Raw and Cooked Vegetable Consumption and Risk of Cardiovascular Disease: a Study of 400,000 Adults in UK Biobank

Qi Feng[1*], Jean H. Kim[2], Wemimo Omiyale[3], **Jelena Bešević**[3], Megan Conroy[3], Margaret May[4], Zuyao Yang[2], Samuel Y. Wong[2], Kelvin Kam-Fai Tsoi[2], Naomi Allen[3], Ben Lacey[3]

[1]Nuffield Department of Population Health, University of Oxford, United Kingdom, [2]The Chinese University of Hong Kong, China, [3]University of Oxford, United Kingdom, [4]University of Bristol, United Kingdom

Submitted to Journal:
Frontiers in Nutrition

Specialty Section:
Clinical Nutrition

Article type:
Original Research Article

Manuscript ID:
831470

Received on:
08 Dec 2021

Revised on:
10 Jan 2022

Journal website link:
www.frontiersin.org

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Author contribution statement

FQ designed the study and analyzed data. FQ, BL, JHK and MM interpreted results. FQ drafted the manuscript. All the coauthors critically reviewed and revised the manuscript.

Keywords

Vegetable intake, Raw vegetable, Cooked vegetable, Cardiovasclar diseases, UK Biobank, Cardiovascular mortality

Abstract

Word count: 260

Objectives: Higher levels of vegetable consumption have been associated with lower risk of cardiovascular disease (CVD), but the independent effect of raw and cooked vegetable consumption remains unclear.
Methods: From the UK Biobank cohort, 399,586 participants without prior CVD were included in analysis. Raw and cooked vegetable intakes were measured with a validated dietary questionnaire at baseline. Multivariable Cox regression was used to estimate the associations between vegetable intake and CVD incidence and mortality, adjusted for socioeconomic status, health status and lifestyle factors. The potential effect of residual confounding was assessed by calculating the percentage reduction in the likelihood-ratio (LR) statistics after adjustment for the confounders.
Results: The mean age was 56 years and 55% were women. Mean intakes of raw and cooked vegetables were 2.3 and 2.8 tablespoons/day, respectively. During 12 years of follow-up, 18,052 major CVD events and 4,406 CVD deaths occurred. Raw vegetable intake was inversely associated with both CVD incidence (adjusted hazard ratio (HR) [95%CI] for highest vs. lowest intake: 0.89 [0.83-0.95]) and CVD mortality (0.85 [0.74-0.97]), while cooked vegetable intake was not (1.00 [0.91-1.09] and 0.96 [0.80-1.13], respectively). Adjustment for potential confounders reduced the LR statistics for the associations of raw vegetables with CVD incidence and mortality by 82% and 87%, respectively.
Conclusions: Higher intakes of raw, but not cooked, vegetables were associated with lower CVD risk. Residual confounding is likely to account for much, if not all, of the observed associations. This study suggests the need to reappraise the evidence on the burden of CVD disease attributable to low vegetable intake in high-income populations.

Contribution to the field

There is substantial evidence on the association between vegetable intake and cardiovascular disease from observational studies and their meta-analyses, but the independent effects of raw and cooked vegetable intake remains unclear. Our analysis included 0.4 million UK Biobank participants without baseline cardiovascular diseases. Daily intakes of raw and cooked vegetables were obtained via a validated touchscreen questionnaire at recruitment. During 12 years of follow-up, 18,053 participants had a major cardiovascular events (11,317 myocardial infarction and 6969 stroke). After multivariable adjustment, we observed inverse associations between incidence of major cardiovascular disease and intake of total and raw vegetables, but no association with cooked vegetables. Statistical models that progressively adjusted for confounders led to large reductions in the predictive value of total and raw vegetable intake, indicating that the observed associations are likely to be accounted for by residual confounding. In this large prospective cohort study, higher levels of raw, but not cooked, vegetable intake were associated with lower cardiovascular risk. However, it is likely that residual confounding partly accounts for the observed associations. The present report highlights the need for rigorous assessment for residual confounding in studies of dietary factors on disease risk.

Ethics statements

Studies involving animal subjects

Generated Statement: No animal studies are presented in this manuscript.

Studies involving human subjects

Generated Statement: The studies involving human participants were reviewed and approved by UK Biobank was approved by the North West Multicenter Research Ethics Committee, the National Information Governance Board for Health and Social Care in England and Wales, and the Community Health Index Advisory Group in Scotland.. The patients/participants provided their written informed consent to participate in this study.

Inclusion of identifiable human data

Generated Statement: No potentially identifiable human images or data is presented in this study.

Data availability statement

Generated Statement: Publicly available datasets were analyzed in this study. This data can be found here: https://www.ukbiobank.ac.uk/.

Raw and Cooked Vegetable Consumption and Risk of Cardiovascular Disease: a Study of 400,000 Adults in UK Biobank

Qi FENG,[1, 2] Jean H KIM, [2] Wemimo OMIYALE, [1] Jelena BEŠEVIĆ, [1] Megan CONROY, [1] Margaret MAY, [3] Zuyao YANG, [2] Samuel Yeung-shan WONG, [2] Kelvin Kam-fai TSOI, [2,4*] Naomi ALLEN, [1*] Ben LACEY [1*]

1. Nuffield Department of Population Health (NDPH), University of Oxford, UK
2. JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China
3. Population Health Sciences, University of Bristol, UK
4. SH Big Data Decision Analytics Research Centre, The Chinese University of Hong Kong, Hong Kong, China

*: These authors share last authorship.

Correspondence:

Dr. Qi Feng

Nuffield Department of Population Health (NDPH),

Big Data Institute, University of Oxford, United Kingdom

Email: qi.feng@ndph.ox.ac.uk

Word count: 3500

Abstract

Objectives: Higher levels of vegetable consumption have been associated with lower risk of cardiovascular disease (CVD), but the independent effect of raw and cooked vegetable consumption remains unclear.

Methods: From the UK Biobank cohort, 399,586 participants without prior CVD were included in analysis. Raw and cooked vegetable intakes were measured with a validated dietary questionnaire at baseline. Multivariable Cox regression was used to estimate the associations between vegetable intake and CVD incidence and mortality, adjusted for socioeconomic status, health status and lifestyle factors. The potential effect of residual confounding was assessed by calculating the percentage reduction in the likelihood-ratio (LR) statistics after adjustment for the confounders.

Results: The mean age was 56 years and 55% were women. Mean intakes of raw and cooked vegetables were 2.3 and 2.8 tablespoons/day, respectively. During 12 years of follow-up, 18,052 major CVD events and 4,406 CVD deaths occurred. Raw vegetable intake was inversely associated with both CVD incidence (adjusted hazard ratio (HR) [95%CI] for highest vs. lowest intake: 0.89 [0.83-0.95]) and CVD mortality (0.85 [0.74-0.97]), while cooked vegetable intake was not (1.00 [0.91-1.09] and 0.96 [0.80-1.13], respectively). Adjustment for potential confounders reduced the LR statistics for the associations of raw vegetables with CVD incidence and mortality by 82% and 87%, respectively.

