Proceedings of the Nutrition Society, Page 1 of 8 doi:10.1017/S0029665122001094

Conference: Plant-rich dietary patterns and health symposium
Meeting: Member-led conference paper

Dietary patterns derived by reduced rank regression and non-communicable disease risk

Carmen Piernas[1,2]*, Min Gao[2] and Susan A. Jebb[2]
[1]*Department of Biochemistry and Molecular Biology II, Center for Biomedical Research (CIBM), Institute of Nutrition and Food Technology (INYTA), University of Granada, Granada, Spain*
[2]*Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK*

Most current nutrition policies and dietary recommendations still reflect decades of research addressing the mechanism of action or health risks of individual nutrients. Yet, most high-income countries including the UK are far from reaching the dietary intakes which are recommended for good health. Food-based dietary patterns (DPs) can help target specific combinations of foods that are associated with disease risk, recognising the coexistence of multiple nutrients within foods and their potential synergistic effects. Reduced rank regression (RRR) has emerged as a useful exploratory approach which uses *a priori* knowledge of the pathway from diet to disease to help identify DPs which are associated with disease risk in a particular population. Here we reviewed the literature with a focus on longitudinal cohort studies using RRR to derive DPs and reporting associations with non-communicable disease risk. We also illustrated the application of the RRR approach using data from the UK Biobank study, where we derived DPs that explained high variability in a set of nutrient response variables. The main DP was characterised by high intakes of chocolate and confectionery, butter and low-fibre bread, and low intakes of fresh fruit and vegetables and showed particularly strong associations with CVD, type 2 diabetes and all-cause mortality, which is consistent with previous studies that derived 'Western' or unhealthy DPs. These recent studies conducted in the UK Biobank population together with evidence from previous cohort studies contribute to the emerging evidence base to underpin food-based dietary advice for non-communicable disease prevention.

Key words: Dietary patterns: Reduced rank regression: CVD: Diabetes

The burden of non-communicable diseases such as heart disease or diabetes continues to increase, and this trend is accelerated by poor dietary habits[1–3]. Many of the top twenty risk factors reported in the Global Burden of Disease study are related to diet, including low intakes of whole grains, legumes, nuts, fruits, fibre or vegetables, and high intakes of red and processed meats, *trans* fat or sodium[4]. Despite efforts to improve population diets, most high-income countries including the UK are far from reaching recommended dietary intakes consistent with good health[5]. Many current policies and dietary recommendations are still based on single nutrients or foods[6,7], and media attention often focuses on one specific nutrient (e.g. sugar or saturated fat). In reality, food intake is a multi-dimensional exposure with many dietary risk factors showing high correlations between themselves, and the biological effects of different nutrients may also have synergistic effects[8,9].

A whole-diet approach to disease prevention is increasingly being adopted in policy and practice. There are several methods to derive dietary patterns (DPs) but this review focuses on reduced rank regression

Abbreviations: DP, dietary pattern; RRR, reduced rank regression; SSB, sugar-sweetened beverage.
***Corresponding author:** Carmen Piernas, email carmen.piernas-sanchez@phc.ox.ac.uk; carmenpiernas@ugr.es

(RRR), which uses *a priori* knowledge about risk factors or nutrients of concern to help identify DPs in a particular study population[10]. We include a detailed explanation of the RRR approach to derive DPs from the research by the authors in the UK Biobank study. The UK Biobank study is a prospective cohort that recruited over 500 000 adults, aged 40–69 at baseline (2006–2010), who provided detailed measures of dietary intake. Together with previous evidence from longitudinal cohort studies using RRR-derived DPs, we present observational associations between RRR-derived DPs and the risk of health outcomes as well as other cardiometabolic risk factors in the UK Biobank population[11,12].

Dietary pattern research

A whole-diet approach is an important tool in nutrition research to capture the inherently complex nature of food intake. Energy intake in human subjects has been shown to be relatively stable, and this implies that changes in one food or nutrient will be associated with changes in another one, also known as substitution effects[13,14]. Therefore, making inferences about a particular food or nutrient in observational studies is problematic and residual confounding will likely affect the observed associations with health outcomes. DP approaches can potentially overcome these issues, while capturing any synergistic effects and the cumulative exposure to different foods and nutrients[9], which in turn may result in stronger associations with health outcomes than what has been observed with single nutrients or foods[8,10].

In observational studies, there are two main approaches used to derive DPs using self-reported measures of dietary intakes: *a priori* or pre-defined DP, and *a posteriori* or data-driven/exploratory DP[8,15]. Pre-defined DP approaches include specific diets (i.e. vegetarian or vegan) as well as dietary indices or scores, which are used as a proxy for adherence to a particular diet (i.e. Mediterranean dietary score, healthy eating index). Exploratory approaches are data-driven because they use different statistical techniques to derive particular combinations of food groups from dietary data collected in a specific population. The most widely used exploratory approaches include factor and principal component analysis, partial least squares and RRR.

Both *a priori* and *a posteriori* approaches aim to reduce a large amount of information from complex diets in order to investigate diet–disease associations, but all these methods have particular pros and cons that have been summarised elsewhere[8,9,15]. In any case, it is critical to choose the most appropriate approach that aligns with a well-developed research question, as this can influence the observed associations with disease risk[8,10,16].

