

8
9 427 **Transparency:** The lead author (KML) affirms that the manuscript is an honest, accurate,
10 and transparent account of the study being reported; that no important aspects of the study
11 428
12 have been omitted and that discrepancies from the study as planned have been explained
13 429
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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](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: 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: 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 Office 2005.
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

- 1
- 2
- 3
- 4 36. Trichopoulou A, Orfanos P, Norat T, et al. Modified Mediterranean diet and survival: EPIC-elderly
- 5 prospective cohort study. 2005;330(7498):991. doi: 10.1136/bmj.38415.644155.8F %J BMJ
- 6 37. Trichopoulou A, Costacou T, Bamia C, et al. Adherence to a Mediterranean Diet and Survival in a
- 7 Greek Population. 2003;348(26):2599-608. doi: 10.1056/NEJMoa025039
- 8 38. Bach-Faig A, Berry EM, Lairon D, et al. Mediterranean diet pyramid today. Science and cultural
- 9 updates. *Public Health Nutr* 2011;14(12A):2274-84. doi: 10.1017/S1368980011002515
- 10 [published Online First: 2011/12/13]
- 11 39. Magliano D, Liew D, Pater H, et al. Accuracy of the Australian National Death Index: comparison
- 12 with adjudicated fatal outcomes among Australian participants in the Long-term
- 13 Intervention with Pravastatin in Ischaemic Disease (LIPID) study. 2003;27(6):649-53. doi:
- 14 10.1111/j.1467-842X.2003.tb00615.x
- 15 40. Schnier C, Bush K, Nolan J, et al. Definitions of Stroke for UK Biobank Phase 1 Outcomes
- 16 Adjudication Documentation prepared by: On behalf of UK Biobank Outcome Adjudication
- 17 Group. 2017. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_mi.pdf
- 18 (accessed 16 September 2020).
- 19 41. Schnier C, Bush K, Nolan J, et al. Definitions of Acute Myocardial Infarction and Main Myocardial
- 20 Infarction Pathological Types UK Biobank Phase 1 Outcomes Adjudication Documentation
- 21 prepared by: On behalf of UK Biobank Outcome Adjudication Group. 2017.
- 22 http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_mi.pdf (accessed 16
- 23 September 2020).
- 24 42. National Health Service. COVID-19 Daily Deaths. 2020.
- 25 <https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-daily-deaths/>
- 26 (accessed 16 September 2020).
- 27 43. Ntalla I, Kanoni S, Zeng L, et al. Genetic Risk Score for Coronary Disease Identifies Predispositions
- 28 to Cardiovascular and Noncardiovascular Diseases. *J Am Coll Cardiol* 2019;73(23):2932-42.
- 29 doi: <https://doi.org/10.1016/j.jacc.2019.03.512>
- 30 44. Dudbridge F. Polygenic Epidemiology. 2016;40(4):268-72. doi: 10.1002/gepi.21966
- 31 45. Townsend P, Phillimore P, Beattie A. Health and Deprivation. *Nurs Stand* 1988;2(17):34. doi:
- 32 10.7748/ns.2.17.34.s66 [published Online First: 1988/01/30]
- 33 46. World Health Organization. BMI classification 2015 [updated 21 July 2015. Available from:
- 34 http://apps.who.int/bmi/index.jsp?introPage=intro_3.html accessed 3rd December 2015.
- 35 47. The IPAQ Group. IPAQ scoring protocol - International Physical Activity Questionnaire. Available
- 36 at: <https://sites.google.com/site/theipaq/scoring-protocol>. Accessed 22 May 2020. [
- 37 48. Neelakantan N, Koh W-P, Yuan J-M, et al. Diet-Quality Indexes Are Associated with a Lower Risk
- 38 of Cardiovascular, Respiratory, and All-Cause Mortality among Chinese Adults. *The Journal of*
- 39 *Nutrition* 2018;148(8):1323-32. doi: 10.1093/jn/nxy094 %J The Journal of Nutrition
- 40 49. Stefler D, Maljutina S, Kubinova R, et al. Mediterranean diet score and total and cardiovascular
- 41 mortality in Eastern Europe: the HAPIEE study. *Eur J Nutr* 2017;56(1):421-29. doi:
- 42 10.1007/s00394-015-1092-x
- 43 50. Lassale C, Gunter MJ, Romaguera D, et al. Diet Quality Scores and Prediction of All-Cause,
- 44 Cardiovascular and Cancer Mortality in a Pan-European Cohort Study. *PLoS One*
- 45 2016;11(7):e0159025. doi: 10.1371/journal.pone.0159025
- 46 51. Kaluza J, Håkansson N, Brzozowska A, et al. Diet quality and mortality: a population-based
- 47 prospective study of men. *Eur J Clin Nutr* 2009;63(4):451-57. doi: 10.1038/sj.ejcn.1602968
- 48 52. Schulze MB, Martínez-González MA, Fung TT, et al. Food based dietary patterns and chronic
- 49 disease prevention. *BMJ* 2018;361 doi: 10.1136/bmj.k2396
- 50 53. Lakkur S, Judd SE. Diet and Stroke. 2015;46(7):2007-11. doi:
- 51 doi:10.1161/STROKEAHA.114.006306
- 52 54. Chiavaroli L, Vigliouk E, Nishi SK, et al. DASH Dietary Pattern and Cardiometabolic Outcomes:
- 53 An Umbrella Review of Systematic Reviews and Meta-Analyses. *Nutrients* 2019;11(2) doi:
- 54 10.3390/nu11020338 [published Online First: 2019/02/16]
- 55
- 56
- 57
- 58
- 59
- 60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
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 Indicator	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 Indicator	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)	>90 g/day=0 ≤90 g/day=1
11. Calcium	Sausage, bacon, ham, liver 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 Indicator	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 indicators 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	
4. Cereals	Unsalted peanuts, unsalted nuts, types of spreads/sauces consumed (Peanut butter), seeds 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	
5. Fish and seafood	White pasta, wholemeal pasta, white rice, brown rice, couscous, other grain Homemade soup, ingredients in homemade soup (pasta) 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 indicators 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	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

