International Journal of Epidemiology, 2018, 1603–1612
doi: 10.1093/ije/dyy030
Advance Access Publication Date: 2 April 2018
Original article

Cardiovascular Risks and Inequalities

Patterns of plant and animal protein intake are strongly associated with cardiovascular mortality: the Adventist Health Study-2 cohort

Marion Tharrey,[1,2] François Mariotti,[2] Andrew Mashchak,[1] Pierre Barbillon,[3] Maud Delattre[3] and Gary E Fraser[1]*

[1]Center for Nutrition, Healthy Lifestyle and Disease Prevention, Loma Linda University, School of Public Health, Loma Linda, CA, USA, [2]UMR PNCA, AgroParisTech, INRA, Université Paris-Saclay, 75005, Paris, France and [3]UMR MIA-Paris, INRA, Université Paris-Saclay, 75005, Paris, France

*Corresponding author. Adventist Health Studies, School of Public Health, 24951 N Circle Dr, NH 2033, Loma Linda University, Loma Linda, CA 92350, USA. E-mail: gfraser@llu.edu

Editorial decision 29 January 2018; Accepted 6 February 2018

Abstract

Background: Current evidence suggests that plant and animal proteins are intimately associated with specific large nutrient clusters that may explain part of their complex relation with cardiovascular health. We aimed at evaluating the association between specific patterns of protein intake with cardiovascular mortality.

Methods: We selected 81 337 men and women from the Adventist Health Study-2. Diet was assessed between 2002 and 2007, by using a validated food frequency questionnaire. Dietary patterns based on the participants' protein consumption were derived by factor analysis. Cox regression analysis was used to estimate multivariate-adjusted hazard ratios (HRs) adjusted for sociodemographic and lifestyle factors and dietary components.

Results: There were 2276 cardiovascular deaths during a mean follow-up time of 9.4 years. The HRs for cardiovascular mortality were 1.61 [98.75% confidence interval (CI), 1.12 2.32; *P*-trend $<$ 0.001] for the 'Meat' protein factor and 0.60 (98.75% CI, 0.42 0.86; *P*-trend $<$ 0.001) for the 'Nuts & Seeds' protein factor (highest vs lowest quintile of factor scores). No significant associations were found for the 'Grains', 'Processed Foods' and 'Legumes, Fruits & Vegetables' protein factors. Additional adjustments for the participants' vegetarian dietary pattern and nutrients related to cardiovascular disease outcomes did not change the results.

Conclusions: Associations between the 'Meat' and 'Nuts & Seeds' protein factors and cardiovascular outcomes were strong and could not be ascribed to other associated nutrients considered to be important for cardiovascular health. Healthy diets can be advocated based on protein sources, preferring low contributions of protein from meat and higher intakes of plant protein from nuts and seeds.

Key words: Nutritional epidemiology, factor analysis, cardiovascular disease, dietary protein, plant protein

Key Messages

- Plant and animal proteins are heterogeneously associated with CVD mortality.
- Protein-based factor analysis showed that a high contribution of protein from meat increased risk of CVD mortality, whereas a high contribution of protein from nuts and seeds is protective.
- These associations were not influenced by other characteristics of the diet, such as vegetarian dietary patterns or nutrients related to CVD outcomes.

Introduction

Recently, the eco-environmental sustainability of a diet containing high amounts of animal protein has been questioned, and it is thought that plant as an alternative to animal proteins will lead to more eco-friendly results.[1–4] Numerous observational and interventional studies have investigated how plant and animal protein may differentially affect CVD risk factors and mortality, given that cardiovascular diseases (CVDs) are the leading cause of death worldwide.[5] However, evidence for a beneficial effect of plant proteins is mixed.[6] Limited and inconsistent results highlight the probable role of confounding by non-protein dietary components, and possibly also relate to the association between dietary protein sources and diet quality.[6,7] Dietary proteins are not consumed in isolation but are embedded in complex food matrices as a part of the overall diet. In particular, each protein food group provides other specific non-protein compounds that can also affect cardiovascular health.[6] Thus, a simple analysis of dietary protein intake as from plant or animal sources may be too broad, and greater consideration of specific protein food sources and the background diet is required to accurately assess associations with CVD risk factors and mortality.[8]

Analysis of dietary patterns—using factor analysis—has proven effective to study the multidimensionality of diet and to give more insight into the relations between diet and disease.[9–11] This approach uses the correlations between food and nutrient intakes to derive non-correlated factors that describe general patterns that might be easier for the public to interpret and translate into relevant policy and guidance. Following the same approach, dietary protein patterns can be identified by analysing the intakes of protein from a variety of food sources. This approach overcomes the high correlation between protein groups and can identify the fundamental characteristics of a healthy protein pattern. The aim of this study was to examine the associations between patterns of protein intake and cardiovascular mortality in the Adventist Health Study-2 (AHS-2) cohort. We hypothesized that the dietary protein patterns that are identified may differentially affect cardiovascular health. We further investigated whether these associations were influenced by other general characteristics of the diet, including vegetarian dietary patterns and specific nutrients.

Methods

Study population

The Adventist Health Study 2 (AHS-2) is a population-based longitudinal study of more than 96 000 Seventh-day Adventist church men and women living in the USA and Canada, recruited between 2002 and 2007. The methods used for the cohort formation and its characteristics have been described elsewhere.[12] The AHS-2 was approved by the institutional review board of Loma Linda University and written consent was obtained from all participants at enrolment. Exclusion criteria for the present analyses were age < 25 years (n $=$ 226); estimated energy intake (not including write-in items) $<$ 500 kcal/d or $>$ 4500 kcal/d or improbable response patterns (e.g., identical responses to all questions on a page or more than 69 missing values in dietary data) (n $=$ 5840); body mass index (BMI) $< 14\,kg/m^2$ or $> 60\,kg/m^2$ ($n = 2539$); and self-reported history of cardiovascular events at baseline ($n = 6182$) (see flow diagram in Supplementary Figure 1, available as Supplementary data at *IJE* online). After these exclusions, 81 337 participants remained for analysis.

Mortality data

Deaths through 31 December 2013 were identified by biennial follow-up of participants and linkage with the National Death Index. The underlying cause of death was coded using the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10). CVD deaths were identified as those starting with the letter I.