Dietary patterns derived by reduced rank regression

RRR has emerged as a useful hybrid approach which uses *a priori* knowledge about nutrients or biomarkers with an established link with disease, to help identify data-driven combinations of foods in a particular population[10]. In common with all the other exploratory approaches, the RRR method aims to extract successive linear combinations of predictor variables (i.e. food groups). However, the main goal of RRR is to explain the maximum variability in response variables (i.e. nutrients of interest with established associations with disease risk); while the other exploratory approaches aim to maximise the variability in the predictors (principal component analysis) or a balance between predictor and response variation (partial least squares)[10].

The RRR model can include one or more response variables, and a number of uncorrelated factors (DPs) will be derived which is equal to the number of specified response variables in the model. The variation in each of the response variables that is explained by each DP gives an indication of the likely 'strength' between each DP and each response variable. Typically, the DPs that explain the maximum variability in all the response variables collectively are retained for analysis. Each predictor variable in the model (i.e. food group) will also receive a factor loading, so that a larger value (positive or negative) indicates that the food group makes a greater contribution to the DP. Participants in a study population can then obtain an overall score (z score) for each of the derived DPs based on the factor loading and their intake of each food group, representing the degree to which their dietary intake adheres to a DP relative to other participants.

A key strength of the RRR approach is that the derived DPs use *a priori* knowledge of established associations between biomarkers or nutrients of interest and disease, hence it is particularly appealing if the aim is to identify DPs in association with disease endpoints. The emerging DPs are food-based, meaning they are easier to interpret and communicate, which helps inform public health advice. However, RRR-derived DPs share limitations in common with the rest of approaches, for example, subjective decisions are made about the number of DPs that will be retained for analysis, and it is unclear which food group(s) may drive the associations. In particular, RRR relies on the availability of response variables with solid evidence of their associations with disease outcomes; and by choosing a specific set of nutrient response variables, there will be some variability explained by other nutrients which may also be involved in the disease pathway of interest[10,15].

Dietary patterns emerging from longitudinal cohort studies

The state of the evidence with regards to RRR-derived DPs and non-communicable diseases such as CVD and diabetes is gradually growing. A literature search on PubMed and Embase up to May 2022 returned seventeen studies using data from longitudinal cohorts looking at prospective associations between RRR-derived DPs with CVD risk and/or all-cause mortality[11,16–31], and nineteen studies investigating associations with the risk of developing diabetes[12,32–49]. The vast majority of these studies have been published since 2015, and

https://doi.org/10.1017/S0029665122001094 Published online by Cambridge University Press

have included samples from a variety of populations including the USA (*n* 8[20,24,25,31,39,40,42,45,46]); Europe (*n* 11[16,18,19,21,23,26,29,35–37,47,48]), Asia (*n* 6[22,27,28,32,33,41,49]), UK (*n* 4[11,12,30,34,43,44]) or others (*n* 2[17,38]). Population specificity is important here, because even when the choice of response variables (e.g. nutrients) is the same, the emerging DPs can vary substantially between different populations, owing to cultural differences in food intakes.

Until very recently, evidence with regards to RRR-derived DPs in association with CVD and all-cause mortality risk in the UK population was limited to one small study[30], while the evidence with regards to diabetes was constrained to two small cohort studies[43,44]. Here we focus on the use of the UK Biobank study to illustrate the application of the RRR approach to derive DPs and summarise the prospective associations between RRR-derived DPs and CVD[11], type 2 diabetes[12] and all-cause mortality[11] in the context of previously conducted studies. We have also investigated associations with cardiometabolic risk factors (e.g. BMI, blood pressure or lipid biomarkers) to help understand potentially underlying mechanisms. The UK Biobank is a prospective cohort that recruited over 500 000 adults between 2006 and 2010, aged 40–69 years at baseline. Together with important information on demographic, lifestyle factors, medical history, physical measures (height and weight) and blood and urine samples, detailed dietary intake measures were collected through an online questionnaire (Oxford WebQ) administered via email on up to five times during approximately 1·5–2 years[50,51]. The WebQ is a validated 24-h dietary assessment tool which collected quantitative information on two hundred and six widely consumed foods and thirty-two beverages[52–54]. We developed a food grouping system to group the number of reported foods and beverages into fifty major food groups before including them into the RRR model[55,56], as this would also facilitate obtaining meaningful and relevant DPs. Four nutrient response variables were included: energy density (kJ/g), saturated fat (% total energy), free sugars (% total energy) and fibre density (g/MJ). These nutrient variables have established associations with the development of obesity, type 2 diabetes and CVD[57–66], which underscores the importance of a well-supported hypothesis underlying the proposed research. From the initial baseline cohort, a subsample of participants that provided at least two 24-h questionnaires (*n* ~117 000[11]; *n* ~120 000[12]) and free of disease at baseline, were included and the mean dietary intake across multiple assessments was calculated to derive DPs.

With four response variables, the RRR model extracted four DPs together explaining a total of 77 % of variability in the response variables. The top two DPs explaining 43 % (DP1) and 20 % (DP2) of variability were retained for analyses. DP1 showed high factor loadings for chocolate and confectionery, butter and other spreads, low-fibre bread, table sugars, grain-based desserts, sugar-sweetened beverages (SSBs) among others; and low factor loadings for fresh fruit, vegetables and high-fibre breakfast cereals. DP1 showed strong positive correlations with energy density, free sugars and saturated fat, but negative correlations with fibre. Conversely, DP2 showed high factor loadings for SSBs, fruit juice, table sugars and preserves, chocolate confectionery; and low factor loadings for high-fat cheese, butter and other animal fat spreads. DP2 showed a strong positive correlation with free sugars, but a negative correlation with saturated fat, while energy density or fibre were not correlated[11,12].