Conclusions: Higher intakes of raw, but not cooked, vegetables were associated with lower CVD risk. Residual confounding is likely to account for much, if not all, of the observed associations. This study suggests the need to reappraise the evidence on the burden of CVD disease attributable to low vegetable intake in high-income populations.

Keywords: vegetable intake; raw vegetables; cooked vegetables; cardiovascular disease; cardiovascular mortality; UK Biobank

Contribution to the Field

There is substantial evidence on the association between vegetable intake and cardiovascular disease from observational studies and their meta-analyses, but the independent effects of raw and cooked vegetable intake remains unclear. Our analysis included 0.4 million UK Biobank participants without baseline cardiovascular diseases. Daily intakes of raw and cooked vegetables were obtained via a validated touchscreen questionnaire at recruitment. During 12 years of follow-up, 18,053 participants had a major cardiovascular events (11,317 myocardial infarction and 6969 stroke). After multivariable adjustment, we observed inverse associations between incidence of major cardiovascular disease and intake of total and raw vegetables, but no association with cooked vegetables. Statistical models that progressively adjusted for confounders led to large reductions in the predictive value of total and raw vegetable intake, indicating that the observed associations are likely to be accounted for by residual confounding. In this large prospective cohort study, higher levels of raw, but not cooked, vegetable intake were associated with lower cardiovascular risk. However, it is likely that residual confounding partly accounts for the observed associations. The present report highlights the need for rigorous assessment for residual confounding in studies of dietary factors on disease risk.

Introduction

There exists a large body of research evidence to suggest that a high vegetable intake may protect against a wide range of health outcomes, including cardiovascular disease (CVD) (1,2). Although dietary guidelines have consistently recommended a high consumption of vegetables to the general population (3,4) as a source of beneficial macronutrients and micronutrients, such as dietary fiber, vitamins, and phytochemicals (5), it is estimated that inadequate vegetable consumption accounts for about 1.5 million premature deaths from cardiovascular disease alone each year (6).

However, little is known about the independent effects of cooked vegetables and raw vegetables on health outcomes. Previous epidemiological studies have demonstrated inconsistent findings. The EPIC study (7) of 450 000 participants recruited across Europe found that both cooked and raw vegetable intake was associated with lower CVD mortality and all-cause mortality. The PURE study (8) of 135 000 participants reported an inverse association with all-cause mortality for raw vegetable intake, but not for cooked vegetable intake, and neither cooked nor raw vegetable intake was associated with CVD incidence. An Australian cohort study (9) of 150 000 participants reported that only cooked vegetable intake was associated with a lower overall mortality, but did not investigate cardiovascular outcomes. The reason for the discrepancies in these findings is unclear, but may reflect variation in dietary patterns between populations as well as methodological difference, such as dietary assessment methods and insufficient adjustment for potential confounders.

UK Biobank is a cohort of half million participants with over a decade of follow-up (10). A wide range of participant characteristics were measured at baseline using standardized methods, minimizing measurement error and allowing for adjustment for a broad set of potential confounders. During follow-up, a large number of incident CVD and CVD deaths have been recorded, allowing for well-powered epidemiological investigations on cardiovascular outcomes (11). The objective of this study was to examine the effect of vegetable intake, and specifically the independent effects of raw and cooked vegetable intake, on CVD incidence and mortality in UK Biobank.

Methods

Study design and participants

UK Biobank is a population-based prospective cohort study (10). Between 2006 and 2010, half million participants aged 40-69 years were recruited across England, Wales and Scotland. Participants attended assessment centres, during which time they completed a touchscreen questionnaire that collected information on sociodemographic characteristics, lifestyle, health status, medication use, reproductive history, and environmental factors. In addition, anthropometric and other physical measures were taken using standardized procedures, and blood, urine and saliva samples were collected.

The participants' health was followed-up via linkage to hospitalization databases (the National Health Service [NHS] Hospital Episode Statistics for participants in England; the Scottish Morbidity Record for participants in Scotland; and the Patient Episode Database for participants in Wales) and national death registries (NHS Information Centre for participants in England and Wales; and NHS Central Registry for participants in Scotland). UK Biobank was approved by the North West Multicenter Research Ethics Committee, the National Information Governance Board for Health and Social Care in England and Wales, and the Community Health Index Advisory Group in Scotland. All participants provided informed consent.

This study excluded participants that withdrew their consents during follow-up, had missing data on vegetable intake, had prior CVDs, had conditions likely to change dietary pattern (e.g., pregnancy and cancer). Furthermore, 5 885 participants had missing data on other key covariates (body mass index [BMI], meat consumption and Townsend deprivation index), and were excluded. In total, 399 586 participants were included in analysis (Figure 1).

Measurement of exposures and outcomes

Information was collected at baseline on total daily intake of raw vegetables and of cooked vegetables. Participants were asked in the dietary questionnaire '*On average how many heaped tablespoons of salad or raw vegetables would you eat per day? (including lettuce, tomato in sandwiches)*' and '*On average how many heaped tablespoons of cooked vegetables would you eat per day? (do not include potatoes)*'. Total vegetable intake was calculated as the sum of raw and cooked vegetable intakes. Vegetable intake was categorized into four levels, using

cutoff values of 0, 1-2, 3-4 and ≥5 tablespoons/day for raw and cooked vegetable intake, and cutoff values of 0-1, 2-3, 4 to 7 and ≥8 for total vegetable intake. Previous analyses have demonstrated high repeatability and validity of vegetable consumption measured in this baseline dietary questionnaire: repeatability over a 4-year period is 82% for cooked vegetables and 72% for raw vegetables, with high agreement also observed when compared with 24-hour dietary assessment (12).

The primary outcomes were CVD incidence and mortality. The secondary outcomes were incident myocardial infarction (MI), incident stroke, and all-cause mortality. Incident CVD was defined as hospitalization or death from MI or stroke (13). CVD mortality was defined as death due to any cardiovascular disease. For analyses of disease incidence, participants were censored at date of hospitalisation, date of death, or last date of follow-up (31 March 2021 for participants from England and Scotland, and 28 February 2018 for participants from Wales), whichever occurred first. In mortality analysis, participants were censored at date of death or last date of follow-up (28 February 2021), whichever occurred first. Health outcomes were defined using the International Classification of Disease (ICD) codes. The exact ICD codes used are shown in Supplementary table 1.