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

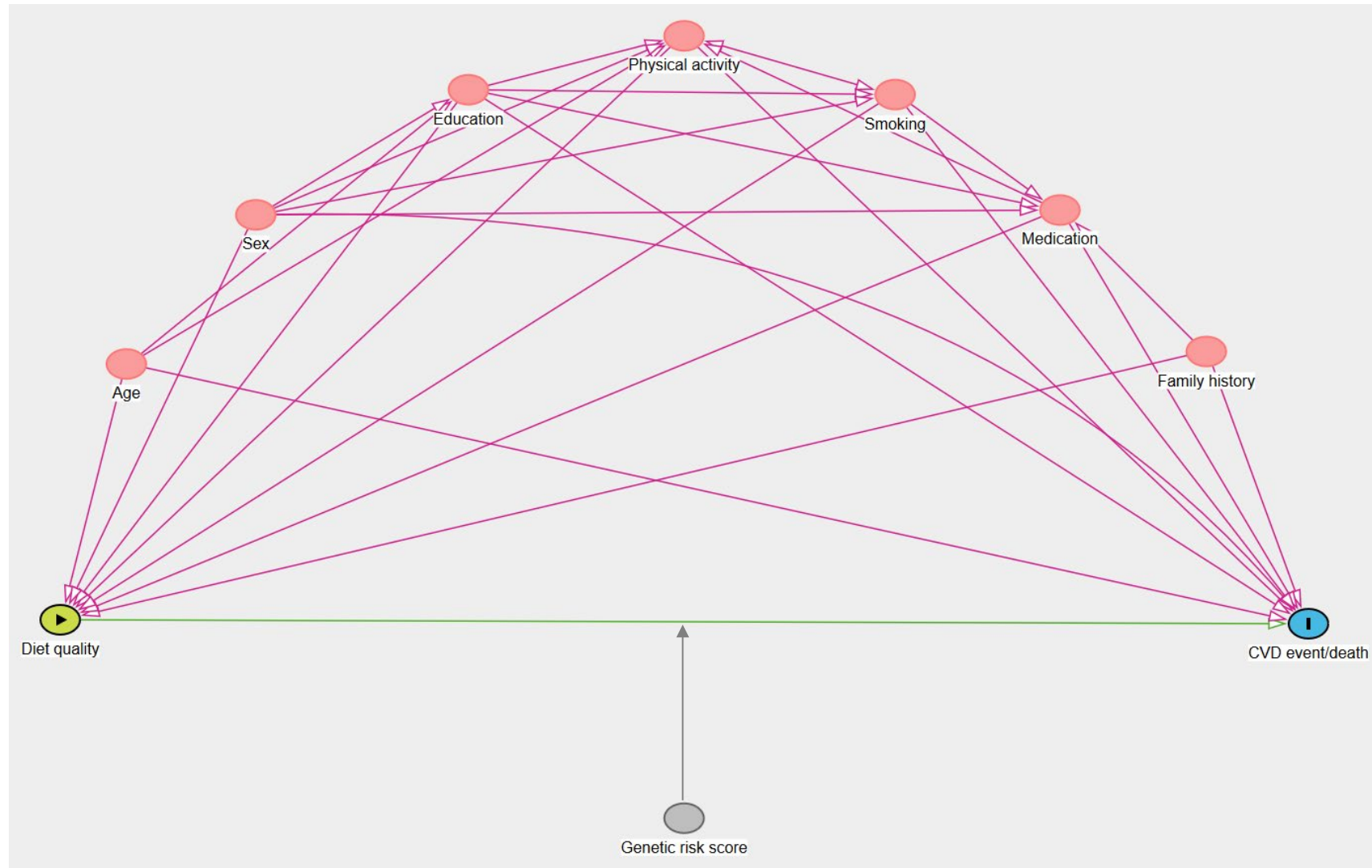
For peer review only

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

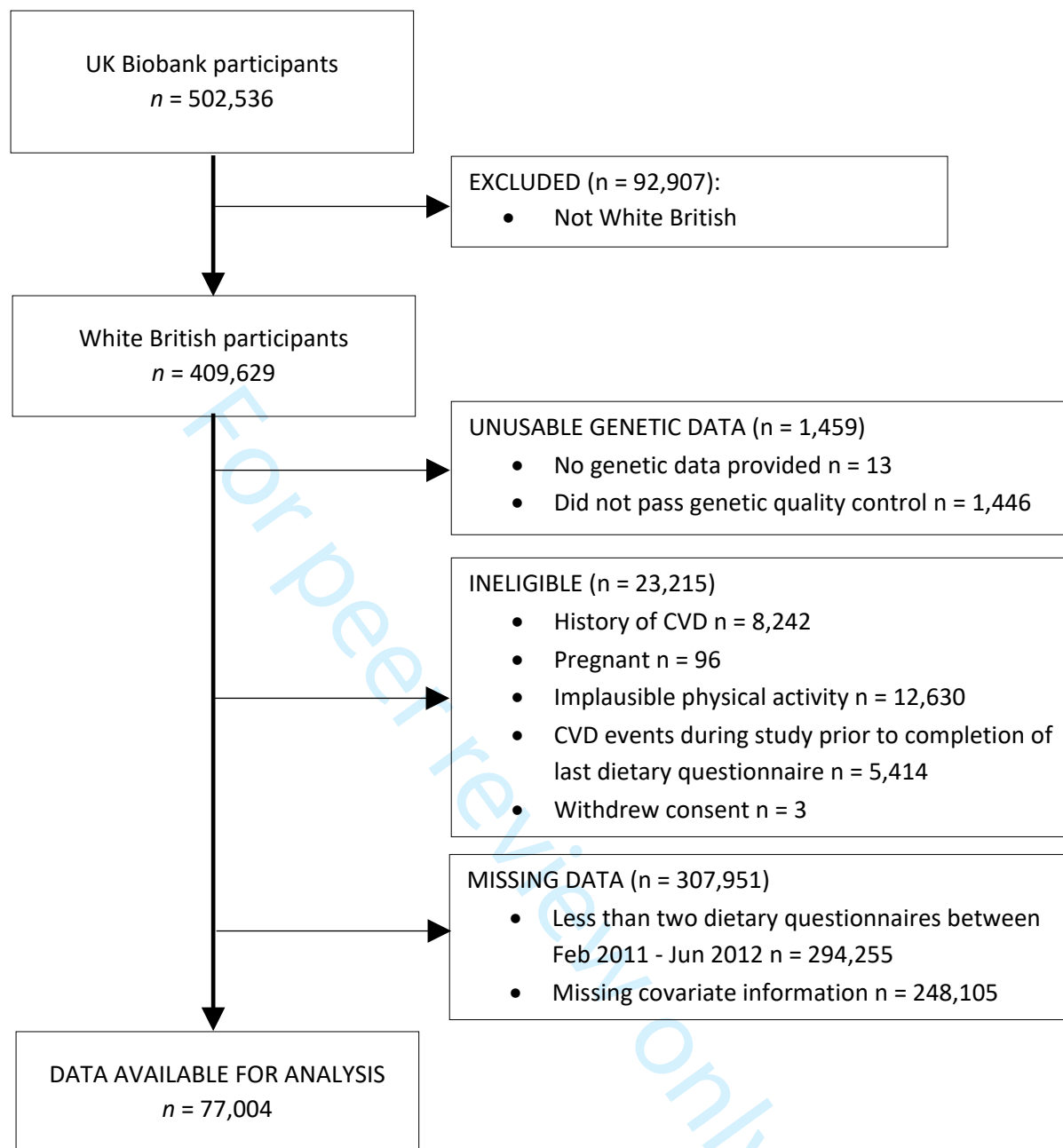
Characteristic	Excluded N (%)	Analytic N (%)
N ¹	425,532	77,004
Sex female	231,382 (54.4)	42,020 (54.6)
Age at recruitment (years), Mean \pm SD	56.6 \pm 8.2	56.2 \pm 7.8
Townsend Deprivation Index		
Least deprived	82,535 (19.4)	18,129 (23.5)
2nd least deprived	82,878 (19.5)	17,227 (22.4)
Medium deprivation	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 \pm SD	27.6 \pm 4.9	26.5 \pm 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			

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

1
2
3 conjunction with this article (freely available on the Web sites of PLoS Medicine at
4 <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and
5 Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at
6 <http://www.strobe-statement.org>.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

BMJ Open

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

Journal:	<i>BMJ Open</i>
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

SCHOLARONE™
Manuscripts



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 Abbott¹, Steven J. Bowe², Joey Ward³, Catherine Milte¹, Sarah A. McNaughton¹

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

1
2
3 1 **ABSTRACT**
4

5
6 2 **Objectives:** To examine associations of three diet quality indices and a polygenic risk score
7
8 3 with incidence of all-cause mortality, cardiovascular disease (CVD) mortality, myocardial
9
10 4 infarction (MI) and stroke.