There is wide heterogeneity in previous studies regarding the choice of response variables in the RRR model, and most studies included combinations of nutrients[16,21,34,44,67] (% energy from monounsaturated fats and polyunsaturated fats; protein and total carbohydrates), as well as biomarkers[25,27,32] or disease risk factors (BMI, waist circumference, blood lipids or blood pressure)[30,35,38,49]. The majority of these previous studies obtained dietary measurements from food frequency questionnaires with fewer using either 24 h recalls or food records/diaries. The resulting DPs yielded combinations of food groups which were population-specific, although most studies identified at least one unhealthy or 'Western' DP characterised by high intakes of soft drinks, alcohol, refined carbohydrates, added sugars, fats, meat (including processed) and low fruit/vegetable intakes.

Associations with health outcomes

CVD

Previous research studies have derived CVD-related DPs and their associations with several outcomes, mostly using response variables such CVD biomarkers or risk factors[18,20,26,29–31,35,68] as oppose to nutrients[11,16,21,25]. In common with the DPs derived in the UK Biobank population, most of these previous DPs were characterised by intakes of high-fat/high-sugar, refined carbohydrates (i.e. low-fibre bread, table sugars, SSBs) and high saturated fat foods (i.e. butter and animal spreads, red and processed meats), but low intakes of fruit, vegetables or higher-fibre foods. Of the studies performed in US cohorts, one reported significant associations between DPs characterised by high meat and refined carbohydrates/potato products but low fruit intakes with CVD mortality[20]; and another one reported a DP characterised by high intakes of fats, SSBs, meat, but low vegetables in association with a higher incidence of stroke[31]. In European populations, a study in Germany reported significant inverse associations with CVD mortality among those following a DP high in whole grain, fruits, cereals and vegetables but low intakes of processed meat, butter and cheese[29]. Another study among Swedish participants with obesity reported a significant association with total CVD in people following an energy-dense, high SFA, low-fibre DP[21]. A study in the UK Whitehall II study reported significant associations with incident CHD in association with a DP characterised by high intakes of white bread, fried potatoes, added sugars, processed meat and SSBs but low intakes of vegetables[30].

The associations between DPs obtained in the UK Biobank population and several health outcomes are

https://doi.org/10.1017/S0029665122001094 Published online by Cambridge University Press

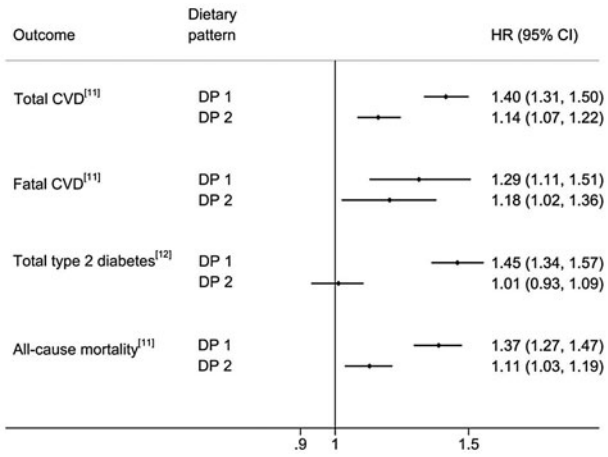

Fig. 1. Risk of health outcomes associated with dietary pattern 1 (DP1) and dietary pattern 2 (DP2) in the UK Biobank study. Adjusted hazard ratios (HR) and 95 % CI, showing the risk of each DP for people in the fifth quintile (indicating the highest adherence to each DP). Estimates were obtained from Cox proportional hazard models using age during study as the underlying timescale. The models were stratified by sex and regions (England, Scotland and Wales) and adjusted for socio-demographic characteristics: ethnicity (white, others, missing), Townsend index of deprivation (quintiles one to five, with higher scores representing greater deprivation), education (higher degree [college or university degree, or professional qualifications], any school degree [A levels, AS levels, O levels, General Certificate of Secondary Educations or Certificate of Secondary Educations], vocational qualifications [National Vocational Qualifications, Higher National Diploma or Higher National Certificate], other [none of the above qualifications], missing); behavioural risk factors: smoking status (never, current, previous, missing), physical activity (low [<600 metabolic equivalent (MET)-minutes per week], moderate [≥600 and <3000 metabolic equivalent (MET)-minutes per week], high [≥3000 metabolic equivalent (MET)-minutes per week], missing), log-transformed energy intake; and health history/conditions: family history of diabetes, menopause in women, hypertension, CVD, high cholesterol.

summarised in Fig. 1[11]. People in the highest quintile of adherence to DP1 showed a 40 % higher risk of total cardiovascular events and 29 % higher risk of fatal CVD events compared to people in the first quintile. DP2 showed non-linear associations, although there were 14 % higher risk of total CVD and 18 % higher risk of fatal CVD among people in the highest quintile of DP2 compared to those in the lowest quintile. Adjustment for the BMI revealed slightly attenuated associations only for the associations observed with DP1, suggesting that the observed association was potentially mediated through excess weight. These observations are broadly consistent with previous evidence from single-nutrient studies or other DP studies using dietary quality scores[69–72].