Dietary assessment and covariable data

Usual dietary intake during the previous year was assessed through the baseline questionnaire using a quantitative food frequency questionnaire (FFQ). Detailed descriptions of the methods of dietary measurement using the questionnaire and its validation against six 24-h recalls have been described elsewhere.[13,14] Animal protein had de-attenuated correlations of 0.68 and 0.76 in Blacks and Whites, respectively. Plant protein had de-attenuated correlation of 0.57 in both races. A guided multiple-imputation approach was used to fill out missing data in the dietary variables.[15] Other sociodemographic and lifestyle factors were also assessed by the baseline questionnaire. The vegetarian status of participants was identified according to their reported intake of foods of animal origin and classified to five categories (vegan, lacto-ovo-vegetarian, pesco-vegetarian, semi-vegetarian, non-vegetarian).[16]

Identification of animal and plant protein categories

Classification of food items from the FFQ into animal and plant protein categories is described in Supplementary Methods 1, available as Supplementary data at *IJE* online. Animal protein groups were defined as coming from the following sources: red meat, processed meat, poultry, seafood, milk, yogurt, cheese, animal fats (to capture low amounts of proteins in cream and butter) and egg products. Plant proteins were defined as coming from: grains, soya, legumes, peanuts, tree nuts and seeds, potatoes, fruits, vegetables, other vegetables (e.g. condiments, spices etc.).[17] Protein contents of food items having a single source of protein were directly assigned to the corresponding protein group. Otherwise, representative recipes were developed and the amount of protein from each constituent ingredient was then assigned to the appropriate protein group. To evaluate the errors originating from the created recipes, we calculated the difference in the amount of protein between the original item and that calculated from the summed protein of the ingredients. We also calculated total protein intake before and after the breakdown of foods to their constituents.

Statistical analysis

Cox proportional hazard regressions were used in preliminary analyses to determine associations between protein intake coming from plant or animal sources and the risk of CVD mortality. Length of follow-up was used as the time variable, terminated either by death or by censoring. For these analyses, plant and animal protein were mutually adjusted. Hazard ratios (HR) and 95% CIs were calculated for 18-g increases of plant or animal protein intake, corresponding to ~1/4 of total protein intake. Plant and animal protein intakes were adjusted for total energy according to the residual method.[18]

Protein intake from each of the 18 protein food groups was expressed as the percentage of total protein intake (% kcal). Protein dietary patterns were then generated by factor analyses using the PROC FACTOR procedure in SAS. Factors were extracted using the principal component method, and an orthogonal transformation (Varimax rotation) was further applied to achieve a simpler structure with greater interpretability. To determine the number of factors to retain, we considered eigenvalues >1,[19] a breakpoint in the Scree test[20] and the interpretability of the factors.[21] For each food group, loadings for factors represented the correlation between the food groups and a factor. The protein dietary patterns were labelled according to food groups that made major contributions to the factor (absolute value of factor loading >0.20). Factor scores for each protein dietary pattern were calculated for each subject by summing the percentage of protein intake from each food group weighted by that food group's factor loading.[22] Participants were then grouped into quintiles of factor scores.

The specific effects of the identified protein factors on cause-specific CVD mortality [i.e. ischaemic heart disease) IHD and stroke separately] were identified using Cox proportional hazard regression analyses. Different models of increasing complexity were tested to investigate the effect of adjustment for additional potential confounding factors identified by the analyses and from the existing literature. Model 1 was adjusted for mean-centred age, sex, race and energy intake, BMI, physical activity, smoking status, alcohol consumption, education, personal income and marital status. Because the source of dietary protein is also related to vegetarian dietary patterns,[23] the second model was further adjusted for the traditional five vegetarian-spectrum dietary categories. In the third and fourth models, we replaced the vegetarian dietary pattern identifiers with other nutrients that are often related to CVD outcomes. Model 3 was adjusted for saturated fatty acids, unsaturated fatty acids,[24,25] fibre,[26,27] sodium,[28,29] vitamin B_6, B_{12}, folates[30,31] and antioxidants (vitamins A, C, E).[32] Model 4 was further adjusted for fat from meat products (fish excluded)[33] and fat from nuts.[34] Differences in the associations between factors and CVD mortality by age, sex, race and BMI were evaluated by testing possible interaction terms. The robustness of the a posteriori approach (using factors) was further checked by an a priori analysis, based on methods typically used to create a priori scores of adherence to a dietary profile.[35] Five scores were created to compare to the five dietary profiles that are the object of the main a posteriori approach here conducted. Details can be found in Supplementary Table 4, available as Supplementary data at *IJE* online. The Cox proportional hazards assumption was evaluated using tests and plots based on the Schoenfeld residuals. We tested for possible nonlinear relationships between protein factors and mortality using stepwise restricted cubic spline analysis.[36] All analyses were performed with the SAS statistical software package V. 9.4 for Windows (SAS Institute, Cary, NC, USA). A P-value <0.05 was considered to be statistically significant.

Results

Associations between plant and animal protein intake and CVD mortality

Among 81 337 participants followed for < 12 years (median follow up of 9.9 years), 2276 deaths were identified as due to CVDs. After controlling for several potential confounders, an 18-g increase in animal protein intake was significantly associated with a slightly higher risk of CVD mortality in models further controlling for the type of vegetarian diet or for a set of nutrients related to CVD risk (Table 1). No significant associations were found with plant protein intake.

Identification of dietary patterns

The difference between the total protein intake as calculated after breakdown using recipe ingredients and the original intake was 0.01 ± 0.15%. Factor analysis retained five main protein factors, which explained 47.2% of the total variance (Table 2). A Scree plot displaying the eigenvalues is shown in Supplementary Figure 2, available as Supplementary data at *IJE* online. As shown in Table 2, the factor labels—'Grains', 'Processed Foods', 'Meat', 'Legumes, Fruits & Vegetables' ('LFV') and 'Nuts & Seeds', respectively—were taken directly from foods with correspondingly heavy weightings. Participants with high factor scores for the 'Grains' protein factor had on average 43.8% of their protein intake coming from grains (see Supplementary Table 1, available as Supplementary data at *IJE* online). Factor 2 was labelled 'Processed Foods', since it was characterized by high loadings for proteins from cheese, animal fat, eggs, potatoes and milk, which came mostly from processed products in our database. The average percentages of protein intake by food group for the five protein factors are given in Supplementary Table 1, available as Supplementary data at *IJE* online.