11
12 5 **Design:** Prospective cohort study.

13
14 6 **Setting:** UK Biobank, UK.

15
16
17 7 **Participants:** 77,004 men and women (40-70 years) recruited between 2006 and 2010.

18
19 8 **Main outcome measures:** A polygenic risk score was created from 300 single nucleotide
20
21 9 polymorphisms (SNP) associated with CVD. Cox proportional hazard ratios (HR) were used to
22
23 10 estimate independent effects of diet quality and genetic risk on all-cause mortality, CVD
24
25 11 mortality, MI and stroke risk. Dietary intake (Oxford WebQ) was used to calculate
26
27 12 Recommended Food Score (RFS), Healthy Diet Indicator (HDI) and Mediterranean Diet Score
28
29 13 (MDS).

30
31 14 **Results:** New all-cause (n=2,409) and CVD (n=364) deaths, and MI (n=1,141) and stroke
32
33 15 (n=748) events were identified during mean follow-ups of 7.9 and 7.8 years, respectively.
34
35 16 The adjusted HR associated with one-point higher RFS for all-cause mortality was 0.96 (95%
36
37 17 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
38
39 18 1.00) and stroke was 0.94 (95% CI: 0.91 to 0.98). The adjusted HR for all-cause mortality
40
41 19 associated with one-point higher HDI and MDS was 0.97 (95% CI: 0.93 to 0.99) and 0.95
42
43 20 (95% CI: 0.91 to 0.98), respectively. The adjusted HR associated with one-point higher MDS
44
45 21 for stroke was 0.93 (95% CI: 0.87 to 1.00). There was little evidence of associations between
46
47 22 HDI and risk of CVD mortality, MI or stroke. There was evidence of an interaction between
48
49 23 diet quality and genetic risk score for MI.

50
51 24 **Conclusion:** Higher diet quality predicted lower risk of all-cause mortality, independent of
52
53 25 genetic risk. Higher RFS was also associated with lower risk of CVD mortality and MI. These
54
55 26 findings demonstrate the benefit of following a healthy diet, regardless of genetic risk.

56 27
57
58
59
60

28

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.

29 INTRODUCTION

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

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

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

59 METHODS

60 **Study design and participants**

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

81 **Study measures**

82 Dietary intake

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

1
2
3 92 collected via four online cycles between February 2011 to June 2012: February 2011 to April
4 93 2011 (online cycle 1); June 2011 to September 2011 (online cycle 2); October 2011 to
5 94 December 2011 (online cycle 3); April 2012 to June 2012 (online cycle 4). The total period of
6
7 95 available dietary data from the Oxford WebQ was 38 months (Apr 2009 - Jun 2012). Email
8
9 96 invitations were sent on different days of the week to capture variation in dietary intakes
10
11 97 and participants were given 3 days to complete the questionnaire for cycles 1 and 2 and 14
12
13 98 days for cycles 3 and 4.
14
15

16
17 99 To establish a baseline dietary intake in the present analysis, we calculated a mean dietary
18
19 100 intake based on the four online Oxford WebQ cycles only. This was because the time
20
21 101 between the 1st and 4th online cycle measurements was 16 months (Feb 2011 - Jun 2012)
22
23 102 and was considered a more credible timeframe for an average baseline than the 38 months
24
25 103 available from all five Oxford WebQ measurements. This resulted in a minimal sample loss
26
27 104 (<10%) while providing a shorter dietary exposure period and a more consistent approach to
28
29 105 the use of the dietary data by using only the online cycles of the OxfordWebQ. To better
30
31 106 capture usual intake, we calculated average nutrient intakes, food group intakes and diet
32
33 107 quality scores for participants who had two or more valid measurements for the four online
34
35 108 cycles of the Oxford WebQ.
36

37 109 38 110 Diet quality

39 111 Information on food and beverage intakes from the Oxford WebQ were used to calculate
40
41 112 three diet quality indices: the Recommended Food Score (RFS), which is based on intakes of
42
43 113 encouraged foods only,²⁶ and the Healthy Diet Indicator (HDI), which scores intakes of a
44
45 114 combination of foods and nutrients from dietary guidelines,²⁷ and the Mediterranean Diet
46
47 115 Score (MDS), representing dietary patterns identified as healthful.²⁸ These indices were
48
49 116 selected as they represent three contrasting diet quality methodologies that have been
50
51 117 applied internationally to assess diet-disease associations.^{9 10 26 27 29 30}
52

53 118 The RFS is a food-based variety index designed to assess consumption of food groups
54
55 119 encouraged in the dietary guidelines.¹⁰ As detailed in Supplemental Table 1, food intakes
56
57 120 were scored according to five food groups: fruits (7 items), vegetables (7 items),
58
59 121 wholegrains (2 items), lean meat and alternatives (3 items) and reduced fat dairy products
60
122 (2 items). Scoring was based on the RFS designed by Kant and Graubaud,²⁶ and has been

1
2
3 123 used elsewhere.^{31 32} We summed intakes of food items within each group to create a total
4
5 124 intake for each food group. Food groups were then assigned a score of 1 if they were
6
7 125 consumed above the minimum amount threshold: 15 g/d for non-beverages and 30 g/d for
8
9 126 beverages. Intakes below these thresholds were scored 0. Scores ranged between 0 and 21,
10
11 127 with higher scores indicating a higher quality diet and a wider consumption of
12
13 128 recommended foods.³³

15 129 The HDI is a food- and nutrient-based index designed to reflect consumption of foods
16
17 130 recommended for a healthy diet by the World Health Organisation.³⁴ The original HDI was
18
19 131 developed and validated in 1997 based on the 1990 World Health Organisation's dietary
20
21 132 recommendations for the prevention of chronic disease.³⁵ We adapted a 12-point Healthy
22
23 133 Diet Score designed by Maynard et al.²⁷ to reflect adherence to the 2020 World Health
24
25 134 Organisation healthy diet fact sheet.³⁴ As cholesterol intake is not part of the 2020
26
27 135 recommendations and information on its intake was not available in the UK Biobank, we
28
29 136 used an 11-item score that included the following groups: saturated fat; poly-unsaturated
30
31 137 fat; protein; total carbohydrates; dietary fibre; fruits and vegetables; pulses and nuts; total
32
33 138 non-milk extrinsic sugars; fish; red meat and meat products; and calcium. Data on intake of
34
35 139 non-milk extrinsic sugars was not available in the UK Biobank and so we adapted the HDI to
36
37 140 score intakes of total sugars instead. Criteria for scoring was based on cut points detailed in
38
39 141 Supplemental Table 2. We assigned intakes within the cut offs a score of 1 and those outside
40
41 142 of the cut offs were assigned a score of 0. The total score ranged from 0 to 11, with a higher
42
43 143 score reflecting a higher diet quality (Supplemental Table 2).