Statistical analysis

Cox proportional hazard models were used to yield hazard ratios (HR) and 95% confidence interval (CI) for the associations between health outcomes and vegetable intake. Models were stratified by age (<50, 50-60, ≥60 years), sex, ethnicity, and region, and adjusted for educational attainment, Townsend deprivation index (continuous), hypertension, diabetes, physical activity level, smoking, alcohol consumption, BMI (continuous), use of mineral supplements, use of vitamin supplements, aspirin/ibuprofen, antihypertensive drugs, statins, insulin treatment, intake of fresh fruits, red meat, processed meat, oily fish and non-oily fish. The definition and measurement of the covariates are shown in the Supplementary Methods. The lowest intake level was used as the referent group. Test of linear trend was obtained by fitting the mean values of each vegetable intake level. The proportional hazards assumption was assessed using scaled Schoenfeld residuals (no violation was found in this study). Raw and cooked vegetable intake were mutually adjusted when investigating their independent effects. Variance inflation factor values were used to examine potential multi-collinearity.

We calculated the increase in the likelihood ratio (LR) $\chi 2$ statistics on addition of the vegetable intake term (raw, cooked and total) to the Cox models with various levels of adjustment of potential confounders. This provides a quantitative measure of the extent to which vegetable intake improve risk prediction for the outcome in different models. Comparisons of the changes in the LR $\chi 2$ statistic between model with minimal adjustments (e.g. age, sex, ethnicity, and region) to those with a more comprehensive set of confounders ('fully-adjusted' models) is therefore measure of the extent to which the confounders account for minimally-adjusted associations between vegetable intake and the outcome of interest. Furthermore, given that many confounders are measured imperfectly, the proportional change in this LR $\chi 2$ statistic is a semi-quantitative method of assessing for residual confounding, as models with perfectly measured confounders would be expected to further reduce the LR $\chi 2$ statistic in fully-adjusted models (14). More details are shown in Supplementary Methods.

For sensitivity analysis, we first excluded participants who developed the outcomes of interest during the first two years of follow-up, to minimize reverse causation. ~~Second, alternative to mutually adjusting for raw and cooked vegetable in multivariable regression, the effect of one kind vegetable intake can be investigated by restricting the sample to participants with zero intake of the other. Therefore, we examined the effects of cooked vegetable intake among the participants with *zero* raw vegetable intake, and the effects of raw vegetable intake among those with *zero* cooked vegetable intake. Third~~Second, we investigated the effect of the proportion of raw vegetable in total vegetable intake (raw vegetables divided by total vegetables), conditional on total vegetable intake and other covariates, after excluding the participants with total vegetable intake of *zero* tablespoon/day (n = 5 304). We conducted subgroup analysis based on ethnicity (White versus non-White), to examine potential ethnic differences in the associations. All analysis were performed using R (version 3.6.0; R Development Core Team, Vienna, Austria).

Results

After exclusion, 399 586 participants were included in the main analysis (Figure 1). The baseline characteristics of these participants is shown in Table 1 (Supplementary table 2). The mean age of participants was 56.1 (standard deviation 8.1) years, 55.4% were women, and 90.9% were of White ethnicity. Mean BMI was 27.3 (4.7) kg/m^2, 41.3% reported high levels of physical activity, and 4.7% had a self-reported history of diabetes. Mean intakes of total

vegetables, raw vegetables and cooked vegetables were 5.0 (3.4), 2.3 (2.2) and 2.8 (2.2) tablespoons/day, respectively; the distributions of total, raw and cooked vegetable intakes are shown in Supplementary figure 1.

Participants with higher levels of total vegetable intake were more likely to be women, better educated, and residents of an affluent area, with lower mean BMI and higher levels of physical activity, and less likely to be smokers. Raw and cooked vegetable intake were weakly correlated (Pearson correlation coefficient = 0.30). Variance inflation factor values for raw and cooked vegetable intake were 1.32 and 1.29, respectively, indicating very low collinearity (< 10). Supplementary table 3 and Supplementary table 4 showed the baseline characteristics of the participants across different raw vegetable intake levels and the cooked vegetable intake levels, respectively. The distributions of baseline characteristics by raw and cooked vegetable intake were similar to the distributions by total vegetable intake.

During a median follow-up of 12.1 years for CVD incidence outcomes, 18 052 participants developed CVD (11 317 MI and 6 969 stroke). There was an inverse association between incident CVD and total and raw vegetable intake, but not cooked vegetable intake (Figure 2; Supplementary figure 3). Compared with the lowest level of total vegetable intake, the highest level was associated with 10% lower CVD incidence (HR [95%CI] 0.90 [0.83 to 0.97]). Higher intake of raw vegetable intake was inversely associated with incident CVD (HR [95%CI] for highest vs. lowest intake: 0.89 [0.83 to 0.95]) and incident MI (0.88 [0.81 to 0.96]; Figure 2), whereas cooked vegetable intake showed null associations with incident CVD (1.00 [0.91 to 1.09]) or incidence MI (0.97 [0.86 to 1.08]). We noted a potential inverse association between raw vegetable intake and incident stroke, although this was not statistically significant. No evidence was found for association between incident stroke and total~~, raw~~ or cooked vegetable intakes (Figure 2).

During a median follow-up of 12.0 years for mortality outcomes, 13 398 participants died, of which 2589 deaths were due to CVD. Consuming 2 or more heaped tablespoons/day of total vegetables was associated with lower risk of CVD mortality (HR [95%CI] for highest vs. lowest intake: 0.83 [0.71-0.96]), but there was little evidence of trend in risk with higher levels of intake (Figure 3). Similarly, there was evidence of an inverse association of CVD mortality with raw vegetable intake (0.85 [0.74-0.97]) but little evidence of a trend (p=0.164), and there was no evidence of an association of CVD mortality with cooked vegetables. For all-cause

mortality, there was a strong inverse association with eating some vegetables (1 or more tablespoons of raw or cooked vegetables per day), and strong evidence of trend with increasing raw vegetable intake (p<0.001) but not cooked vegetables (p=0.932).

Progressive adjustment for potential confounders attenuated HR estimates and substantially reduced the LR $\chi 2$ statistics in adjusted models (Table 2). For models of CVD incidence and raw vegetable intake, covariate adjustment attenuated HR (highest vs. lowest intake groups) from 0.79 (0.74 to 0.84) to 0.88 (0.83, 0.94), with the LR $\chi 2$ statistic declining by 81.9%. This substantial attenuation suggests that were the potential confounders measured perfectly, much, if not all, of the observed association with reported vegetable intake would be explained by residual confounding, although one cannot rule out the possibility of a true causal protective effect. Similar findings were observed for MI, CVD mortality and all-cause mortality with both raw and cooked vegetable intake, with the proportional changes in the LR $\chi 2$ statistic of about 80% or more (Table 2, Supplementary table 5). Adjustment for socioeconomic (including educational attainment, and Townsend deprivation index) and lifestyle factors (including physical activity, smoking, drinking, use of mineral supplements, use of vitamin supplements, fruit intake, oily fish intake, non-oily fish intake, red meat intake, processed meat intake) results in most of the reductions in LR $\chi 2$ statistic, while further adjustment for BMI and baseline health status resulted in only slight further reductions (Supplementary table 5), suggesting that the observed associations are likely to be accounted for by residual confounding from socioeconomic status and lifestyle factors.