Diabetes

Most previous studies have investigated associations between DPs and incident type 2 diabetes using response variables such as biomarkers[38,39,42,49], diabetes markers such as HbA1c or the homoeostasis model assessment of insulin resistance[32,37,41,43] or nutrients[12,34,44,67]. In the US population, two studies found DPs characterised by high intakes of red meat/meat products, refined carbohydrates, SSBs that were significantly associated with type 2 diabetes[38,42], while the opposite DP characterised by high intakes of whole grains, fruit, vegetables and low-fat dairy was inversely associated with type 2 diabetes[39]. A study among German adults reported a DP rich in fresh fruit but low in SSBs, alcohol, red/processed meat that was inversely associated with type 2 diabetes[37]; while less traditional DPs among Asian populations were generally associated with type 2 diabetes incidence[32,41,49]. Two small studies including UK middle-aged adults reported a high fat/glycaemic index DP[44] and a DP with high intakes of SSBs, fast foods/snacks, white bread[43] which were both significantly associated with type 2 diabetes. The associations observed in the UK Biobank study with type 2 diabetes showed that those in the highest quintile of adherence to DP1 showed a 45 % higher risk of incident diabetes compared to people in the first quintile (Fig. 1)[12]. However, there were non-significant associations between DP2 and diabetes incidence. Adjustment for BMI in the main model also attenuated the association observed with DP1, while the association with DP2 remained not statistically significant.

Most of the food groups and underlying nutrients that characterise DP1 in the UK Biobank study as well as other unhealthy DPs derived in previous cohorts have been shown to be significantly associated with diabetes and adverse metabolic effects[73,74], particularly for chocolate confectionery[75], SSBs[76,77] or added sugars[78]; as well as low consumption of fruits, vegetables and higher-fibre bread[79,80]. For the unexpected lack of association between DP2 and diabetes, it is important to note that this DP was characterised by high intakes of foods rich in free sugars but in the context of low intakes of saturated fat rich foods and energy density, and adequate intakes of fibre. This finding supports previous evidence suggesting that free sugars increase diabetes risk through increased energy intake and consequently weight gain, rather than through independent effects[64,81]. This has also been observed in short- and medium-term isoenergetic intervention trials showing no effects of free sugars on body weight or blood pressure[82].

All-cause mortality

The associations with CVD and diabetes observed in previous studies as well as the UK Biobank support the associations observed with the overall risk of mortality. Previous US and European cohort studies have generally reported significant associations between unhealthy/Western DPs and all-cause mortality[16,20,25,26]. In the UK Biobank[11], people in the highest quintile of adherence to DP1 showed a 37 % higher risk of all-cause mortality compared to people in the first quintile (Fig. 1). DP2 showed somewhat weaker associations, with 11 % higher risk of all-cause mortality among people in the highest quintile compared to those in the lowest quintile. These associations generally mirror the ones observed with total and fatal CVD events, which supports the

https://doi.org/10.1017/S0029665122001094 Published online by Cambridge University Press

evidence showing that CVD contributes a higher burden to total mortality than diabetes[83].

For this particular outcome, diet is one of the multiple factors that may act synergistically with others (i.e. smoking, physical activity, socioeconomic status) which can influence mortality risk, hence the results observed with all-cause mortality have to be carefully interpreted given the possibility of residual confounding in these associations due to non-dietary differences between the groups.

Cardiometabolic risk factors

Several studies have reported associations between DPs and markers of cardiometabolic risk which collectively help understand potential underlying mechanisms in the observed associations with non-communicable disease risk. A US study showed that a DP characterised by high SSBs, meat, potatoes was associated with markers of arterial stiffness[68]; while other studies in Swedish adults showed a higher incidence of metabolic syndrome among people following a Western DP[35]; and positive associations with cardiometabolic risk factors among people with obesity and high adherence to an energy-dense high SFA low-fibre DP[21]. Cross-sectional analyses in the UK Biobank population showed significant associations between the DPs and several cardiometabolic risk factors measured at baseline[11]. For DP1, people adhering the most to this pattern showed particularly large differences in baseline BMI (BMI 27·1 kg/m2 in the highest quintile compared to BMI 25·4 kg/m2 in the lowest quintile), and somewhat smaller differences in diastolic blood pressure (82 mmHg in the highest quintile compared to 79·7 mmHg in the lowest quintile) and HDL cholesterol (1·41 mM/l in the highest quintile compared to 1·59 mM/l in the lowest quintile). For DP2, there were no clear associations with baseline markers, except for a small clinical difference in LDL cholesterol (3·62 mM/l in the highest quintile compared to 3·70 mM/l in the lowest quintile) and HDL (1·45 mM/l in the highest quintile compared to 1·57 mM/l in the lowest quintile).

Strengths and limitations

The evidence summarised here comes from longitudinal cohort studies, with wide variability in population characteristics and sample sizes. The UK Biobank is one of the largest contemporary cohorts of UK adults who provided detailed measures of dietary intake, which is essential to investigate the associations between food-based DPs and long-term disease risk. The follow-up times since completion of the dietary questionnaires (from 5–8 years) was sufficient to accumulate enough cases of CVD, diabetes and all-cause mortality and to achieve more precise estimates of association. Linkages with electronic health records also allowed for an objective ascertainment of cases.