44 144 The MDS is a food- and nutrient-based score designed to reflect adherence to a
45
46 145 Mediterranean style diet. The present study used the 9-item index developed and validated
47
48 146 by Trichopoulou et al. as it is the first and most widely used version of the MDS.^{36 37} Food
49
50 147 and nutrient intakes were scored according to nine components: vegetables, legumes, fruits
51
52 148 and nuts, cereals, fish and seafood, monounsaturated fats to saturated fats ratio, dairy
53
54 149 products, meat and meat products and alcohol (Supplemental Table 3). As used by
55
56 150 Trichopoulou et al.,³⁶ we used sex-specific median intakes as cut off points for intakes of
57
58 151 each component. A score of 1 was assigned to participants whose intake of vegetables,
59
60 152 legumes, fruits and nuts, cereals, fish and seafood and monounsaturated: saturated fats was
153
154 153 above the median. A score of 1 was assigned to intake of dairy products, meat and meat

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

159 Cardiovascular events and mortality

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

176 Polygenic risk score

177 We used the March 2018 release of the imputed genetic data from UK Biobank
178 (downloaded 11 November 2019). From the resulting dataset, we excluded those who self-
179 reported ancestry other than white British, those who were missing more than 10% of the
180 genetic data and those who were defined by UK Biobank as being heterozygosity outliers.
181 Additionally, for every pair of who were individuals who were second cousins or closer (i.e.
182 those with a kinship coefficient > 0.042) one was excluded at random. We used information
183 on 300 single nucleotide polymorphisms (SNPs) known to be associated with coronary
184 artery disease⁴³ to create a PRS for CVD for each individual.⁴⁴ Evidence indicates that a

1
2
3 185 genetic risk score estimated from these 300 SNPs is associated with traditional risk factors
4
5 186 for CVD, such as type 2 diabetes and hypertension, contributes to the development of CVD
6
7 187 conditions that have their origins in atherosclerosis, such as peripheral arterial disease and
8
9 188 stroke, and is associated with premature mortality.⁴³ The PRS was estimated using PLINK, an
10
11 189 open-source tool for genomic research,⁴⁵ by generating the sum of the number of risk
12
13 190 alleles present at each locus and weighting by the log of the odds for that locus¹⁸ estimated
14
15 191 from the list of 300 SNPs using the PLINK “-score” command – with no-mean-imputation
16
17 192 flag. PRS were available for all participants included in the final study sample, where PRS
18
19 193 were transformed to standardised Z scores and were treated as a continuous variable in all
20
21 194 modelling.

22 195

23 196 Demographic and health information

24
25 197 Information on demographics, medical history and health behaviours were collected using
26
27 198 interview-administered questionnaires at recruitment and follow ups. Participant age at
28
29 199 recruitment and sex were self-reported. No adjustments were made for discrepancies
30
31 200 between self-reported sex and genetic sex. Education was assessed by asking “Which of the
32
33 201 following qualifications do you have? (You can select more than one),” with the options
34
35 202 college or university degree, A levels or equivalent, O levels or GCSEs or equivalent, CSEs,
36
37 203 NVQ/HND/HNC, other professional qualifications (e.g., nursing or teaching). We
38
39 204 operationalised this into 5 categories based on the highest level of education: i) college or
40
41 205 university degree, ii) all professional qualification (NVQ/HND/HNC, other professional
42
43 206 qualifications), iii) A levels or equivalent, iv) O levels, GCSEs or equivalent or CSEs and v)
44
45 207 none of the above or prefer not to answer. The Townsend deprivation index, a composite
46
47 208 measure of deprivation based on unemployment, non-car ownership, non-home ownership,
48
49 209 and household overcrowding,⁴⁶ was estimated from the preceding national census data,
50
51 210 with each participant assigned a score corresponding to the postcode of their home
52
53 211 dwelling and a negative value representing high socioeconomic status. We operationalised
54
55 212 the Townsend deprivation index as quintiles.

56
57 213 Information on smoking (never, previous and current smoker), previous doctor diagnosis of
58
59 214 any type of diabetes or a CVD event (yes, no) and use of medication (anti-hypertensive,
60
215 lipid-lowering or exogenous hormones or diabetes; yes, no) were collected. We created a

1
2
3 216 binary variable for family history (of father, mother and siblings) of CVD and related diseases
4
5 217 (yes, no). Body Mass Index (BMI) was calculated as weight/height² collected during the
6
7 218 Assessment Centre visit. We created a binary variable to indicate overweight or obese
8
9 219 according to standard World Health Organisation cut offs.⁴⁷ Physical activity was estimated
10
11 220 using Metabolic Equivalents (METs), the ratio of a person's working metabolic rate relative
12
13 221 to their resting metabolic rate. One MET was defined as the energy cost of sitting quietly
14
15 222 and is equivalent to a caloric consumption of 1kcal/kg/hour. We used standard cut-offs to
16
17 223 categorised participants as meeting physical activity guidelines of 150 min per week if their
18
19 224 METs were ≥ 600 MET-min/week.⁴⁸
20
21 225

22 226 **Statistical analysis**

23 227 Complete case analysis was used. We investigated missingness by comparing demographic
24
25 228 characteristics of the excluded sample with the analytic sample. Descriptive analyses
26
27 229 included mean (SD) for continuous variables and number (%) for categorical variables. We
28
29 230 created sample-based tertiles of diet quality for RFS, HDI and MDS for descriptive purposes
30
31 231 only. Unadjusted linear regression analyses were used to examine intakes of encouraged
32
33 232 food groups and total energy and nutrient intakes across tertiles of diet quality indices. This
34
35 233 descriptive analysis aimed to show that diet quality scores reflect differences in underlying
36
37 234 food and nutrient intakes, thus assisting with interpretation and translation into actual food
38
39 235 intakes.

40 236 We used multivariable Cox proportional hazard regression models to estimate hazard ratios
41
42 237 (HR) and 95% Confidence Intervals (CI) of all-cause mortality, CVD mortality and risk of CVD
43
44 238 events (MI and stroke) according to each diet quality index separately (RFS, HDI and MDS).
45
46 239 We treated diet quality indices as continuous independent variables. CVD events and
47
48 240 mortality were treated as time-to-event outcome/dependent variables. We estimated the
49
50 241 duration of follow up as the time between the last day of dietary data and the first event of
51
52 242 either an MI, stroke, mortality, or the censoring date (4 March 2020). In participants who
53
54 243 had multiple events during the study period, the first event date was used. We adjusted the
55
56 244 Cox regression analyses for covariates identified using a directed acyclic graph
57
58 245 (Supplemental Figure 1). These included age (continuous), sex, deprivation (categorical),
59
60 246 smoking status (categorical), physical activity (continuous), medication use (binary), family

1
2
3 247 history of CVD (binary) and energy intake (continuous). The role of sex by diet quality and by
4
5 248 PRS interactions were further tested by adding an interaction term to each model.
6
7 249 Consistent with recommendations for sex differences in cardiovascular associations,⁴⁹
8
9 250 analyses were presented stratified by sex of whether there were any apparent sex
10
11 251 differences. The Cox proportional hazards models also included PRS as an independent
12
13 252 variable and were additionally adjusted for the first 8 principal components of ancestry and
14
15 253 genotyping batch.¹⁴ We included an interaction term in the models to test for statistical
16
17 254 interaction between each diet quality score and PRS. Interactions were further inspected by
18
19 255 conducting post-hoc estimation of the effects of diet quality indices on events at 'low' and
20
21 256 'high' PRS score of -1 and +1 which, given PRS was a standard score, represent minus and
22
23 257 plus one SD of PRS. Data were analysed using Stata (version 16.0; StataCorp., College
24
25 258 Station, TX, USA). To address possible reverse causation, sensitivity analyses excluded
26
27 259 deaths and incident cases of MI and stroke within the first 2 years of follow up.
28
29 260