In sensitivity analyses, when adjusting for total vegetable intake, higher proportion of raw vegetable intake in total vegetable intake was associated with lower CVD incidence and all-cause mortality, but not with other outcomes (Supplementary table 6). ~~In analyses restricted to participants with *zero* raw vegetable intake (n = 39 436), cooked vegetable intake showed inverse association with all-cause mortality, but null associations with CVD incidence and mortality were found, consistent with the primary results (Supplementary table 7). Among the participants with *zero* cooked vegetable intake (n = 12 102), raw vegetable intake's association with the outcomes became null, although there was limited power to assess these association reliably (Supplementary table 8).~~ Furthermore, excluding the participants who had outcome events within the first two years of follow-up did not materially change the main findings (Supplementary table 7~~9~~). Subgroup analyses restricted to White participants (n = 378 028) showed similar results to the primary analysis (Supplementary table 8); and there was no

evidence that the associations differed to those of non-White ethnicity, although there were substantially fewer non-White participants (n = 21 558) (Supplementary table 9), and as such limited power to assess for heterogeneity.

Discussion

In this large prospective cohort study, total vegetable intake was associated with reduced risks of CVD incidence, CVD mortality, and all-cause mortality. When assessing the independent effect of raw and cooked vegetable intake, only raw vegetable intake showed inverse associations with CVD outcomes, whereas cooked vegetables showed no association. However, given the large reductions in the predictive values of total and raw vegetable intake after adjustment for socioeconomic and lifestyle factors, residual confounding is likely to account for much, if not all, of the remaining associations.

The modest inverse associations of total vegetable intake with CVD outcomes and all-cause mortality in our analyses are consistent with previous large-scale observational evidence. For example, a meta-analysis of 24 cohort studies estimated that high vegetable intake reduced all-cause mortality by about 13% (relative risk 0.87 [0.82-0.92]) (15). Previous systematic reviews showed total vegetable consumption was associated with a risk reduction in CVD incidence by 11% (15) to 18% (16), similar to the ~10% lower risk in the present study. Our findings of the inverse association with MI are also in line with published meta-analyses with effect sizes ranging from 9% to 15% (15–17). Although previous studies have also demonstrated an association with reduced risk of stroke (15–17), we did not find sufficient evidence for such an association.

In contrast to the large number of studies on total vegetable intake, there is limited evidence on the independent effect of raw and cooked vegetables on all-cause mortality. Aune *et al* (15) conducted a meta-analysis that found cooked vegetable was associated with 13% (relative risk 0.87 [0.80-0.94]) lower risk of all-cause mortality, and raw vegetable was associated with 12% (relative risk 0.88 [0.79-0.98]) lower risk of mortality, although the analyses of raw and cooked vegetables were not mutually adjusted. Studies which have attempted to assess the independent effects of raw and cooked vegetable intakes on all-cause mortality have reported conflicting results. Our results are broadly consistent with the EPIC study (7), in which both raw vegetable intake and cooked vegetable intake were associated with reduced risk of all-cause mortality.

By contrast, the PURE study (8) reported an inverse association with all-cause mortality for raw vegetable intake, but not for cooked vegetable intake, whilst an Australian cohort study (9) reported that only cooked vegetable intake was associated with a lower overall mortality. The characteristics and main findings of these studies are summarized in Supplementary table 10.

In this study, cooked vegetable intake and raw vegetable intake showed different associations with cardiovascular outcomes. We found inverse associations of raw vegetables with CVD incidence and mortality, but null associations with cooked vegetables. This is consistent with the MORGEN study, a Dutch cohort (18), in which raw, but not processed, vegetables were associated with a reduced risk of ischaemic stroke. In the EPIC cohort (7), there was a stronger inverse association of CVD mortality with raw than cooked vegetables. Whereas the PURE study (8) found no evidence of an association of cardiovascular disease and raw vegetable intake, and high intakes levels of cooked vegetable was positively associated with CVD incidence.

Previous studies that reported associations of higher levels of vegetable intake with lower risk of CVD have proposed various mechanisms by which these associations may be mediated. For example, it has been suggested that diets high in vegetables have, on average, fewer calories and replaces foods that are high in fat, sodium and glycemic load (15,19). It has also been hypothesized that the lower risk might be mediated by micronutrients, namely higher intake of vitamins, polyphenols and antioxidant compounds (2,5), which are required for regulating various biological processes, including anti-oxidation, anti-inflammation, lipid metabolism, and endothelial function (20). As for the different associations of raw and cooked vegetables observed in the present and other studies, several possible mechanisms have been proposed in previous studies. Firstly, it has been proposed that the kinds of the vegetables that are usually consumed cooked (e.g. beans, peas, eggplant, etc.) may differ from those typically consumed raw (e.g. lettuce, etc.). Secondly, cooking processes can alter the digestibility of food as well as bioavailability of nutrients (21). For example, Lee *et al.* found that vitamin C retention after cooking ranged from 0% to 91% for various combinations of cooking method and vegetable, with higher retention after microwaving and lower retention after boiling (22). Thirdly, the seasoning and oils used in cooking vegetables often increase intake of sodium and fat, which are known risk factors for CVD incidence and mortality (23,24).

Despite these proposed mechanisms, the present study indicates that observed associations of vegetable intakes with CVD risk and all-cause mortality are likely to be mostly accounted for by residual confounding. Studies using Mendelian randomisation (which are less susceptible to confounding, and other biases of observational studies) might be particular usefully in reliably assessing the associations of diet on disease risk. For example, a recent Mendelian randomization study that used genetic determinants of plasma vitamin C concentration as a surrogate for vegetable intake reported a null association with ischemic heart disease (odds ratio 0.90 [0.75 to 1.08]) and all-cause mortality (odds ratio 0.88 [0.72 to 1.08]), despite strong inverse associations between vitamin C and these outcomes in observational analyses (25).

The present study found the observed associations were mainly accounted for by socioeconomic status and lifestyle factors (26). Both low socioeconomic status and major lifestyle factors, such as smoking and alcohol intake, are established risk factors for CVD, and there is strong evidence that the effect of socioeconomic status is partially mediated by the known lifestyle factors (27). For example, one study reported that an unhealthy lifestyle (including smoking, drinking, obesity, physical inactivity and others) mediated 34-38% of the association between socioeconomic status and all-cause death (28). Therefore, given the complicated inter-relationship between socioeconomic status, lifestyle and health outcomes, adjustment of measures of both socioeconomic status and lifestyle factors is likely to be important.