However, previous research as well as recent studies by the authors in the UK Biobank share common limitations that are inherent to observational studies. The derived DPs are based on self-reported measures of dietary intakes, which can be over- and/or underreported by participants introducing biases, although people with implausible dietary intakes were excluded in the studies by the authors. Day-to-day variation in diet was partly accounted for by including at least two dietary questionnaires, but some degree of bias due to random variation will remain[84]. In terms of investigating the associations between dietary exposures and longer-term health, it is clear that a lifetime exposure to a poor diet is a major determinant. However, most studies collected measurements of diet over a short time-period before disease manifestation and this may only reflect the most recent portion of this exposure. It is also unknown if DPs are stable over time or if changes preceding disease may affect the results. As with any observational research, caution is needed to avoid any causal inference and residual confounding (even after careful adjustment is made) cannot be ruled out. The cross-sectional associations investigating potential mechanisms through the associations with cardiometabolic risk factors measured at baseline may be affected by reversed causality, therefore careful interpretation is warranted. Finally, DPs identified in each study are culturally specific and relevant to each population, hence generalisability to other populations or countries is limited.

The main advantage of using RRR over other DP methods is that it allows an investigation of the cumulative effects of the whole diet, hence reducing confounding due to other dietary aspects, while also incorporating the mechanistic evidence through the integration of nutrient response variables[10]. The interpretation of the results is based on foods rather than nutrients, providing a more specific evidence base to inform the development of food-based dietary guidelines and making it easier to develop public health nutrition policies. However, the RRR method hinges on an imperfect evidence base to identify the underlying response variables and their association with disease endpoints. By focusing on a limited set of nutrients and foods which usually cannot explain 100 % of the nutrient response variance, it is possible that the emerging DPs contain other important nutrients which are involved in the biological pathway to disease but cannot be specified. For most chronic diseases there might be complex interactions in different metabolic pathways linking diet and disease, which may not be captured in a single RRR model with a limited number of response variables.

Implications and conclusions

In the context of previous evidence, recent DP research by the authors in the UK Biobank study supports current dietary recommendations for good health which emphasise reductions in saturated fat and free sugars, and increases in vegetables, fruits or grains, among others[6,7]. But this new evidence is food-based, and points towards diets which are lower in chocolate and confectionery, butter, low-fibre bread, added sugars and SSBs, and

https://doi.org/10.1017/S0029665122001094 Published online by Cambridge University Press

higher in fresh fruit and vegetables and whole grain foods to reduce the risk of CVD, diabetes and limit premature mortality. This should also help avoid confusing messages to the public regarding specific nutrients (e.g. saturated fat and sugars), recognising that foods are mixtures of nutrients and there might be synergistic effects between them. By identifying patterns in population-specific cohorts the hope is that food-based dietary advice may help accelerate behaviour change because it helps target specific foods and their combinations which are culturally relevant.

In conclusion, a whole-diet approach to disease prevention is attracting increasing interest from researchers, practitioners and policymakers. RRR is a particularly useful tool to derive population-specific food-based DPs using *a priori* knowledge about biomarkers or nutrients of concern which mediate the associations between the emerging DPs and disease outcomes and to help provide culturally relevant information to underpin behaviour change strategies to reduce non-communicable diseases.

Acknowledgements

We thank all participants in the UK Biobank for their invaluable contribution. The research summarised here was conducted using the UK Biobank resource under application number 14990.

Financial Support

This manuscript was prepared without specific funding. C. P. is funded by a Ramon y Cajal Fellowship RYC2020-028818-I (Ministry of Science and Innovation, Spain). S. A. J. and M. G. are funded by the Oxford and Thames Valley NIHR Applied Research Centre and the NIHR Oxford Biomedical Research Centre. The funders had no role in the study, or the decision to submit the paper. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Funding for open access charge: Universidad de Granada.

Conflict of Interest

None.

Authorship

The authors had sole responsibility for all aspects of preparation of the paper.