261 **Patient and public involvement**

262 The development of the research question or outcome measures was not informed by
263 patients' priorities, experience, or preferences. No patients were involved in the design and
264 conduct of the present study. There are no plans to disseminate the results to study
265 participants.
266

267 **RESULTS**

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

1
2
3 277 Fifty-nine per cent of the participants were overweight or obese and 85% met physical
4
5 278 activity guidelines.
6

7 279 Mean RFS, HDI and MDS were 6.78 (SD 2.40), 3.57 (SD 1.26) and 5.31 (SD 1.04), respectively.
8
9 280 Intake of fruits, vegetables, wholegrains and lean meat were higher with increasing tertile of
10
11 281 RFS, HDI and MDS (Table 2). Intake of low-fat dairy were higher with increasing tertile of RFS
12
13 282 and HDI and lower with increasing tertile of MDS. Intakes of total fat, saturated fat,
14
15 283 carbohydrates and sugars were higher with increasing tertile of RFS, HDI and MDS. Intakes
16
17 284 of protein were lower with increasing tertile of HDI and MDS, while intakes were higher with
18
19 285 increasing tertile of RFS. Intakes of PUFA were lower with increasing tertile of RFS, while
20
21 286 intakes were higher with increasing tertile of HDI and MDS (Table 2).
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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 \pm SD	56.2 \pm 7.8	57.0 \pm 7.8	55.6 \pm 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 \pm SD ¹	26.5 \pm 4.4	27.1 \pm 3.9	26.0 \pm 4.7
Waist circumference (cm), Mean \pm SD ¹	88.1 \pm 13.0	95.2 \pm 10.8	82.3 \pm 11.6
Total PA (MET min), Mean \pm SD	2477 \pm 2326	2542 \pm 2439	2423 \pm 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 \pm SD	8853 \pm 2172	9574 \pm 2253	8252 \pm 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.

1
2
3 287 During a mean follow-up of 7.8 years (a total of 600,193 person-years), we observed 1,141
4
5 288 new MI events and 748 new stroke events. During a mean follow-up of 7.9 years (a total of
6
7 289 604,431 person-years), we observed 364 deaths due to CVD and 2,409 all-cause deaths. Of
8
9 290 these, the majority of MI (72%) and stroke (60%) events, and CVD (72%) and all-cause (59%)
10 291 deaths were in males.
11
12

13 292

15 293 **RFS and risk of all-cause mortality, CVD mortality, MI and stroke**

17
18 294 The adjusted HR associated with a one-point higher RFS for all-cause mortality was 0.96
19
20 295 (95% CI: 0.94, 0.98), for CVD mortality was 0.94 (95% CI: 0.90, 0.98), for MI was 0.97 (95%
21
22 296 CI: 0.95, 1.00) and for stroke was 0.94 (95% CI: 0.91, 0.98) (Table 3). There was limited
23
24 297 evidence (all p-values > 0.1) of sex by diet interactions. When stratified by sex, associations
25
26 298 were comparable in men, while there was only evidence of an association between RFS and
27
28 299 all-cause mortality and stroke in females. The adjusted HR associated with a one-point
29
30 300 higher PRS for MI was 1.33 (95% CI: 1.25, 1.41); when stratified by sex, there was evidence
31
32 301 of a stronger association in males. When an interaction term for PRS was added to the
33
34 302 models, there was no evidence (at the p<0.05 level) of interaction between RFS and PRS for
35
36 303 any outcomes (p-interaction=0.40 [all-cause mortality], p-interaction=0.77 [CVD mortality],
37
38 304 p-interaction=0.17 [MI], and p-interaction=0.10 [stroke]). The interaction of sex by PRS
39
40 305 showed evidence that the effect of higher PRS on higher risk of MI was more pronounced
41
42 306 for males (HR=1.21 [95% CI: 1.06, 1.37], p=.004). Effect sizes were consistent when deaths
43
44 307 and incident cases of MI and stroke within the first 2 years of follow up were excluded (data
45
46 308 not shown).
47

309

48 310 **HDI and risk of all-cause mortality, CVD mortality, MI and stroke**

49
50 311 The adjusted HR associated with a one-point higher HDI for all-cause mortality was 0.97
51
52 312 (95% CI: 0.93, 0.99). There was little evidence of associations between HDI and risk of CVD
53
54 313 mortality, MI, or stroke (Table 4). There was limited evidence (all p-values > 0.1) of sex by
55
56 314 diet interactions. When stratified by sex, there was evidence of an association between HDI
57
58 315 and all-cause mortality in males only. The adjusted HR associated with a one-point higher
59
60 316 PRS for MI was 1.33 (95% CI: 1.25, 1.41), which when stratified by sex, there was evidence

1
2
3 317 of a stronger association in males. When an interaction term for PRS was added to the
4
5 318 models, there was no evidence of interaction between HDI and PRS for other outcomes (p-
6
7 319 interaction=0.66 [all-cause mortality], p-interaction=0.86 [CVD mortality] and p-
8
9 320 interaction=0.17 [stroke]). There was some evidence of interaction between HDI and PRS for
10
11 321 MI events (p-interaction=0.049). While there was no evidence of an effect of HDI on MI for
12
13 322 participants with low (-1 SD) PRS (HR=1.02 [95% CI: 0.95, 1.10], p=0.61), there was some
14
15 323 evidence of an association between higher HDI and reduced risk of MI events for those with
16
17 324 high (+1 SD) PRS (HR=0.93 [95% CI: 0.88, 0.99], p=0.017). The interaction of sex by PRS
18
19 325 showed evidence that the effect of higher PRS on higher risk of MI was more pronounced
20
21 326 for males (HR=1.21 [95% CI: 1.06, 1.37], p=.004). Effect sizes were consistent when deaths
22
23 327 and incident cases of MI and stroke within the first 2 years of follow up were excluded (data
24
25 328 not shown).

26 329

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

31 331 The adjusted HR associated with a one-point higher MDS for all-cause mortality was 0.95
32
33 332 (95% CI: 0.91, 0.98) and for stroke was 0.93 (95% CI: 0.87, 1.00) (Table 5). There was limited
34
35 333 evidence of associations between MDS and risk of CVD mortality and MI. There was limited
36
37 334 evidence (all p-values > 0.1) of sex by diet interactions. When stratified by sex, there was
38
39 335 evidence of an association between MDS and all-cause mortality and MI in males only. The
40
41 336 adjusted HR associated with a one-point higher PRS for MI was 1.33 (95% CI: 1.25, 1.41);
42
43 337 when stratified by sex, there was evidence of a stronger association in males. When an
44
45 338 interaction term for PRS was added to the models, there was no evidence of interaction
46
47 339 between MDS and PRS for other outcomes (p-interaction=0.58 [all-cause mortality], p-
48
49 340 interaction=0.72 [CVD mortality] and p-interaction=0.12 [stroke]). There was evidence of
50
51 341 interaction between MDS and PRS for MI events (p-interaction=0.026). While there was no
52
53 342 evidence of an effect of MDS on MI for those with low (-1 SD) PRS (HR=1.03 [95% CI: 0.94,
54
55 343 1.12], p=0.56) there was strong evidence of an association between higher MDS and
56
57 344 reduced risk of MI events for those with high (+1 SD) PRS (HR=0.91 [95% CI: 0.85, 0.97],
58
59 345 p=0.004). The interaction of sex by PRS showed evidence that the effect of higher PRS on
60
346 higher risk of MI was more pronounced for males (HR=1.21 [95% CI: 1.06, 1.37], p=.004).