Table 1. Multivariate-adjusted hazard ratios of CVD mortality for each 18-g increase (corresponding to one-quarter of total protein intake) in animal and plant protein intake[+] in 81 337 participants of the Adventist Health Study-2

		Animal protein		Plant protein	
		HR	(95% CI)	HR	(95% CI)
Deaths/ person-years	2276/ 767 487				
Model 1[a]		1.06	(0.99 1.14)	0.99	(0.94 1.06)
Model 2[b]		1.07	(1.00 1.15)	0.98	(0.94 1.05)
Model 3[c]		1.12	(1.05 1.19)	0.95	(0.89 1.02)

[+]Animal and plant protein intake were energy-adjusted using the residual method and in the same model.

[a]Adjusted on mean-centred age (years), sex (men, women), race (Black, White), energy intake (kcal/day), BMI (kg/m^2), physical activity (min/week), smoking status (current smoker, quit < 1 year, quit 1–4 years, quit 5–9 years, quit 10–19 years, quit 20–29 years, quit 30 years, never smoked), alcohol consumption (never, past, current), income ($\leq$ 10, 000, >10, 000–30, 000, $\geq$30, 000 USD per year), education ($\leq$ high school, some college, $\geq$ Bachelor's degree), marital status (single, divorced and widowed, married and common law).

[b]Model 1 further adjusted for the type of diet in the vegetarian spectrum (vegans, lacto-ovo-vegetarian, pesco-vegetarian, semi-vegetarian, non-vegetarian).

[c]Model 1 further adjusted on polyunsaturated fatty acids (PUFA), saturated fatty acids (SFA), sodium and vitamins A, C, E, B_6, B_9 and B_{12} (intakes of nutrients were energy-adjusted with the residual method).

Characteristics of population

Baseline characteristics of the study participants across quintiles of the identified protein factors are given in Table 3. For instance, subjects in the upper quintile for the 'Meat' or 'Processed Foods' protein factors tended to have greater BMIs and to be more physically inactive, and were more likely to be smokers and current alcohol consumers (Ps < 0.001) compared with those in the lower quintile, whereas opposite trends were found for the 'Grains', 'LFV' and 'Nuts & Seeds' protein factors. The vegetarian status of participants and intakes of various nutrients across protein factors are presented in Supplementary Tables 2 and 3, available as Supplementary data at *IJE* online.

Associations between dietary protein factors and CVD mortality

Table 4 shows HRs of CVD mortality according to quintiles of the identified protein factors. Restricted cubic spline analyses did not identify evidence of nonlinearity in any of these associations. Significant age-interactions were found for the 'Meat' and 'Nuts & Seeds' protein factors and for BMI. After multiple adjustments for lifestyle confounders (Model 1), subjects belonging to the fifth quintile of the 'Meat' protein factor had a 61% higher risk of CVD death compared with those in the first quintile. By contrast, participants in the high quintile of 'Nuts & Seeds' protein factor had a 40% lower CVD mortality risk. No significant associations were found for the 'Grains', 'Processed Foods' and 'LFV' protein factors. After additional adjustment on the five vegetarian categories, or for other nutrients that are related to CVD outcomes, the estimates changed very little (Models 2 and 3). Further

Table 2. Factor loadings of the 18 protein food groups identified by factor analysis in 81 337 participants of the Adventist Health Study-2[a]

	Factor loadings				
	Factor1 (Grains)	Factor2 (Processed Foods)	Factor3 (Meat)	Factor4 (LFV)	Factor5 (Nuts & Seeds)
Grains	0.77	–	−0.32	–	–
Other vegetables	0.63	–	–	–	–
Seafood	−0.46	–	0.33	–	0.35
Yogurt	−0.47	–	−0.26	–	–
Cheese	–	0.64	–	–	–
Animal fat	–	0.57	–	–	–
Eggs	–	0.48	–	–	−0.24
Potatoes	0.36	0.43	–	0.40	–
Milk	–	0.40	–	−0.37	–
Soya	–	−0.53	−0.40	–	–
Red meat	–	–	0.73	–	–
Processed meat	–	–	0.62	–	–
Poultry	−0.34	–	0.57	–	−0.27
Vegetables	–	–	–	0.75	–
Legumes	–	–	–	0.59	–
Fruits	–	–	–	0.56	–
Peanuts	–	–	–	–	0.70
Treenuts & seeds	–	–	–	–	0.64
Variance explained, %	16.7	9.7	7.6	7.1	5.7

LFV, Legumes, Fruits & Vegetables.
[a]Positive loadings <0.20 and negative loadings >-0.20 were omitted for simplicity.

adjustment on fat from meat products and fat from nuts somewhat attenuated the associations for the 'Meat' factor, but left the nut protein result unchanged. The analysis based on the a priori method (using scores) showed good agreement with the a posteriori approach (using factors; see Supplementary Table 4, available as Supplementary data at *IJE* online). The CVD deaths identified in the population consisted mainly of ischaemic heart diseases (37.8%), other forms of heart disease (24.2%) and cerebrovascular diseases (23.7%). Sensitivity analysis showed clear trends in the expected directions when investigating associations between the factors and IHD and stroke separately, although these were not statistically significant (data not shown).

Since significant interaction terms were found between 'Meat' and 'Nuts & Seeds' factors and age, HRs across age categories were estimated (Figure 1). Strong associations between these protein factors and CVD death were found among young adults aged 25–44 (2-fold higher risk for the 'Meat' protein factor and almost 3-fold lower risk for the 'Nuts & Seeds' protein factor). However, the strength of these associations decreased with increasing age and were no longer apparent by age 80 and above.

Discussion

In the present study, a thorough analysis of specific protein dietary factors, derived by factor analysis, disclosed clearly contrasting associations of specific animal and plant protein factors on CVD outcomes. Our results show that high scores on the 'Meat' protein factor are associated with increased risk of CVD mortality, whereas high scores on the 'Nuts & Seeds' protein factor are associated with lower risk. When considering protein intake only divided into plant and animal sources, we found a weak positive association between animal protein intake and CVD mortality and no association for the plant proteins. These results emphasize that general statements about plant or animal protein may lack specificity, and that greater consideration of specific protein food sources and patterns is required, as had been proposed by some authors.[6] Specific characteristics and lifestyles of the population, different types of dietary substitution and the background diets have made it difficult to clearly ascribe benefits to the type and source of protein and likely contributed to the lack of conclusive evidence relating animal and plant protein consumption to CVD risk.[8,37–40]

Our results are consistent with other studies that reported either an increased risk of CVD mortality

Table 3. Baseline characteristics among 81 337 Adventist Health Study-2 participants by quintile (Q) of the protein factors[a]