This study has some limitations. First, we did not measure intake of specific types of raw or cooked vegetables, nor were we able to account for differences in cooking methods. Second, vegetable intakes are self-reported in baseline dietary questionnaire, although the validity and repeatability of the UK Biobank baseline dietary questionnaire has been evaluated and confirmed in previous studies (12). Third, we did not adjust for total calorie intake because such information was not available from the baseline dietary questionnaire, but we did adjusted for physical activity level and BMI, which has been shown as a valid method for isocaloric adjustment (29). Future studies should seek to address these limitations. However, such studies should also be aware of the importance of assessing reliably for residual confounding using similar methods to the present report, or other approaches, such as Mendelian randomisation.

Although the present report does not find strong evidence of an association between higher vegetable intake and reduced risk of major cardiovascular disease, the wider literature suggests

that increasing vegetable intake is likely to reduce risk of some other common diseases (4). As such, the public health recommendations on the benefits to health and the environment of a diet that is high in vegetable intake remains.

Conclusion

In this study of 0.4 million middle-age adults with 12-year follow-up, higher intakes of raw but not cooked vegetables were associated with lower CVD risk. However, given the large reductions in the predictive values of raw vegetable intake after adjustment for socioeconomic and lifestyle factors, residual confounding is likely to account for much, if not all, of the remaining associations. This study highlights the need for rigorous assessment for residual confounding in studies of the effects of diet and other lifestyle factors on disease risk, and suggests the need to reappraise the evidence on the burden of CVD disease attributable to low vegetable intake in high-income populations.

Contributors

FQ designed the study and analyzed data. FQ, BL, JHK and MM interpreted results. FQ drafted the manuscript. All the coauthors critically reviewed and revised the manuscript.

Declaration of interests

All authors declare none conflict of interest.

Acknowledgements

We thank the participants and investigators in UK Biobank. This analysis was conducted under the UK Biobank application 65563.

This research was funded in whole, or in part, by the Wellcome Trust [205339/Z/16/Z]. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

Data availability

Individual level data are available at UK Biobank website https://www.ukbiobank.ac.uk/ , upon application to and approval from UK Biobank. The analytic codes are available upon request to the corresponding author.

Figure Legends

Figure 1: Flowchart of participant included in the main analysis

Figure 2: Incident cardiovascular disease (CVD), myocardial infarction (MI) and stroke versus vegetable consumption

Hazard ratios (HR; fully adjusted models) and 95% confidence interval (CI) by level of total, raw, and cooked vegetable consumption (heaped tablespoons/day), relative to the lowest (reference) consumption level. Exclusions as in Table 1. The model was stratified by age (<50, 50-60, ≥60 years), sex, ethnicity and region, and adjusted for educational attainment, Townsend deprivation index (continuous), hypertension, diabetes, physical activity level, smoking, alcohol consumption, BMI (continuous), use of mineral supplements, use of vitamin supplements, aspirin/ibuprofen, antihypertensive drugs, statins, insulin treatment, intake of fresh fruits, red meat, processed meat, oily fish and non-oily fish.

Figure 3: Cardiovascular disease (CVD) mortality and all-cause mortality versus vegetable consumption

Hazard ratios (HR; fully adjusted models) and 95% confidence interval (CI) by level of total, raw, and cooked vegetable consumption (heaped tablespoons/day), relative to the lowest (reference) consumption level. Exclusions as in Table 1. The model was stratified by age (<50, 50-60, ≥60 years), sex, ethnicity and region, and adjusted for educational attainment, Townsend deprivation index (continuous), hypertension, diabetes, physical activity level, smoking, alcohol consumption, BMI (continuous), use of mineral supplements, use of vitamin supplements, aspirin/ibuprofen, antihypertensive drugs, statins, insulin treatment, intake of fresh fruits, red meat, processed meat, oily fish and non-oily fish.

References

1. Angelino D, Godos J, Ghelfi F, Tieri M, Titta L, Lafranconi A, Marventano S, Alonzo E, Gambera A, Sciacca S, et al. Fruit and vegetable consumption and health outcomes: an umbrella review of observational studies. *Int J Food Sci Nutr* (2019) 70:652–667. doi:10.1080/09637486.2019.1571021

2. Wallace TC, Bailey RL, Blumberg JB, Burton-Freeman B, Chen C-YO, Crowe-White KM, Drewnowski A, Hooshmand S, Johnson E, Lewis R, et al. Fruits, vegetables, and health: A comprehensive narrative, umbrella review of the science and recommendations for enhanced public policy to improve intake. *Crit Rev Food Sci Nutr* (2020) 60:2174–2211. doi:10.1080/10408398.2019.1632258

3. US Department of Health and Human Services and US Department of Agriculture. *Dietary guidelines for Americans 2015-2020, 8th edition*. (2015). Available at: https://health.gov/our-work/food-nutrition/previous-dietary-guidelines/2015

4. World Health Organization. Healthy diet. *World Health Organization* (2020) Available at: https://www.who.int/news-room/fact-sheets/detail/healthy-diet [Accessed January 12, 2021]

5. Slavin JL, Lloyd B. Health Benefits of Fruits and Vegetables. *Advances in Nutrition* (2012) 3:506–516. doi:10.3945/an.112.002154

6. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, Mullany EC, Abate KH, Abbafati C, Abebe Z, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* (2019) 393:1958–1972. doi:10.1016/S0140-6736(19)30041-8

7. Leenders M, Sluijs I, Ros MM, Boshuizen HC, Siersema PD, Ferrari P, Weikert C, Tjønneland A, Olsen A, Boutron-Ruault M-C, et al. Fruit and vegetable consumption and mortality: European prospective investigation into cancer and nutrition. *Am J Epidemiol* (2013) 178:590–602. doi:10.1093/aje/kwt006

8. Miller V, Mente A, Dehghan M, Rangarajan S, Zhang X, Swaminathan S, Dagenais G, Gupta R, Mohan V, Lear S, et al. Fruit, vegetable, and legume intake, and cardiovascular disease and deaths in 18 countries (PURE): a prospective cohort study. *Lancet* (2017) 390:2037–2049. doi:10.1016/S0140-6736(17)32253-5

9. Nguyen B, Bauman A, Gale J, Banks E, Kritharides L, Ding D. Fruit and vegetable consumption and all-cause mortality: evidence from a large Australian cohort study. *Int J Behav Nutr Phys Act* (2016) 13:9. doi:10.1186/s12966-016-0334-5

10. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* (2015) 12:e1001779. doi:10.1371/journal.pmed.1001779