References

1. Afshin A, Sur PJ, Fay KA *et al.* (2019) Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **393**, 1958–1972.
2. Meier T, Gräfe K, Senn F *et al.* (2019) Cardiovascular mortality attributable to dietary risk factors in 51 countries in the WHO European region from 1990 to 2016: a systematic analysis of the global burden of disease study. *Eur J Epidemiol* **34**, 37–55.
3. Ezzati M & Riboli E (2013) Behavioral and dietary risk factors for noncommunicable diseases. *N Engl J Med* **369**, 954–964.
4. Collaborators GRF (2020) Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *Lancet* **396**, 1223–1249.
5. (2018) National Diet and Nutrition Survey: results from years 7 and 8 (combined). https://www.gov.uk/government/statistics/ndns-results-from-years-7-and-8-combined (accessed 16 March 2018).
6. Services USDoAaUSDoHaH (2020) Dietary Guidelines for Americans, 2020–2025. 9th Edition. DietaryGuidelines.gov.
7. Service UKNH (2019) Eat well. https://www.nhs.uk/live-well/eat-well/.
8. Zhao J, Li Z, Gao Q *et al.* (2021) A review of statistical methods for dietary pattern analysis. *Nutr J* **20**, 1–18.
9. Schulze MB & Hoffmann K (2006) Methodological approaches to study dietary patterns in relation to risk of coronary heart disease and stroke. *Br J Nutr* **95**, 860–869.
10. Hoffmann K, Schulze MB, Schienkiewitz A *et al.* (2004) Application of a new statistical method to derive dietary patterns in nutritional epidemiology. *Am J Epidemiol* **159**, 935–944.
11. Gao M, Jebb SA, Aveyard P *et al.* (2021) Associations between dietary patterns and the incidence of total and fatal cardiovascular disease and all-cause mortality in 116,806 individuals from the UK Biobank: a prospective cohort study. *BMC Med* **19**, 83.
12. Gao M, Jebb SA, Aveyard P *et al.* (2022) Associations between dietary patterns and incident type 2 diabetes: prospective cohort study of 120,343 UK biobank participants. *Diabetes Care* **45**, 1315–1325.
13. Willett W (2012) *Nutritional Epidemiology*, vol. **40**. Oxford, UK: Oxford University Press.
14. Willett WC, Howe GR & Kushi LH (1997) Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* **65**, 1220S–1228S.
15. Schulze MB, Martínez-González MA, Fung TT *et al.* (2018) Food based dietary patterns and chronic disease prevention. *Br Med J* **361**, k2396.
16. Hoffmann K, Boeing H, Boffetta P *et al.* (2005) Comparison of two statistical approaches to predict all-cause mortality by dietary patterns in German elderly subjects. *Br J Nutr* **93**, 709–716.
17. Appannah G, Pot G, Huang R-C *et al.* (2015) Identification of a dietary pattern associated with greater cardiometabolic risk in adolescence. *Nutr, Metab Cardiovasc Dis* **25**, 643–650.
18. Biesbroek S, Van Der A DL, Brosens MC *et al.* (2015) Identifying cardiovascular risk factor-related dietary patterns with reduced rank regression and random forest in the EPIC-NL cohort. *Am J Clin Nutr* **102**, 146–154.
19. Cai Q, Osté MC, Gomes-Neto AW *et al.* (2021) Metabolic syndrome-related dietary pattern and risk of mortality in kidney transplant recipients. *Nutr, Metab Cardiovasc Dis* **31**, 1129–1136.
20. Héroux M, Janssen I, Lam M *et al.* (2010) Dietary patterns and the risk of mortality: impact of cardiorespiratory fitness. *Int J Epidemiol* **39**, 197–209.
21. Johns DJ, Lindroos AK, Jebb SA *et al.* (2015) Dietary patterns, cardiometabolic risk factors, and the incidence of

https://doi.org/10.1017/S0029665122001094 Published online by Cambridge University Press