Characteristic	Factor 1 (Grains)			Factor 2 (Processed Foods)			Factor 3 (Meat)			Factor 4 (LFV)			Factor 5 (Nuts & Seeds)		
	Q1	Q3	Q5	Q1	Q3	Q5	Q1	Q3	Q5	Q1	Q3	Q5	Q1	Q3	Q5
Mean (SD) age, years	55.9 (13.8)	57.2 (14.1)	57.8 (14.3)	55.9 (13.5)	57.0 (14.2)	57.8 (14.5)	56.7 (14.4)	58.5 (14.1)	53.8 (13)	55.5 (14.3)	57 (14.1)	58.4 (13.8)	53.5 (13.4)	56.6 (14.0)	60.8 (14.0)
Sex, *n* women (%)	12 122 (74.5)	10 514 (64.6)	9866 (60.7)	10 978 (67.5)	10 668 (65.6)	10 569 (65)	11 645 (71.6)	10 476 (64.4)	10 121 (62.2)	10 096 (62.1)	10 686 (65.7)	11 516 (70.8)	11 187 (68.8)	10 929 (67.2)	9993 (61.4)
Race, *n* Black (%)	6073 (37.3)	3972 (24.4)	2846 (17.5)	6569 (40.4)	4324 (26.6)	1917 (11.8)	3137 (19.3)	3564 (21.9)	6046 (37.2)	4045 (24.9)	3989 (24.5)	4971 (30.6)	7840 (48.2)	3533 (21.7)	2202 (13.5)
Mean (SD) BMI, kg/m^2	27.6 (5.9)	27.1 (5.8)	26.4 (5.8)	26.0 (5.3)	27.1 (5.9)	28.0 (6.1)	26.6 (5.5)	26.2 (5.4)	29.1 (6.5)	27.9 (6.2)	27.1 (5.8)	26.1 (5.5)	28.4 (6.3)	27.1 (5.7)	25.5 (5.1)
Smoking status, *n* never smokers (%)	13 196 (81.1)	13 119 (80.6)	13 542 (83.2)	13 542 (83.2)	13 082 (80.4)	12 996 (79.9)	14 077 (86.5)	13 590 (83.5)	11 369 (69.9)	12 554 (77.2)	13 332 (82)	13 527 (83.2)	12 734 (78.3)	13 266 (81.5)	13 456 (82.7)
Alcohol status, *n* never drinkers (%)	9763 (60)	10 673 (65.6)	11 249 (69.2)	10 976 (67.5)	10 476 (64.4)	10 268 (63.1)	11 690 (71.9)	11 314 (69.5)	7812 (48.0)	9653 (59.3)	10 644 (65.4)	11 341 (69.7)	9730 (59.8)	10 540 (64.8)	11 347 (69.8)
Mean (SD) physical activity, (min/week)	13 098 (80.5)	13 094 (80.5)	12 741 (78.3)	13 594 (83.6)	13 060 (80.3)	12 222 (75.1)	13 329 (81.9)	13 390 (82.3)	12 065 (74.2)	12 255 (75.3)	13 190 (81.1)	13 286 (81.7)	12 337 (75.8)	13 087 (80.4)	13 267 (81.6)
Personal income, *n* (%) USD/year:															
≤10,000	2944 (18.1)	3175 (19.5)	4122 (25.3)	3220 (19.8)	3337 (20.5)	3475 (21.4)	3235 (19.9)	3519 (21.6)	3176 (19.5)	3387 (20.8)	3337 (20.5)	3669 (22.6)	3334 (20.5)	3447 (21.2)	3332 (20.5)
>10,000–30,000	5945 (36.5)	6182 (38)	6395 (39.3)	5925 (36.4)	6246 (38.4)	6204 (38.1)	5762 (35.4)	6290 (38.7)	6151 (37.8)	6209 (38.2)	6012 (37.0)	6290 (38.7)	6072 (37.3)	6033 (37.1)	6393 (39.3)
>30000	7379 (45.4)	6911 (42.5)	5750 (35.4)	7123 (43.8)	6685 (41.1)	6588 (40.5)	7271 (44.7)	6460 (39.7)	6940 (42.7)	6672 (41)	6919 (42.5)	6308 (38.8)	6862 (42.2)	6789 (41.7)	6543 (40.2)
Education, *n* (%):															
High school or less	3358 (21)	3115 (19.4)	3615 (22.5)	3006 (18.7)	3356 (20.9)	3540 (22.0)	2509 (15.6)	3171 (19.7)	4165 (25.9)	3649 (22.7)	3160 (19.7)	3483 (21.7)	3889 (24.3)	3214 (20.0)	3071 (19.1)
Some college	6489 (40.5)	6357 (39.6)	6304 (39.2)	6169 (38.4)	6412 (39.9)	6569 (40.8)	5978 (37.2)	6310 (39.3)	6991 (43.5)	6445 (40.1)	6277 (39)	6410 (40)	6601 (41.2)	6406 (39.9)	6018 (37.4)
Bachelor's degree or higher	6164 (38.5)	6584 (41)	6178 (38.4)	6881 (42.9)	6294 (39.2)	5993 (37.2)	7585 (47.2)	6589 (41)	4912 (30.6)	5997 (37.3)	6644 (41.3)	6133 (38.3)	5543 (34.6)	6453 (40.2)	6984 (43.5)
Marital status, *n* currently married (%)	1351 (8.3)	990 (6.1)	916 (5.6)	1330 (8.2)	1008 (6.2)	874 (5.4)	1053 (6.5)	836 (5.1)	1426 (8.8)	1270 (7.8)	985 (6.1)	1068 (6.6)	1508 (9.3)	946 (5.8)	823 (5.1)

LFV, egumes, Fruits & Vegetables.

[a]Categories used in the models were as follows: sex (men, women), race (Black, White), smoking status (current smoker, quit < 1 year, quit 1–4 years, quit 5–9 years, quit 10–19 years, quit 20–29 years, quit 30 years, never smoked), alcohol status (never, past, current), income (≤10,000, >10, 000–30 000, ≥30,000 USD per year), education (≤ high school, some college, ≥ Bachelor's degree), marital status (single, divorced and widowed, married and common law). Tests for trend among protein factors were all < 0.001 (except for age and sex for Factor 2, not significant), and obtained by analysis of variance (for age, BMI and physical activity), binomial logistic regression (for sex and race) or multinomial logistic regression (for smoking status, alcohol status, personal income, education and marital status). All models were adjusted on age, sex and race (as appropriate).