11. Littlejohns TJ, Sudlow C, Allen NE, Collins R. UK Biobank: opportunities for cardiovascular research. *European Heart Journal* (2019) 40:1158–1166. doi:10.1093/eurheartj/ehx254

12. Bradbury KE, Young HJ, Guo W, Key TJ. Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire. *J Nutr Sci* (2018) 7:e6. doi:10.1017/jns.2017.66

13. Fan M, Sun D, Zhou T, Heianza Y, Lv J, Li L, Qi L. Sleep patterns, genetic susceptibility, and incident cardiovascular disease: a prospective study of 385 292 UK biobank participants. *European Heart Journal* (2020) 41:1182–1189. doi:10.1093/eurheartj/ehz849

14. Floud S, Balkwill A, Moser K, Reeves GK, Green J, Beral V, Cairns BJ. The role of health-related behavioural factors in accounting for inequalities in coronary heart disease risk by education and area deprivation: prospective study of 1.2 million UK women. *BMC Med* (2016) 14:145. doi:10.1186/s12916-016-0687-2

15. Aune D, Giovannucci E, Boffetta P, Fadnes LT, Keum N, Norat T, Greenwood DC, Riboli E, Vatten LJ, Tonstad S. Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality-a systematic review and dose-response meta-analysis of prospective studies. *Int J Epidemiol* (2017) 46:1029–1056. doi:10.1093/ije/dyw319

16. Zhan J, Liu Y-J, Cai L-B, Xu F-R, Xie T, He Q-Q. Fruit and vegetable consumption and risk of cardiovascular disease: A meta-analysis of prospective cohort studies. *Crit Rev Food Sci Nutr* (2017) 57:1650–1663. doi:10.1080/10408398.2015.1008980

17. Bechthold A, Boeing H, Schwedhelm C, Hoffmann G, Knüppel S, Iqbal K, De Henauw S, Michels N, Devleesschauwer B, Schlesinger S, et al. Food groups and risk of coronary heart disease, stroke and heart failure: A systematic review and dose-response meta-analysis of prospective studies. *Crit Rev Food Sci Nutr* (2019) 59:1071–1090. doi:10.1080/10408398.2017.1392288

18. Oude Griep LM, Verschuren WMM, Kromhout D, Ocké MC, Geleijnse JM. Raw and processed fruit and vegetable consumption and 10-year stroke incidence in a population-based cohort study in the Netherlands. *Eur J Clin Nutr* (2011) 65:791–799. doi:10.1038/ejcn.2011.36

19. Jenkins DJA, Dehghan M, Mente A, Bangdiwala SI, Rangarajan S, Srichaikul K, Mohan V, Avezum A, Díaz R, Rosengren A, et al. Glycemic Index, Glycemic Load, and Cardiovascular Disease and Mortality. *N Engl J Med* (2021)NEJMoa2007123. doi:10.1056/NEJMoa2007123

20. Tang G-Y, Meng X, Li Y, Zhao C-N, Liu Q, Li H-B. Effects of Vegetables on Cardiovascular Diseases and Related Mechanisms. *Nutrients* (2017) 9: doi:10.3390/nu9080857

21. Palermo M, Pellegrini N, Fogliano V. The effect of cooking on the phytochemical content of vegetables. *J Sci Food Agric* (2014) 94:1057–1070. doi:10.1002/jsfa.6478

22. Lee S, Choi Y, Jeong HS, Lee J, Sung J. Effect of different cooking methods on the content of vitamins and true retention in selected vegetables. *Food Sci Biotechnol* (2018) 27:333–342. doi:10.1007/s10068-017-0281-1

23. He FJ, Li J, Macgregor GA. Effect of longer term modest salt reduction on blood pressure: Cochrane systematic review and meta-analysis of randomised trials. *BMJ* (2013) 346:f1325. doi:10.1136/bmj.f1325

24. Chanita U, Prapimporn S, Daruneewan W, Vijj K, Thakkinstian A. Oil Consumption and Cardiovascular Disease: An Umbrella Review of Systematic Reviews and Meta-Analyses. *Current Developments in Nutrition* (2020) 4:571–571. doi:10.1093/cdn/nzaa046_071

25. Kobylecki CJ, Afzal S, Davey Smith G, Nordestgaard BG. Genetically high plasma vitamin C, intake of fruit and vegetables, and risk of ischemic heart disease and all-cause mortality: a Mendelian randomization study. *Am J Clin Nutr* (2015) 101:1135–1143. doi:10.3945/ajcn.114.104497

26. Tomson J, Emberson J, Hill M, Gordon A, Armitage J, Shipley M, Collins R, Clarke R. Vitamin D and risk of death from vascular and non-vascular causes in the Whitehall study and meta-analyses of 12,000 deaths. *Eur Heart J* (2013) 34:1365–1374. doi:10.1093/eurheartj/ehs426

27. Wang J, Geng L. Effects of socioeconomic status on physical and psychological health: lifestyle as a mediator. *Int J Environ Res Public Health* (2019) 16: doi:10.3390/ijerph16020281

28. Laine JE, Baltar VT, Stringhini S, Gandini M, Chadeau-Hyam M, Kivimaki M, Severi G, Perduca V, Hodge AM, Dugué P-A, et al. Reducing socio-economic inequalities in all-cause mortality: a counterfactual mediation approach. *Int J Epidemiol* (2020) 49:497–510. doi:10.1093/ije/dyz248

29. Jakes RW, Day NE, Luben R, Welch A, Bingham S, Mitchell J, Hennings S, Rennie K, Wareham NJ. Adjusting for energy intake--what measure to use in nutritional epidemiological studies? *Int J Epidemiol* (2004) 33:1382–1386. doi:10.1093/ije/dyh181