1
2
3 347 Effect sizes were consistent when deaths and incident cases of MI and stroke within the first
4
5 348 2 years of follow up were excluded (data not shown).
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

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}

1
2
3 380 Future diet-disease research should extend this to better understand the role of specific
4
5 381 plant and animal foods as part of overall dietary patterns. Moreover, the UK Biobank
6
7 382 participants are predominately white and highly-educated and genetic risk profiles and their
8
9 383 associations with risk of CVD may be different in more diverse populations. Thus, further
10
11 384 research in diverse populations is needed to investigate the applicability of these diet
12
13 385 quality methodologies for examining CVD risk independent of genetic risk.⁷

14
15 386 Although there was limited evidence of interactions for sex by diet quality index, our
16
17 387 stratified results showed large effect sizes for associations between diet quality and genetic
18
19 388 CVD risk in males. This warrants further investigation as previous research shows stronger
20
21 389 associations in males than females.^{14 60} The role of sex may be partly explained by the high
22
23 390 prevalence of diabetes and unhealthy behaviours in men,⁶¹ and in this study may be due the
24
25 391 lower number of events and deaths in women compared with men in the present study.
26
27 392 Nonetheless, it is possible that the biological and behavioural pathways in which risk factors
28
29 393 exert their effects on CVD risk are different between men and women.⁶⁰

30 394 **Strengths and limitations**

31
32
33 395 Our main strength was the large sample size and inclusion of genetic data. This enabled
34
35 396 investigation of a genetic risk score created 300 SNPs known to be associated with CVD,
36
37 397 more than any previous publications in the UK Biobank.^{14 57} While the PRS used was specific
38
39 398 to coronary disease, it has been used to identify predispositions to a wide variety of CVD
40
41 399 and non-CVDs, as well as premature mortality, given these may develop in parallel with
42
43 400 coronary disease for the same genetic origins. The dietary questionnaire has been
44
45 401 previously validated and included sufficient detail to allow us to create three contrasting
46
47 402 diet quality indices. There are a number of limitations that should be acknowledged. While
48
49 403 the dietary assessment method is a short-term measure of intake, our use of up to four
50
51 404 instances of dietary assessments provided an estimate of longer-term intake. Although the
52
53 405 present analysis is likely to be subject to self-selection bias associated with the number of
54
55 406 participants who completed the dietary assessment and the low response rate, associations
56
57 407 between demographic and behavioural risk factors and mortality in the UK Biobank have
58
59 408 been shown to be comparable to those from national health survey data from England and
60
60 409 Scotland.⁶² Whilst we adjusted analyses based on a range of confounders identified using a

1
2
3 410 directed acyclic graph, we cannot discount the possibility of residual or unmeasured
4
5 411 confounding.

6
7
8 412 **Conclusion**

9
10 413 This prospective population-based cohort study provided evidence that higher diet quality
11 414 (RFS, HDI and MDS) was associated with lower risk of all-cause mortality, regardless of
12 415 genetic CVD risk. Diet quality, when estimated using the RFS only, was associated with lower
13 416 risk of CVD mortality and MI, independent of genetic CVD risk. The diet quality indices
14 417 investigated in this study have common food-based scoring components, providing further
15 418 evidence for the best practice design and implementation of food-based diet quality indices
16 419 for assessing health outcomes. Further research in diverse populations is needed to
17 420 investigate the applicability of different diet quality methodologies for examining CVD risk
18 421 independent of genetic susceptibility.
19
20
21
22
23
24
25
26
27
28
29

30 423 **Authors' contributions:** KML, GA, SB, JW, CM and SAM designed the analysis. KML, SB, JW
31 424 and GA conducted the statistical analysis. KML drafted the manuscript. All authors
32 425 contributed to the critical review of the manuscript and approved the final version of the
33 426 manuscript.
34
35
36

37 427 **Funding:** KML is supported by a National Health and Medical Research Council Emerging
38 428 Leadership Fellowship (APP1173803). JW is funded by the Lister Prize Fellowship (173096).
39 429 The funding source had no role in the design or conduct of the study; collection,
40 430 management, analysis, and interpretation of the data; or preparation, review, or approval of
41 431 the manuscript.
42
43
44
45
46

47 432 **Competing interests:** All authors have completed the ICMJE uniform disclosure form at
48 433 www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the
49 434 submitted work; KML is a consultant for HeadUp Labs; no other financial relationships with
50 435 any organisations that might have an interest in the submitted work in the previous three
51 436 years; no other relationships or activities that could appear to have influenced the
52 437 submitted work.
53
54
55
56

57 438 **Ethical approval:** UK Biobank received ethical approval from the research ethics committee
58 439 (reference 13/NW/0382). All participants provided informed consent to participate. An
59
60

1
2
3 440 ethics exemption was granted by Deakin University Human Research Ethics Committee
4
5 441 (Reference number 2019_293).
6

7 442 **Data availability statement:** The genetic and phenotypic UK Biobank data are available on
8
9 443 application to the UK Biobank. This research has been conducted using the UK Biobank
10
11 444 Resource under Application 34894.

12 445 **Transparency:** The lead author (KML) affirms that the manuscript is an honest, accurate,
13
14 446 and transparent account of the study being reported; that no important aspects of the study
15
16 447 have been omitted and that discrepancies from the study as planned have been explained
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

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](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: 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: 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 Office 2005.
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