Table 4. Multivariate-adjusted hazard ratio of CVD mortality by quintile (Q) of identified protein dietary factors in 81 337 participants of the Adventist Health Study-2

	Q1	Q2	Q3	Q4	Q5	*P-trend*
Factor 1 (Grains)						
Deaths/person-years	435/151 978	433/153 063	469/153 566	447/154 380	492/154 710	
Model 1 HR (98.75% CI)[a]	1.00 (ref.)	0.91 (0.77 1.08)	0.98 (0.77 1.24)	0.86 (0.72 1.03)	0.89 (0.74 1.08)	*0.067*
Model 2 HR (98.75% CI)[b]	1.00 (ref.)	0.90 (0.76 1.07)	0.97 (0.76 1.24)	0.82 (0.68 0.99)	0.88 (0.74 1.09)	*0.077*
Model 3 HR (98.75% CI)[c]	1.00 (ref.)	0.90 (0.75 1.07)	0.99 (0.78 1.24)	0.84 (0.70 1.01)	0.90 (0.74 1.10)	*0.101*
Model 4 HR (98.75% CI)[d]	1.00 (ref.)	0.91 (0.76 1.08)	1.00 (0.80 1.26)	0.86 (0.71 1.03)	0.93 (0.76 1.13)	*0.303*
Factor 2 (Processed Foods)						
Deaths/person-years	375/153 564	418/153 677	466/152 920	497/153 835	520/153 701	
Model 1 HR (98.75% CI)[a]	1.00 (ref.)	1.00 (0.84 1.19)	1.01 (0.84 1.22)	1.03 (0.86 1.23)	0.98 (0.81 1.19)	*0.722*
Model 2 HR (98.75% CI)[b]	1.00 (ref.)	0.98 (0.82 1.18)	1.02 (0.84 1.23)	1.06 (0.88 1.28)	0.99 (0.81 1.22)	*0.628*
Model 3 HR (98.75% CI)[c]	1.00 (ref.)	1.01 (0.84 1.21)	1.06 (0.88 1.27)	1.14 (0.94 1.39)	1.10 (0.88 1.39)	*0.115*
Model 4 HR (98.75% CI)[d]	1.00 (ref.)	1.01 (0.84 1.22)	1.06 (0.88 1.28)	1.15 (0.94 1.40)	1.12 (0.90 1.41)	*0.073*
Factor 3 (Meat)						
Deaths/person-years	473/ 153 868	474/153 843	477/ 154 046	484/152 869	368/ 153 059	
Model 1 HR (98.75% CI)[a]	1.00 (ref.)	0.97 (0.68 1.37)	1.08 (0.70 1.65)	1.09 (0.78 1.52)	1.61 (1.12 2.32)	*<0.001*
Model 2 HR (98.75% CI)[b]	1.00 (ref.)	1.07 (0.75 1.54)	1.09 (0.71 1.67)	1.22 (0.85 1.76)	1.64 (1.13 2.38)	*<0.001*
Model 3 HR (98.75% CI)[c]	1.00 (ref.)	1.07 (0.75 1.53)	1.10 (0.71 1.69)	1.22 (0.86 1.73)	1.67 (1.16 2.41)	*<0.001*
Model 4 HR (98.75% CI)[d]	1.00 (ref.)	1.05 (0.73 1.51)	1.06 (0.69 1.64)	1.15 (0.80 1.64)	1.46 (0.98 2.18)	*0.012*
Factor 4 (LFV)						
Deaths/person-years	460/153 884	407/154 438	438/153 637	469/153 260	502/152 478	
Model 1 HR (98.75% CI)[a]	1.00 (ref.)	0.93 (0.78 1.10)	0.97 (0.81 1.16)	1.01 (0.85 1.19)	1.06 (0.89 1.26)	*0.366*
Model 2 HR (98.75% CI)[b]	1.00 (ref.)	0.92 (0.78 1.10)	0.97 (0.81 1.16)	1.01 (0.85 1.21)	1.05 (0.87 1.26)	*0.491*
Model 3 HR (98.75% CI)[c]	1.00 (ref.)	0.91 (0.76 1.08)	0.95 (0.78 1.14)	0.97 (0.80 1.17)	1.00 (0.81 1.24)	*0.916*
Model 4 HR (98.75% CI)[d]	1.00 (ref.)	0.92 (0.77 1.10)	0.96 (0.80 1.16)	0.99 (0.82 1.20)	1.04 (0.84 1.28)	*0.560*
Factor 5 (Nuts & Seeds)						
Deaths/person-years	328/152 922	412/153 778	459/153 357	484/153 936	593/153 704	
Model 1 HR (98.75% CI)[a]	1.00 (ref.)	0.73 (0.52 1.02)	0.86 (0.59 1.25)	0.69 (0.49 0.97)	0.60 (0.42 0.86)	*<0.001*
Model 2 HR (98.75% CI)[b]	1.00 (ref.)	0.90 (0.66 1.23)	0.85 (0.59 1.25)	0.73 (0.52 1.02)	0.60 (0.41 0.86)	*<0.001*
Model 3 HR (98.75% CI)[c]	1.00 (ref.)	0.90 (0.66 1.23)	0.86 (0.59 1.25)	0.73 (0.52 1.03)	0.59 (0.41 0.85)	*<0.001*
Model 4 HR (98.75% CI)[d]	1.00 (ref.)	0.88 (0.64 1.20)	0.84 (0.57 1.22)	0.70 (0.49 0.99)	0.56 (0.38 0.81)	*<0.001*

LFV, Legumes, Fruits & Vegetables.

All results are shown as HR and 98.75% CIs to account for multiple comparisons (Bonferroni correction: significance criterion 0.05/4 = 0.0125 for each quintile).

[a]Adjusted on mean-centred age (years), sex (men, women), race (Black, White), energy intake (kcal/day), BMI (kg/m^2), physical activity (min/week), smoking status (current smoker, quit <1 year, quit 1–4 years, quit 5–9 years, quit 10–19 years, quit 20–29 years, quit 30 years, never smoked), alcohol consumption (never, past, current), income (≤ 10,000, >10,000–30,000, ≥30,000 USD per year), education (≤ high school, some college, ≥ Bachelor's degree), marital status (single, divorced and widowed, married and common law).