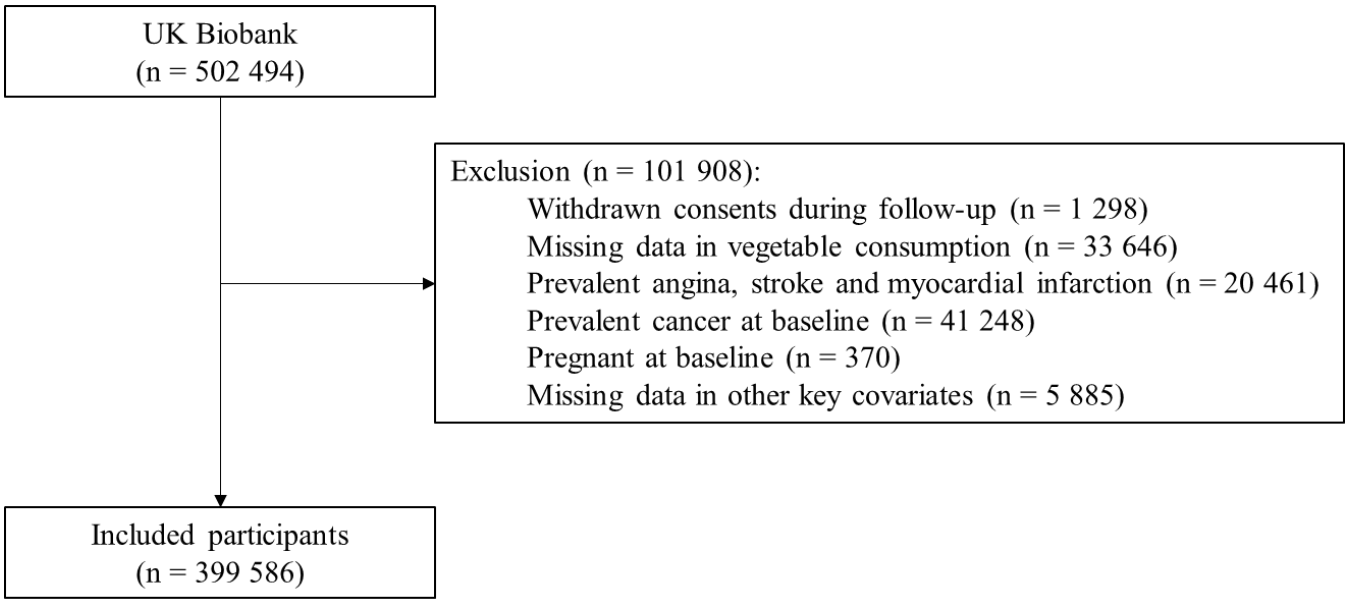

Figure 1: Flowchart of participant included in the main analysis

Table 1: Baseline characteristics of the 399 586 participants in the main analysis, by total vegetable consumption

	≤1 tablespoon/ day (n=15 902)	2-3 tablespoons/ day (n=109 536)	4-7 tablespoons/ day (n=216 499)	≥8 tablespoons/ day (n = 57 649)	Overall (n=399 586)
Female (%)	6 174 (38.8)	54 948 (50.2)	126 375 (58.4)	33 997 (59.0)	221 494 (55.4)
Age (years)	54.0 (8.1)	55.3 (8.2)	56.5 (8.0)	56.4 (8.0)	56.1 (8.1)
Total vegetable intake (tablespoons/day)	0.7 (0.5)	2.6 (0.5)	5.1 (1.0)	10.7 (5.0)	5.0 (3.4)
Raw vegetable intake (tablespoons/day)	0.1 (0.3)	0.9 (0.6)	2.2 (1.1)	5.5 (3.5)	2.3 (2.5)
Cooked vegetable intake (tablespoons/day)	0.5 (0.5)	1.7 (0.6)	2.8 (1.0)	5.3 (3.4)	2.8 (1.9)
White ethnicity (%)	14 782 (93.3)	104 731 (95.9)	206 372 (95.6)	52 143 (90.9)	378 028 (94.9)
Townsend Deprivation index*	-0.2 (3.5)	-1.4 (3.0)	-1.5 (2.9)	-1.1 (3.1)	-1.4 (3.0)
University educated (%)	3321 (21.3)	37040 (34.3)	73733 (34.6)	19483 (34.5)	133577 (34.0)
Body mass index (kg/m^2)	28.0 (5.2)	27.2 (4.7)	27.3 (4.7)	27.4 (4.8)	27.3 (4.7)
Current smoker (%)	3 485 (22.0)	11 828 (10.8)	19 427 (9.0)	5 506 (9.6)	40 246 (10.1)
Current drinker (%)	13 817 (87.1)	101 873 (93.1)	201 964 (93.3)	52 123 (90.5)	369 777 (92.6)
High physical activity level (%)[†]	3 971 (32.2)	31 459 (35.0)	75 528 (42.5)	24 328 (51.0)	135 286 (41.3)
Self-reported hypertension (%)	4 172 (26.2)	26 482 (24.2)	55 071 (25.4)	15 131 (26.2)	100 856 (25.2)
Self-reported diabetes (%)	994 (6.2)	4 859 (4.4)	9 904 (4.6)	3 009 (5.2)	18 766 (4.7)
Regular use of aspirin/ibuprofen (%)	4 065 (25.6)	26 039 (23.8)	53 667 (24.8)	14 394 (25.0)	98 165 (24.6)
Regular use of mineral supplement (%)	2 869 (18.0)	25 789 (23.5)	61 980 (28.6)	17 955 (31.1)	108 593 (27.2)
Regular use of vitamin supplement (%)	1 760 (11.1)	13 223 (12.1)	30 534 (14.1)	9 756 (16.9)	55 273 (13.8)
Use of antihypertensive drugs (%)	938 (5.9)	8 133 (7.4)	20 449 (9.4)	5 628 (9.8)	35 148 (8.8)
Use of statin (%)	696 (4.4)	5 426 (4.9)	13 443 (6.2)	3 822 (6.6)	23 387 (5.9)
Use of insulin (%)	58 (0.4)	375 (0.3)	855 (0.4)	275 (0.5)	1 563 (0.4)
Fruit intake ≥5 pieces/day (%)	684 (4.3)	4 659 (4.3)	15 781 (7.3)	10 076 (17.5)	31 200 (7.8)
Oily fish intake >1 times/week (%)	3 267 (20.7)	39 335 (36.0)	88 514 (41.0)	21 682 (37.7)	72 515 (18.2)
Non-oily fish intake >1 times /week (%)	1 565 (9.9)	13 441 (12.3)	37 587 (17.4)	13 698 (23.8)	66 291 (16.6)
Processed meat intake ≥2 times/week (%)	6 949 (43.8)	38 331 (35.0)	62 278 (28.8)	14 132 (24.5)	121 690 (30.5)
Red meat intake (times/week)	2.0 (1.6)	2.1 (1.4)	2.1 (1.4)	2.0 (1.6)	2.1 (1.4)

Data are mean (standard deviation) or frequency (percentage). Vegetable consumption was self-reported in number of heaped tablespoons/day. People with baseline angina, stroke, myocardial infarction, cancer, pregnancy, and missing data on vegetable consumption and other important covariates were excluded. *Area-level measure of material deprivation. † High physical activity defined based on International Physical Activity Questionnaire and WHO guideline.