cardiovascular disease in severe obesity. *Obesity* **23**, 1063–1070.
22. Lee HA, An H & Lee E (2020) Dietary patterns related to cardiovascular disease based on reduced rank regression analysis of healthy middle-aged Koreans: data from the community-based Korean Genome and Epidemiology Study (KoGES) cohort. *Am J Clin Nutr* **111**, 1159–1169.
23. Leermakers ET, van den Hooven EH, Franco OH *et al.* (2018) *A priori* and *a posteriori* derived dietary patterns in infancy and cardiometabolic health in childhood: the role of body composition. *Clin Nutr* **37**, 1589–1595.
24. Li X, Chen B, Zhang J *et al.* (2022) Association of dietary inflammatory potential with risk of overall and cause-specific mortality. *Br J Nutr* **127**, 1878–1887.
25. Mazidi M, Katsiki N, Mikhailidis DP *et al.* (2019) Association of empirical dietary atherogenic indices with all-cause and cause-specific mortality in a multi-ethnic adult population of the United States. *Nutrients* **11**, 2323.
26. Meyer J, Döring A, Herder C *et al.* (2011) Dietary patterns, subclinical inflammation, incident coronary heart disease and mortality in middle-aged men from the MONICA/KORA Augsburg cohort study. *Eur J Clin Nutr* **65**, 800–807.
27. Seah JY, Ong CN, Koh W-P *et al.* (2019) A dietary pattern derived from reduced rank regression and fatty acid biomarkers is associated with lower risk of type 2 diabetes and coronary artery disease in Chinese adults. *J Nutr* **149**, 2001–2010.
28. Shi Z & Ganji V (2020) Dietary patterns and cardiovascular disease risk among Chinese adults: a prospective cohort study. *Eur J Clin Nutr* **74**, 1725–1735.
29. Drogan D, Hoffmann K, Schulz M *et al.* (2007) A food pattern predicting prospective weight change is associated with risk of fatal but not with nonfatal cardiovascular disease. *J Nutr* **137**, 1961–1967.
30. McNaughton SA, Mishra GD & Brunner EJ (2009) Food patterns associated with blood lipids are predictive of coronary heart disease: the Whitehall II study. *Br J Nutr* **102**, 619–624.
31. Nazari SSH, Mokhayeri Y, Mansournia MA *et al.* (2018) Associations between dietary risk factors and ischemic stroke: a comparison of regression methods using data from the multi-ethnic study of atherosclerosis. *Epidemiol Health* **40**, e2018021.
32. Batis C, Mendez MA, Gordon-Larsen P *et al.* (2016) Using both principal component analysis and reduced rank regression to study dietary patterns and diabetes in Chinese adults. *Public Health Nutr* **19**, 195–203.
33. Batis C, Mendez MA, Sotres-Alvarez D *et al.* (2014) Dietary pattern trajectories during 15 years of follow-up and HbA1c, insulin resistance and diabetes prevalence among Chinese adults. *J Epidemiol Community Health* **68**, 773–779.
34. Brayner B, Kaur G, Keske MA *et al.* (2021) Dietary patterns characterized by fat type in association with obesity and type 2 diabetes: a longitudinal study of UK Biobank participants. *J Nutr* **151**, 3570–3578.
35. Drake I, Sonestedt E, Ericson U *et al.* (2018) A Western dietary pattern is prospectively associated with cardiometabolic traits and incidence of the metabolic syndrome. *Br J Nutr* **119**, 1168–1176.
36. Duan M-J, Dekker LH, Carrero J-J *et al.* (2021) Blood lipids-related dietary patterns derived from reduced rank regression are associated with incident type 2 diabetes. *Clin Nutr* **40**, 4712–4719.
37. Heidemann C, Hoffmann K, Spranger J *et al.* (2005) A dietary pattern protective against type 2 diabetes in the European prospective investigation into cancer and nutrition (EPIC) – Potsdam study cohort. *Diabetologia* **48**, 1126–1134.
38. Imamura F, Lichtenstein AH, Dallal GE *et al.* (2009) Generalizability of dietary patterns associated with incidence of type 2 diabetes mellitus. *Am J Clin Nutr* **90**, 1075–1083.
39. Jacobs S, Kroeger J, Schulze MB *et al.* (2017) Dietary patterns derived by reduced rank regression are inversely associated with type 2 diabetes risk across 5 ethnic groups in the multiethnic cohort. *Curr Dev Nutr* **1**, e000620.
40. Johnson RK, Vanderlinden LA, DeFelice BC *et al.* (2020) Metabolomics-related nutrient patterns at seroconversion and risk of progression to type 1 diabetes. *Pediatr Diabetes* **21**, 1202–1209.
41. Lee HA, Son N, Lee WK *et al.* (2019) A diabetes-related dietary pattern is associated with incident diabetes in obese men in the Korean genome epidemiology study. *J Nutr* **149**, 323–329.
42. Liese AD, Weis KE, Schulz M *et al.* (2009) Food intake patterns associated with incident type 2 diabetes: the Insulin Resistance Atherosclerosis Study. *Diabetes Care* **32**, 263–268.
43. McNaughton SA, Mishra GD & Brunner EJ (2008) Dietary patterns, insulin resistance, and incidence of type 2 diabetes in the Whitehall II study. *Diabetes Care* **31**, 1343–1348.
44. Pastorino S, Richards M, Pierce M *et al.* (2016) A high-fat, high-glycaemic index, low-fibre dietary pattern is prospectively associated with type 2 diabetes in a British birth cohort. *Br J Nutr* **115**, 1632–1642.
45. Schulze MB, Hoffmann K, Manson JE *et al.* (2005) Dietary pattern, inflammation, and incidence of type 2 diabetes in women. *Am J Clin Nutr* **82**, 675–684.
46. Shin D, Lee KW & Song WO (2015) Dietary patterns during pregnancy are associated with risk of gestational diabetes mellitus. *Nutrients* **7**, 9369–9382.
47. Sluijs I, Beulens JW, Van Der ADL *et al.* (2013) Plasma uric acid is associated with increased risk of type 2 diabetes independent of diet and metabolic risk factors. *J Nutr* **143**, 80–85.
48. de ICkd (2014) Adherence to predefined dietary patterns and incident type 2 diabetes in European populations: EPIC-InterAct study. *Diabetologia* **57**, 321–333.
49. Song S & Lee JE (2018) Dietary patterns related to triglyceride and high-density lipoprotein cholesterol and the incidence of type 2 diabetes in Korean men and women. *Nutrients* **11**, 8.
50. Collins R (2012) What makes UK biobank special? *Lancet* **379**, 1173–1174.
51. Sudlow C, Gallacher J, Allen N *et al.* (2015) UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779.
52. Galante J, Adamska L, Young A *et al.* (2015) The acceptability of repeat internet-based hybrid diet assessment of previous 24-h dietary intake: administration of the Oxford WebQ in UK Biobank. *Br J Nutr* **115**, 681–686.
53. Greenwood DC, Hardie LJ, Frost GS *et al.* (2019) Validation of the Oxford WebQ online 24-h dietary questionnaire using biomarkers. *Am J Epidemiol* **188**, 1858–1867.
54. Liu B, Young H, Crowe FL *et al.* (2011) 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* **14**, 1998–2005.
55. Piernas C, Perez-Cornago A, Gao M *et al.* (2021) Describing a new food group classification system for UK Biobank: analysis of food groups and sources of

https://doi.org/10.1017/S0029665122001094 Published online by Cambridge University Press