[b]Model 1 further adjusted on type of diet in the vegetarian spectrum (vegans, lacto-ovo-vegetarian, pesco-vegetarian, semi-vegetarian, non-vegetarian).

[c]Model 1 further adjusted on polyunsaturated fatty acids (PUFA), saturated fatty acids (SFA), sodium, fiber, and vitamins A, C, E, B_6, B_9 and B_{12} (intakes of nutrients were energy-adjusted with the residual method).

[d]Model 3 further adjusted on fat from meat products (fish excluded) and fat from nuts (intakes of fats were energy-adjusted with the residual method).

associated with red and processed meat consumption;[8,41,42] or a protective effect of nuts and seeds on CVD risk.[43,44] However, in most cases these studies looked at a single food group (e.g. total meat) or energy-based patterns (e.g. Western diet), and did not explicitly consider protein content and sources. Here we focused on factor analyses to find protein variables that in our data were approximately independent. Without strong a priori hypotheses that focused on proteins from a particular source, this was useful to further define the protein structure of our data. In fact these factors clearly weighted heavily on easily recognized dietary sources. Furthermore, we found similar results in relation to CVD risk when we compared with an a priori approach based on scores, which supports the validity of our approach. Previous studies have also investigated different dietary patterns that could be interpreted by considering their protein sources. For instance, the Mediterranean-style dietary pattern, as with other 'pro-vegetarian' patterns, is characterized by a substantial intake of proteins from plant sources which could in part

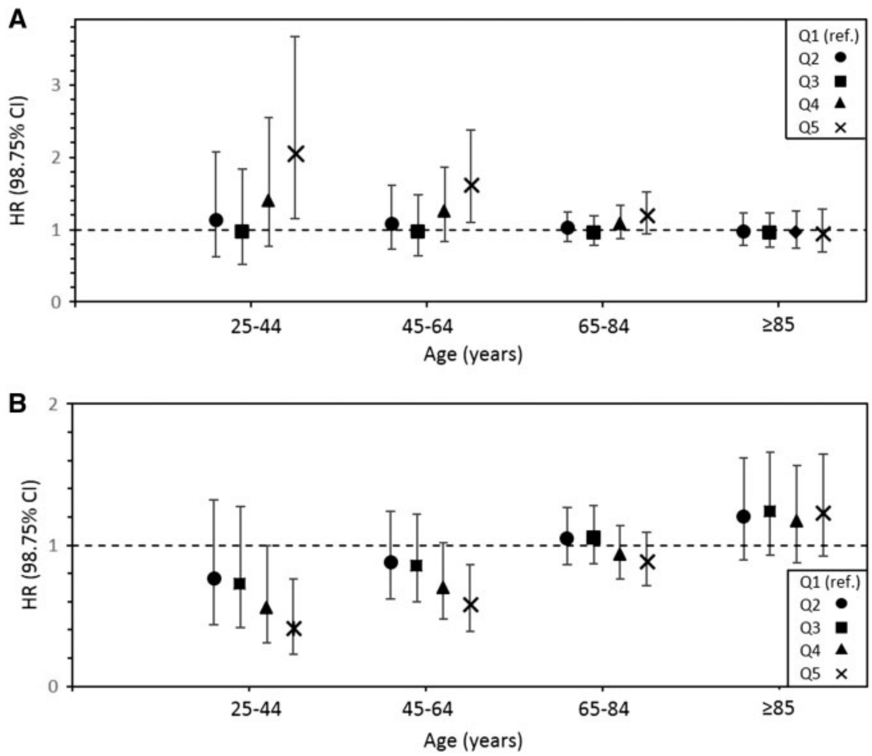

Figure 1. Multivariate-adjusted hazard ratio (HR) of CVD mortality (Model 2) by quintile of the 'Meat' (A) and 'Nuts & Seeds' (B) protein factors and by age categories in 81 337 participants of the Adventist Health Study-2. Significant age interactions were found for the 'Meat' and 'Nuts & Seeds' protein factors ($P = 0.003$ and $P < 0.001$, respectively). HRs were estimated at the mean age of each age category. All confidence intervals were calculated to reflect the 98.75% confidence interval range (following Bonferroni corrections with significance criterion being 0.05/4, i.e. 0.0125 for each quintile).

account for the favourable CVD and type 2 diabetes outcomes associated with this diet.[45–47] However, to our knowledge, protein-based factor analysis is novel and offers a specific insight into the underlying eating patterns that combine a variety of protein food sources. Further, it may offer practical conclusions about the likely value of selecting proteins from different food sources. As each factor is independent one from the other by construction, the models avoid the multicollinearity that is found when analysing protein groups individually. Thus the association of each factor with CVD mortality is more easily interpretable. The interpretation of the factors may still, however, be limited by the fact that they are associated in a complex way with vegetarian diets or other nutrients consumed along with these protein sources (i.e. certain protein types may act as markers of healthy diets). Nevertheless, after controlling for the five vegetarian-spectrum dietary categories and several nutrients (including dietary fatty acids) thought to be related to CVD outcomes, significant associations persisted between the protein dietary factors and CVD mortality. This strengthens the idea that these analyses have captured unique aspects of protein-based patterns, which are attributable to the protein per se. In particular, the 'Nuts & Seeds' protein factor appeared to be associated with the quality of the diet even across different levels of plant-based dietary patterns in this population, suggesting that focusing on more specific plant protein-based diets may improve the ability of dietary recommendations to prevent CVD.

Based on our findings, its appears that choices of dietary protein that favour nuts and seeds compared with meat could be an effective means of reducing the risk of CVD deaths under causal hypotheses, and could be considered for future guidance in dietary public health recommendations. Among potential mechanisms, the amino acid composition of specific foods may affect CVD health by their specific physiological effects. Previous studies have reported an inverse relation of dietary glutamic acid intake with blood pressure. This is an amino acid predominant in plant protein.[48,49] Nuts also provide a high content of L-arginine, a precursor of nitric oxide, known to play fundamental role in vascular health.[50,51] In addition, glycine may be an important contributor to a direct relationship of meat products with blood pressure.[52]

We found that the significant associations between the protein dietary factors and CVD mortality tended to weaken with age. Among possible explanations, it is likely that participants who have reached an advanced age without experiencing any previous cardiovascular events may possess constitutional advantages, and modifiable CVD risk factors such as diet, then, have proportionately less influence on CVD health outcomes. It is significant for public health that the protein factors appear to show very strong associations with premature CVD risk, as early deaths represent large losses of productive years of life.