- 1
2
3 36. Trichopoulou A, Orfanos P, Norat T, et al. Modified Mediterranean diet and survival: EPIC-elderly
4 prospective cohort study. 2005;330(7498):991. doi: 10.1136/bmj.38415.644155.8F %J BMJ
5
6 37. Trichopoulou A, Costacou T, Bamia C, et al. Adherence to a Mediterranean Diet and Survival in a
7 Greek Population. 2003;348(26):2599-608. doi: 10.1056/NEJMoa025039
8
9 38. Bach-Faig A, Berry EM, Lairon D, et al. Mediterranean diet pyramid today. Science and cultural
10 updates. *Public Health Nutr* 2011;14(12A):2274-84. doi: 10.1017/S1368980011002515
11 [published Online First: 2011/12/13]
12
13 39. Magliano D, Liew D, Pater H, et al. Accuracy of the Australian National Death Index: comparison
14 with adjudicated fatal outcomes among Australian participants in the Long-term
15 Intervention with Pravastatin in Ischaemic Disease (LIPID) study. 2003;27(6):649-53. doi:
16 10.1111/j.1467-842X.2003.tb00615.x
17
18 40. Schnier C, Bush K, Nolan J, et al. Definitions of Stroke for UK Biobank Phase 1 Outcomes
19 Adjudication Documentation prepared by: On behalf of UK Biobank Outcome Adjudication
20 Group. 2017. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_mi.pdf
21 (accessed 16 September 2020).
22
23 41. Schnier C, Bush K, Nolan J, et al. Definitions of Acute Myocardial Infarction and Main Myocardial
24 Infarction Pathological Types UK Biobank Phase 1 Outcomes Adjudication Documentation
25 prepared by: On behalf of UK Biobank Outcome Adjudication Group. 2017.
26 http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_mi.pdf (accessed 16
27 September 2020).
28
29 42. National Health Service. COVID-19 Daily Deaths. 2020.
30 <https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-daily-deaths/>
31 (accessed 16 September 2020).
32
33 43. Ntalla I, Kanoni S, Zeng L, et al. Genetic Risk Score for Coronary Disease Identifies Predispositions
34 to Cardiovascular and Noncardiovascular Diseases. *J Am Coll Cardiol* 2019;73(23):2932-42.
35 doi: <https://doi.org/10.1016/j.jacc.2019.03.512>
36
37 44. Dudbridge F. Polygenic Epidemiology. 2016;40(4):268-72. doi: 10.1002/gepi.21966
38
39 45. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and
40 population-based linkage analyses. *Am J Hum Genet* 2007;81(3):559-75. doi:
41 10.1086/519795 [published Online First: 2007/08/19]
42
43 46. Townsend P, Phillimore P, Beattie A. Health and Deprivation. *Nurs Stand* 1988;2(17):34. doi:
44 10.7748/ns.2.17.34.s66 [published Online First: 1988/01/30]
45
46 47. World Health Organization. BMI classification 2015 [updated 21 July 2015. Available from:
47 http://apps.who.int/bmi/index.jsp?introPage=intro_3.html accessed 3rd December 2015.
48
49 48. The IPAQ Group. IPAQ scoring protocol - International Physical Activity Questionnaire. Available
50 at: <https://sites.google.com/site/theipaq/scoring-protocol>. Accessed 22 May 2020. [
51
52 49. Woodward M. Rationale and tutorial for analysing and reporting sex differences in cardiovascular
53 associations. *Heart* 2019;105(22):1701-08. doi: 10.1136/heartjnl-2019-315299
54
55 50. Neelakantan N, Koh W-P, Yuan J-M, et al. Diet-Quality Indexes Are Associated with a Lower Risk
56 of Cardiovascular, Respiratory, and All-Cause Mortality among Chinese Adults. *The Journal of*
57 *Nutrition* 2018;148(8):1323-32. doi: 10.1093/jn/nxy094 %J The Journal of Nutrition
58
59 51. Stefler D, Malyutina S, Kubinova R, et al. Mediterranean diet score and total and cardiovascular
60 mortality in Eastern Europe: the HAPIEE study. *Eur J Nutr* 2017;56(1):421-29. doi:
10.1007/s00394-015-1092-x
52
53 52. Lassale C, Gunter MJ, Romaguera D, et al. Diet Quality Scores and Prediction of All-Cause,
54 Cardiovascular and Cancer Mortality in a Pan-European Cohort Study. *PLoS One*
55 2016;11(7):e0159025. doi: 10.1371/journal.pone.0159025
56
57 53. Kaluza J, Håkansson N, Brzozowska A, et al. Diet quality and mortality: a population-based
58 prospective study of men. *Eur J Clin Nutr* 2009;63(4):451-57. doi: 10.1038/sj.ejcn.1602968
59
60 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

- 1
2
3 55. Lakkur S, Judd SE. Diet and Stroke. 2015;46(7):2007-11. doi:
4 doi:10.1161/STROKEAHA.114.006306
5
6 56. Chiavaroli L, Viguioliouk E, Nishi SK, et al. DASH Dietary Pattern and Cardiometabolic Outcomes:
7 An Umbrella Review of Systematic Reviews and Meta-Analyses. *Nutrients* 2019;11(2) doi:
8 10.3390/nu11020338 [published Online First: 2019/02/16]
9
10 57. Said MA, Verweij N, van der Harst P. Associations of Combined Genetic and Lifestyle Risks With
11 Incident Cardiovascular Disease and Diabetes in the UK Biobank Study. *JAMA Cardiol*
12 2018;3(8):693-702. doi: 10.1001/jamacardio.2018.1717 [published Online First: 2018/06/30]
13
14 58. Baden MY, Liu G, Satija A, et al. Changes in Plant-Based Diet Quality and Total and Cause-Specific
15 Mortality. *Circulation* 2019;140(12):979-91. doi: doi:10.1161/CIRCULATIONAHA.119.041014
16
17 59. Huang T, Yang B, Zheng J, et al. Cardiovascular Disease Mortality and Cancer Incidence in
18 Vegetarians: A Meta-Analysis and Systematic Review. *Ann Nutr Metab* 2012;60(4):233-40.
19 doi: 10.1159/000337301
20
21 60. Den Ruijter HM, Haitjema S, Asselbergs FW, et al. Sex matters to the heart: A special issue
22 dedicated to the impact of sex related differences of cardiovascular diseases. *Atherosclerosis*
23 2015;241(1):205-07. doi: 10.1016/j.atherosclerosis.2015.05.003
24
25 61. Andersen KK, Andersen ZJ, Olsen TS. Age- and Gender-Specific Prevalence of Cardiovascular Risk
26 Factors in 40 102 Patients With First-Ever Ischemic Stroke. *Stroke* 2010;41(12):2768-74. doi:
27 doi:10.1161/STROKEAHA.110.595785
28
29 62. Batty GD, Gale CR, Kivimäki M, et al. Comparison of risk factor associations in UK Biobank against
30 representative, general population based studies with conventional response rates:
31 prospective cohort study and individual participant meta-analysis. *BMJ* 2020;368:m131. doi:
32 10.1136/bmj.m131
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

Dietary Indicator	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. <hr/> 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. <hr/> 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 Indicator	Indicator foods¹	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)	<400 g/day=0 ≥400 g/day=1
	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	
7. Pulses and nuts	Baked bean, pulses, broad bean	<30 g/day=0 ≥30 g/day=1
	Salted peanuts, unsalted peanuts, salted nuts, unsalted nuts, seeds, types of spreads/sauces consumed (peanut butter)	
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	<32 g/day=0 ≥32 g/day=1
	Homemade soup, ingredients in homemade soup (fish)	
10. Red meat and meat products	Beef, pork, lamb, other meat	>90 g/day=0 ≤90 g/day=1
	Poultry intake (skin removed from poultry (no); fat removed from poultry(no))	
	Homemade soup, ingredients in homemade soup (meat)	
11. Calcium	Sausage, bacon, ham, liver	<700 mg/day=0 ≥700 mg/day=1
	Calcium	

1. As available in the UK Biobank

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

Dietary Indicator	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 indicators 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 indicators 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	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