macro- and micronutrients in 208,200 participants. *Eur J Nutr* **60**, 2879–2890.
56. Perez-Cornago A, Pollard Z, Young H *et al.* (2020) Description of the updated nutrition calculation of the Oxford WebQ questionnaire and comparison with the previous version among 207,144 participants in UK Biobank. *medRxiv*.
57. Macdonald IA (2016) A review of recent evidence relating to sugars, insulin resistance and diabetes. *Eur J Nutr* **55**, 17–23.
58. Neuenschwander M, Ballon A, Weber KS *et al.* (2019) Role of diet in type 2 diabetes incidence: umbrella review of meta-analyses of prospective observational studies. *Br Med J* **366**, l2368.
59. Thom G & Lean M (2017) Is there an optimal diet for weight management and metabolic health? *Gastroenterology* **152**, 1739–1751.
60. Reynolds AN, Akerman AP & Mann J (2020) Dietary fibre and whole grains in diabetes management: systematic review and meta-analyses. *PLoS Med* **17**, e1003053.
61. van Dam RM, Willett WC, Rimm EB *et al.* (2002) Dietary fat and meat intake in relation to risk of type 2 diabetes in men. *Diabetes Care* **25**, 417–424.
62. NatCen Social Research, MRC Elsie Widdowson Laboratory, University College London. Medical School. (2015). National Diet and Nutrition Survey Years 1–6, 2008/09–2013/14. [data collection]. 7th Edition. UK Data Service. SN: 6533, http://doi.org/105.255/UKDA-SN-6533-7.
63. Hooper L, Martin N, Jimoh OF *et al.* (2020) Reduction in saturated fat intake for cardiovascular disease. *Cochrane Database Syst Rev*, (8).
64. Te Morenga L, Mallard S & Mann J (2013) Dietary sugars and body weight: systematic review and meta-analyses of randomised controlled trials and cohort studies. *Br Med J* **346**, e7492.
65. Te Morenga LA, Howatson AJ, Jones RM *et al.* (2014) Dietary sugars and cardiometabolic risk: systematic review and meta-analyses of randomized controlled trials of the effects on blood pressure and lipids. *Am J Clin Nutr* **100**, 65–79.
66. Threapleton DE, Greenwood DC, Evans CE *et al.* (2013) Dietary fibre intake and risk of cardiovascular disease: systematic review and meta-analysis. *Br Med J* **347**, f6879.
67. Huisman MJ, Soedamah-Muthu SS, Vermeulen E *et al.* (2018) Does a high sugar high fat dietary pattern explain the unequal burden in prevalence of type 2 diabetes in a multi-ethnic population in The Netherlands? The HELIUS study. *Nutrients* **10**, 92.
68. Lamichhane AP, Liese AD, Urbina EM *et al.* (2014) Associations of dietary intake patterns identified using reduced rank regression with markers of arterial stiffness among youth with type 1 diabetes. *Eur J Clin Nutr* **68**, 1327–1333.
69. Yang Q, Zhang Z, Gregg EW *et al.* (2014) Added sugar intake and cardiovascular diseases mortality among US adults. *JAMA Intern Med* **174**, 516–524.
70. Zong G, Li Y, Wanders AJ *et al.* (2016) Intake of individual saturated fatty acids and risk of coronary heart disease in US men and women: two prospective longitudinal cohort studies. *Br Med J* **355**, i5796.
71. Siervo M, Lara J, Chowdhury S *et al.* (2015) Effects of the Dietary Approach to Stop Hypertension (DASH) diet on cardiovascular risk factors: a systematic review and meta-analysis. *Br J Nutr* **113**, 1–15.
72. Schwingshackl L, Bogensberger B & Hoffmann G (2018) Diet quality as assessed by the healthy eating index, alternate healthy eating index, dietary approaches to stop hypertension score, and health outcomes: an updated systematic review and meta-analysis of cohort studies. *J Acad Nutr Diet* **118**, 74–100. e111.
73. Forouhi NG, Krauss RM, Taubes G *et al.* (2018) Dietary fat and cardiometabolic health: evidence, controversies, and consensus for guidance. *Br Med J* **361**, k2139.
74. Forouhi NG, Misra A, Mohan V *et al.* (2018) Dietary and nutritional approaches for prevention and management of type 2 diabetes. *Br Med J* **361**, k2234.
75. Curtis PJ, Sampson M, Potter J *et al.* (2012) Chronic ingestion of flavan-3-ols and isoflavones improves insulin sensitivity and lipoprotein status and attenuates estimated 10-year CVD risk in medicated postmenopausal women with type 2 diabetes: a 1-year, double-blind, randomized, controlled trial. *Diabetes care* **35**, 226–232.
76. Imamura F, O'Connor L, Ye Z *et al.* (2015) Consumption of sugar sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: systematic review, meta-analysis, and estimation of population attributable fraction. *Br Med J* **351**, h3576.
77. Malik VS, Li Y, Pan A *et al.* (2019) Long-term consumption of sugar-sweetened and artificially sweetened beverages and risk of mortality in US adults. *Circulation* **139**, 2113–2125.
78. Kmietowicz Z (2018) Type 2 diabetes: sweetened drinks pose greater risk than other sugary foods. *Br Med J* **363**, 4943.
79. Bazzano LA, Li TY, Joshipura KJ *et al.* (2008) Intake of fruit, vegetables, and fruit juices and risk of diabetes in women. *Diabetes care* **31**, 1311–1317.
80. Schulze MB, Schulz M, Heidemann C *et al.* (2007) Fiber and magnesium intake and incidence of type 2 diabetes: a prospective study and meta-analysis. *Arch Intern Med* **167**, 956–965.
81. Choo VL, Viguiliouk E, Blanco Mejia S *et al.* (2018) Food sources of fructose-containing sugars and glycaemic control: systematic review and meta-analysis of controlled intervention studies. *Br Med J* **363**, k4644.
82. Fattore E, Botta F, Agostoni C *et al.* (2016) Effects of free sugars on blood pressure and lipids: a systematic review and meta-analysis of nutritional isoenergetic intervention trials. *Am J Clin Nutr* **105**, 42–56.
83. World Health Organization. Noncommunicable diseases (accessed 08/02/2022).
84. Carter JL, Lewington S, Piernas C *et al.* (2019) Reproducibility of dietary intakes of macronutrients, specific food groups, and dietary patterns in 211 050 adults in the UK Biobank study. *J Nutr Sci* **8**, e34.

Proceedings of the Nutrition Society

https://doi.org/10.1017/S0029665122001094 Published online by Cambridge University Press