Our study has potential limitations. First, reporting bias in self-reported dietary assessment and other lifestyle-related data is inevitable, but as these biases are expected to be non-differential they would usually bias toward the null. More importantly, diet was assessed only at baseline and may have changed over time. Second, missing responses in the FFQ were filled by multiple imputation. Nonetheless, multiple imputation does not create bias under the assumption that errors are missing at random. This assumption was approximately satisfied here by the use of guided multiple imputation.[15,53] Finally, although appropriate adjustments for confounding factors were performed, unknown and unmeasured confounders are always possible. Notably, the contribution of other plant food components intimately related to protein intake, such as phytochemicals, may also be relevant.[54] Strengths of this study include the large number of participants and a relatively long follow-up period allowing the accumulation of many cardiovascular deaths, and also the diversity of dietary habits in this population leading to the identification of a wide range of protein food sources.

Conclusion

Our study appears to identify heterogeneous associations of certain plant and animal proteins with CVD risk. Strong associations were found between CVD outcomes and the animal protein factor that weighted heavily on meat products, whereas a specific plant protein factor weighting on nuts and seeds was associated with a lower risk of CVD mortality. These relationships were mostly apparent before old age, thus impacting on premature CVD death. This strengthens the idea that protein sources may be key components of diet quality, possibly largely independent of other confounders, including vegetarian diet categories. Associations of these dietary protein factors with CVD death cannot be easily ascribed to their correlations with other nutrients considered important for cardiovascular health. Our results suggest that healthy choices can be advocated based on protein sources, specifically preferring diets low in meat intake and with a higher intake of plant proteins from nuts and seeds.

Supplementary data

Supplementary data are available at *IJE* online

Funding

This work was supported by AgroParisTech Foundation; the ADEPRINA/AgroParisTech., and the U.S. National Institutes of Health grant #5U01CA152939. M.T. received predoctoral fellowship funding from these two agencies.

Conflict of interest: No authors had any conflict of interest to declare.

References

1. Aiking H. Protein production: planet, profit, plus people? *Am J Clin Nutr* 2014;**100**(**Suppl 1**):483–89S.
2. Pimentel D, Pimentel M. Sustainability of meat-based and plant-based diets and the environment. *Am J Clin Nutr* 2003;**78**:660S–63.
3. Scarborough P, Appleby PN, Mizdrak A *et al*. Dietary greenhouse gas emissions of meat-eaters, fish-eaters, vegetarians and vegans in the UK. *Clim Change* 2014;**125**:179–92.
4. Nelson ME, Hamm MW, Hu FB, Abrams SA, Griffin TS. Alignment of healthy dietary patterns and environmental sustainability: a systematic review. *Adv Nutr* 2016;**7**:1005–25.
5. World Health Organization. *Cardiovascular Diseases (CVDs) Fact Sheet No 317*. http://www.who.int/mediacentre/factsheets/fs317/en/ (18 February 2018, date last accessed).
6. Richter CK, Skulas-Ray AC, Champagne CM, Kris-Etherton PM. Plant protein and animal proteins: do they differentially affect cardiovascular disease risk? *Adv Nutr* 2015;**6**:712–28.
7. Dietary Guidelines Advisory Committee. *Report of the Dietary Guidelines Advisory Committee on the Dietary Guidelines for Americans, 2010, to the Secretary of Agriculture and the Secretary of Health and Human Services*. Washington, DC: U.S. Department of Agriculture, Agricultural Research Service, 2010.
8. Song M, Fung TT, Hu FB *et al*. Association of animal and plant protein intake with all-cause and cause-specific mortality. *JAMA Intern Med* 2016;**176**:1453–63.
9. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* 2002;**13**:3–9.
10. Michels KB, Schulze MB. Can dietary patterns help us detect diet-disease associations? *Nutr Res Rev* 2005;**18**:241–48.
11. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: a review. *Nutr Rev* 2004;**62**:177–203.
12. Butler TL, Fraser GE, Beeson WL *et al*. Cohort profile: the Adventist Health Study-2 (AHS-2). *Int J Epidemiol* 2008;**37**:260–65.
13. Jaceldo-Siegl K, Fan J, Sabaté J *et al*. Race-specific validation of food intake obtained from a comprehensive FFQ: the Adventist Health Study-2. *Public Health Nutr* 2011;**14**:1988–97.
14. Jaceldo-Siegl K, Knutsen SF, Sabaté J *et al*. Validation of nutrient intake using an FFQ and repeated 24 h recalls in black and white subjects of the Adventist Health Study-2 (AHS-2). *Public Health Nutr* 2010;**13**:812–19.
15. Fraser G, Yan R. Guided multiple imputation of missing data: using a subsample to strengthen the missing-at-random assumption. *Epidemiology* 2007;**18**:246–52.
16. Orlich MJ, Singh PN, Sabaté J *et al*. Vegetarian dietary patterns and mortality in Adventist Health Study 2. *JAMA Intern Med* 2013;**173**:1230–38.
17. Camilleri GM, Verger EO, Huneau J-F *et al*. Plant and animal protein intakes are differently associated with nutrient adequacy of the diet of French adults. *J Nutr* 2013;**143**:1466–73.