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

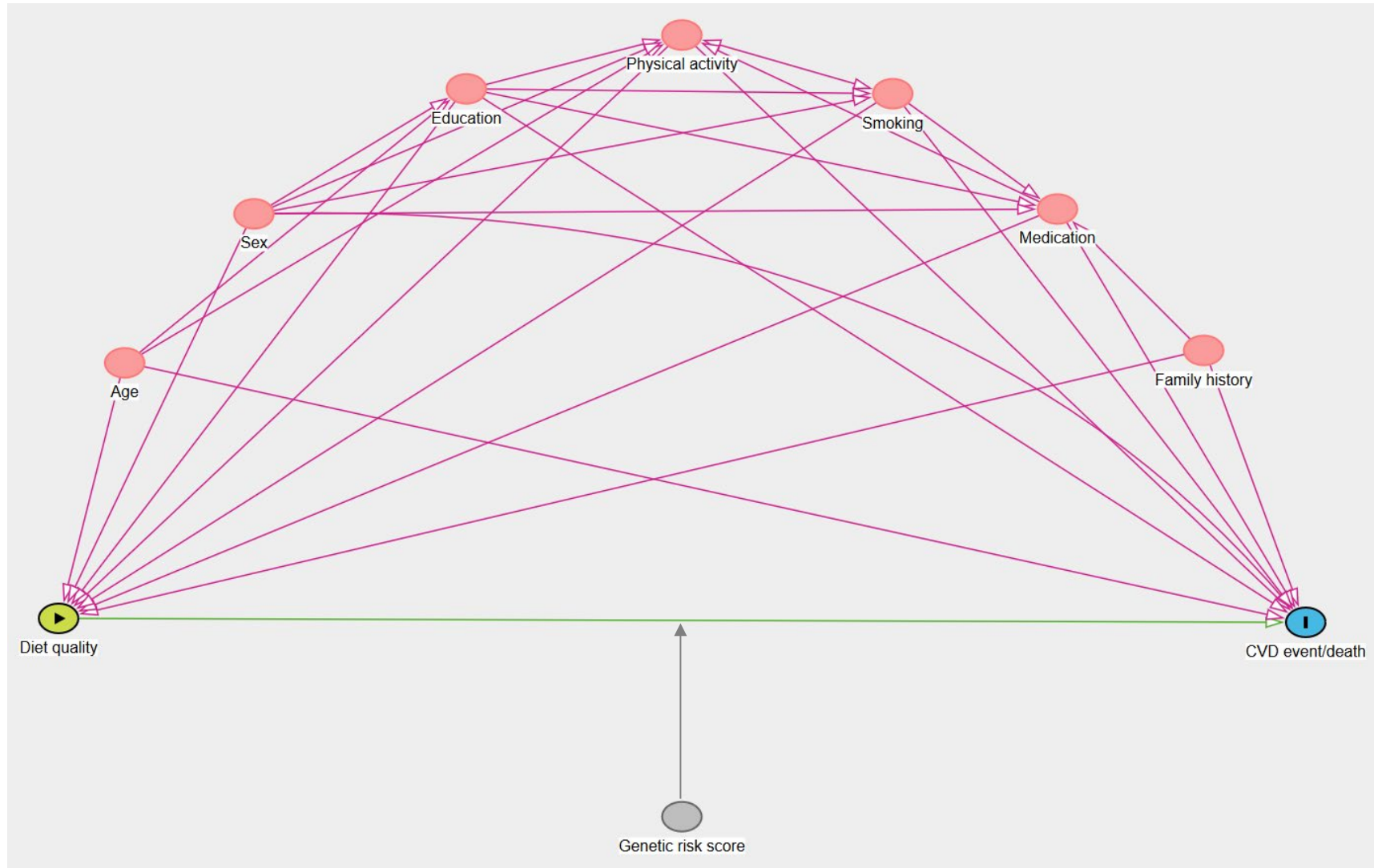
For peer review only

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

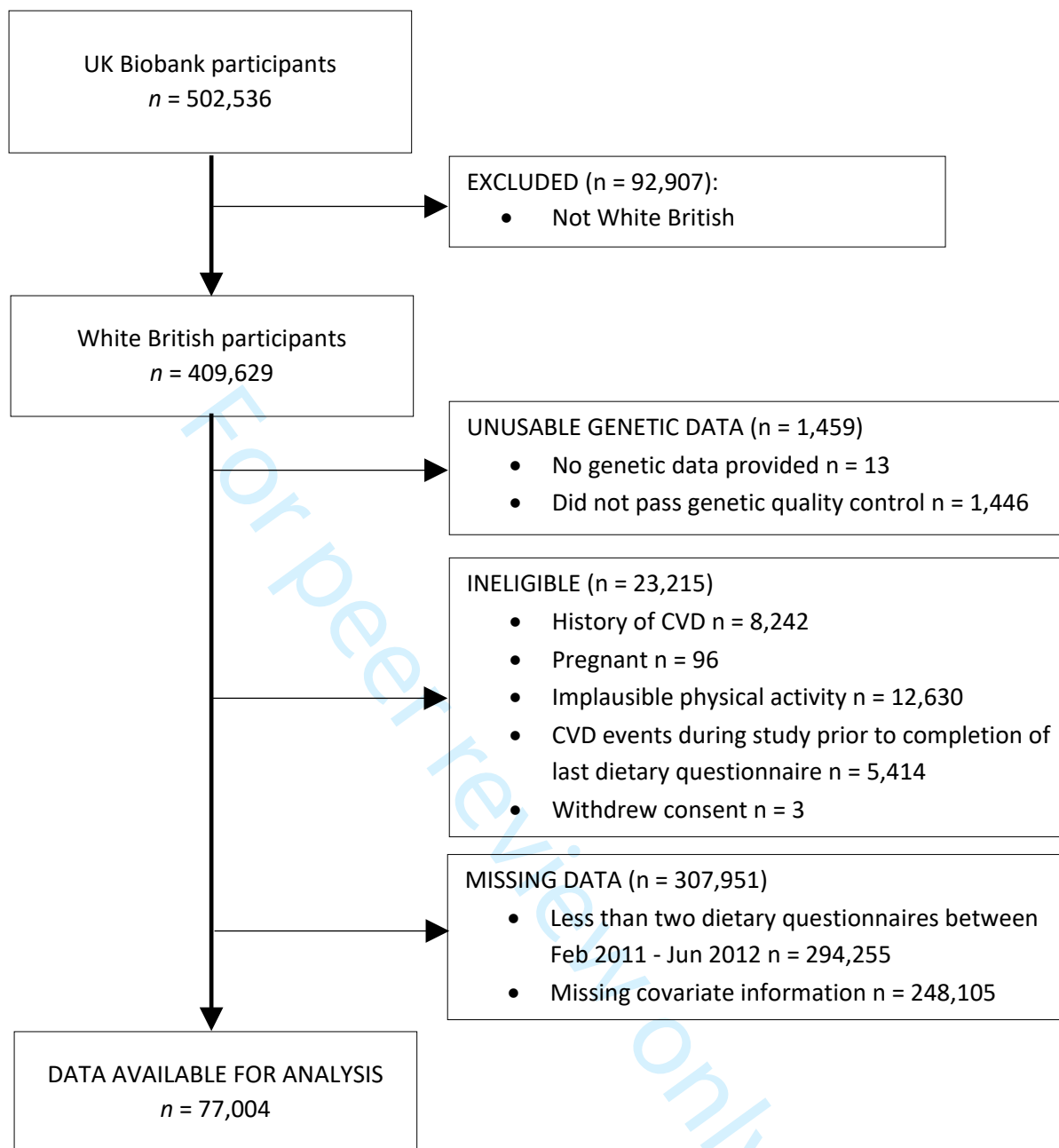
Characteristic	Excluded N (%)	Analytic N (%)
N ¹	425,532	77,004
Sex female	231,382 (54.4)	42,020 (54.6)
Age at recruitment (years), Mean \pm SD	56.6 \pm 8.2	56.2 \pm 7.8
Townsend Deprivation Index		
Least deprived	82,535 (19.4)	18,129 (23.5)
2nd least deprived	82,878 (19.5)	17,227 (22.4)
Medium deprivation	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 \pm SD	27.6 \pm 4.9	26.5 \pm 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

1
2
3 conjunction with this article (freely available on the Web sites of PLoS Medicine at
4 <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and
5 Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at
6 <http://www.strobe-statement.org>.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only