18. Thiébaut A, Kesse E, Com-Nougué C, Clavel-Chapelon F, Bénichou J. Adjustment for energy intake in the assessment of dietary risk factors. *Rev Épidémiol Santé Publique* 2004;**52**:539–57.
19. Kaiser HF. The application of electronic computers to factor analysis. *Educ Psychol Meas* 1960;**20**:141–51.
20. Cattell RB. The scree test for the number of factors. *Multivariate Behav Res* 1966;**1**:245–76.
21. Kline P. *An Easy Guide to Factor Analysis*. New York, NY: Routledge, 1994.
22. Kim J-O, Mueller CW. *Factor Analysis: Statistical Methods and Practical Issues*. Beverly Hills, CA, and London: SAGE Publications, 1978.
23. Rizzo NS, Jaceldo-Siegl K, Sabate J, Fraser GE. Nutrient profiles of vegetarian and nonvegetarian dietary patterns. *J Acad Nutr Diet* 2013;**113**:1610–19.
24. Hooper L, Martin N, Abdelhamid A *et al*. Reduction in saturated fat intake for cardiovascular disease. *Cochrane Database Syst Rev* 2015;**6**:CD011737.
25. Siri-Tarino PW, Sun Q, Hu FB, Krauss RM. Saturated fatty acids and risk of coronary heart disease: modulation by replacement nutrients. *Curr Atheroscler Rep* 2010;**12**:384–90.
26. Park Y, Subar AF, Hollenbeck A, Schatzkin A. Dietary fiber intake and mortality in the NIH-AARP Diet and Health Study. *Arch Intern Med* 2011;**171**:1061–68.
27. Grooms KN, Ommerborn MJ, Pham DQ, Djoussé L, Clark CR. Dietary fiber intake and cardiometabolic risks among US adults, NHANES 1999–2010. *Am J Med* 2013;**126**:1059.
28. Tuomilehto J, Jousilahti P, Rastenyte D *et al*. Urinary sodium excretion and cardiovascular mortality in Finland: a prospective study. *Lancet* 2001;**357**:848–45.
29. Penz ED, Joffres MR, Campbell NR. Reducing dietary sodium and decreases in cardiovascular disease in Canada. *Can J Cardiol* 2008;**24**:497–501.
30. Saposnik G, Ray JG, Sheridan P, McQueen M, Lonn E. Homocysteine-lowering therapy and stroke risk, severity, and disability: additional findings from the HOPE 2 trial. *Stroke* 2009;**40**:1365–72.
31. Cui R, Iso H, Date C, Kikuchi S, Tamakoshi A. Dietary folate and vitamin B6 and B12 intake in relation to mortality from cardiovascular diseases: Japan collaborative cohort study. *Stroke* 2010;**41**:1285–89.
32. Jha P, Flather M, Lonn E, Farkouh M, Yusuf S. The antioxidant vitamins and cardiovascular disease: a critical review of epidemiologic and clinical trial Data. *Ann Intern Med* 1995;**123**:860–72.
33. Erlinger TP, Appel L. *The Relationship Between Meat Intake and Cardiovascular Disease*. CLF White Paper. Baltimore, MD: Johns Hopkins Center for a Livable Future, 2003.
34. Fraser GE, Sabaté J, Beeson WL, Strahan TM. A possible protective effect of nut consumption on risk of coronary heart disease. The Adventist Health Study. *Arch Intern Med* 1992;**152**:1416–24.
35. Trichopoulou A, Costacou T, Bamia C, Trichopoulos D. Adherence to a Mediterranean diet and survival in a Greek population. *N Engl J Med* 2003;**348**:2599–608.
36. Desquilbet L, Mariotti F. Dose-response analyses using restricted cubic spline functions in public health research. *Stat Med* 2010; **29**:1037–57.
37. Iso H, Sato S, Kitamura A, Naito Y, Shimamoto T, Komachi Y. Fat and protein intakes and risk of intraparenchymal hemorrhage among middle-aged Japanese. *Am J Epidemiol* 2003;**157**:32–39.
38. Sauvaget C, Nagano J, Hayashi M, Yamada M. Animal protein, animal fat, and cholesterol intakes and risk of cerebral infarction mortality in the adult health study. *Stroke J Cereb* 2004;**35**: 1531–37.
39. Appel LJ, Sacks FM, Carey VJ *et al*. Effects of protein, monounsaturated fat, and carbohydrate intake on blood pressure and serum lipids, results of the OmniHeart randomized trial. *JAMA* 2005;**294**:2455–64.
40. Li S, Flint A, Pai JK *et al*. Low carbohydrate diet from plant or animal sources and mortality among myocardial infarction survivors. *J Am Heart Assoc* 2014;**3**:e001169.
41. Pan A, Sun Q, Bernstein AM *et al*. Red meat consumption and mortality, results from 2 prospective cohort studies. *Arch Intern Med* 2012;**172**:555–63.
42. Abete I, Romaguera D, Vieira AR, Lopez de Munain A, Norat T. Association between total, processed, red and white meat consumption and all-cause, CVD and IHD mortality, a meta-analysis of cohort studies. *Br J Nutr* 2014;**112**:762–75.
43. Mayhew AJ, de Souza RJ, Meyre D, Anand SS, Mente A. A systematic review and meta-analysis of nut consumption and incident risk of CVD and all-cause mortality. *Br J Nutr* 2016;**115**: 212–25.
44. Grosso G, Yang J, Marventano S, Micek A, Galvano F, Kales SN. Nut consumption on all-cause, cardiovascular, and cancer mortality risk, a systematic review and meta-analysis of epidemiologic studies. *Am J Clin Nutr* 2015;**101**:783–93.
45. Martinez-Gonzalez MA, Sanchez-Tainta A, Corella D *et al*. A provegetarian food pattern and reduction in total mortality in the Prevencion con Dieta Mediterranea (PREDIMED) study. *Am J Clin Nutr* 2014;**100**:320–28S.
46. Martínez-González MÁ, Ruiz-Canela M, Hruby A, Liang L, Trichopoulou A, Hu FB. Intervention trials with the mediterranean diet in cardiovascular prevention: understanding potential mechanisms through metabolomic profiling. *J Nutr* 2015;**146**: 913–19S.
47. Satija A, Bhupathiraju SN, Rimm EB *et al*. Plant-based dietary patterns and incidence of type 2 diabetes in US men and women: results from three prospective cohort studies. *PLoS Med* 2016; **13**:e1002039.
48. Elliott P, Stamler J, Dyer AR *et al*. Association between protein intake and blood pressure, the INTERMAP Study. *Arch Intern Med* 2006;**166**:79–87.
49. Jennings A, MacGregor A, Welch A, Chowienczyk P, Spector T, Cassidy A. Amino acid intakes are inversely associated with arterial stiffness and central blood pressure in women. *J Nutr* 2015;**145**:2130–38.
50. Ros E. Health benefits of nut consumption. *Nutrients* 2010;**2**: 652–82.
51. Huynh NN, Chin-Dusting J. Amino acids, arginase and nitric oxide in vascular health. *Clin Exp Pharmacol Physiol* 2006;**33**: 1–8.
52. Stamler J, Brown IJ, Daviglus ML *et al*. Dietary glycine and blood pressure, the International Study on Macro/Micronutrients and Blood Pressure. *Am J Clin Nutr* 2013;**98**:136–45.
53. Fraser GE, Yan R, Butler TL, Jaceldo-Siegl K, Beeson WL, Chan J. Missing data in a long food frequency questionnaire, are imputed zeroes correct? *Epidemiology* 2009;**20**:289–94.
54. Howard BV, Kritchevsky D. Phytochemicals and cardiovascular disease. *Circulation* 1997;**95**:2591–93.