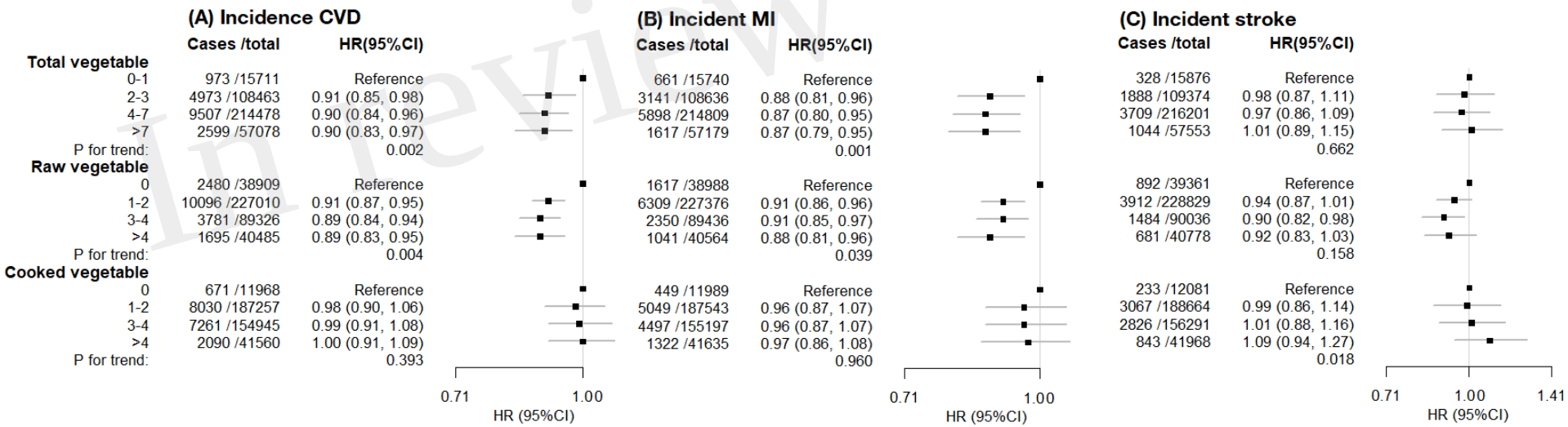

Figure 2: Incident cardiovascular disease (CVD), myocardial infarction (MI) and stroke versus vegetable consumption

Hazard ratios (HR; fully adjusted models) and 95% confidence interval (CI) by level of total, raw, and cooked vegetable consumption (heaped tablespoons/day), relative to the lowest (reference) consumption level. Exclusions as in Table 1. The model was stratified by age (<50, 50-60, ≥60 years), sex, ethnicity and region, and adjusted for educational attainment, Townsend deprivation index (continuous), hypertension, diabetes, physical activity level, smoking, alcohol consumption, BMI (continuous), use of mineral supplements, use of vitamin supplements, aspirin/ibuprofen, antihypertensive drugs, statins, insulin treatment, intake of fresh fruits, red meat, processed meat, oily fish and non-oily fish.

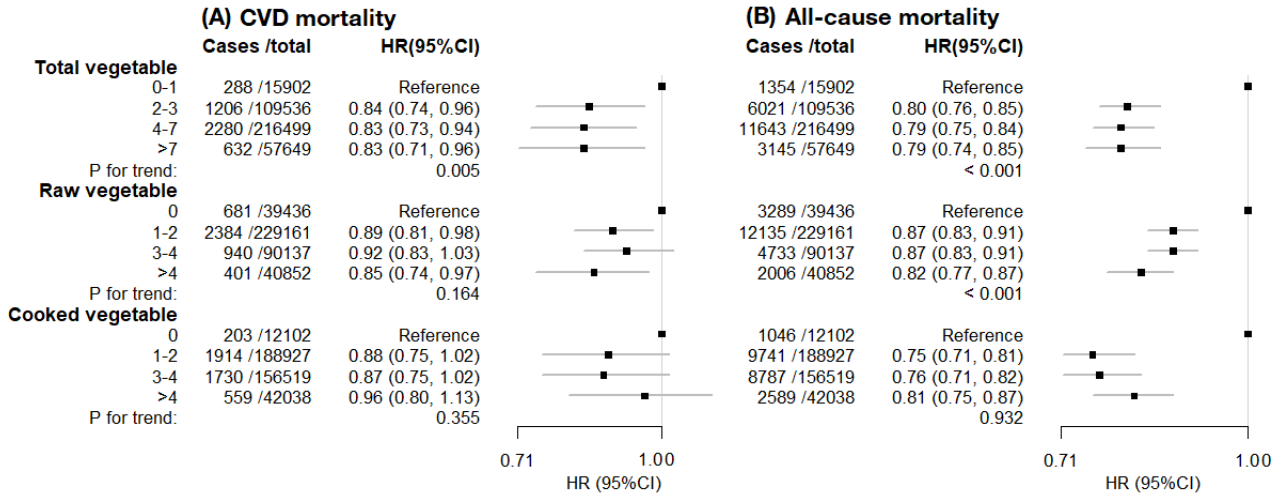

Figure 3: Cardiovascular disease (CVD) and all-cause mortality versus vegetable consumption
Hazard ratios (HR; fully adjusted models) and 95% confidence interval (CI) by level of total, raw, and cooked vegetable consumption (heaped tablespoons/day), relative to the lowest (reference) consumption level. Exclusions as in Table 1. The model was stratified by age (<50, 50-60, ≥60 years), sex, ethnicity and region, and adjusted for educational attainment, Townsend deprivation index (continuous), hypertension, diabetes, physical activity level, smoking, alcohol consumption, BMI (continuous), use of mineral supplements, use of vitamin supplements, aspirin/ibuprofen, antihypertensive drugs, statins, insulin treatment, intake of fresh fruits, red meat, processed meat, oily fish and non-oily fish.

Table 2: Associations between vegetable intake with CVD incidence, myocardial infarction incidence, stroke incidence, CVD mortality and all-cause mortality in basic model and fully-adjusted model

	Basic model		Fully-adjusted model		Attenuation (% reduction in χ^2) †
	Improvement in fit (χ^2)	HR (95% CI)*	Improvement in fit (χ^2)	HR (95% CI)*	
CVD incidence					
Total vegetable intake	87.8	0.74 (0.69, 0.80)	10.1	0.90 (0.83, 0.97)	88.6
Raw vegetable intake	127.9	0.79 (0.74, 0.84)	23.2	0.89 (0.83, 0.95)	81.9
Cooked vegetable intake	53.0	0.77 (0.71, 0.84)	1.5	1.00 (0.91, 1.09)	97.2
MI incidence					
Total vegetable intake	75.1	0.71 (0.65, 0.78)	11.1	0.87 (0.79, 0.95)	85.2
Raw vegetable intake	88.8	0.78 (0.72, 0.84)	13.3	0.88 (0.81, 0.96)	85.0
Cooked vegetable intake	42.8	0.74 (0.67, 0.83)	0.6	0.97 (0.86, 1.08)	98.5
Stroke incidence					
Total vegetable intake	18.8	0.84 (0.74, 0.95)	2.2	1.01 (0.89, 1.15)	88.1
Raw vegetable intake	31.7	0.85 (0.77, 0.94)	5.6	0.92 (0.83, 1.03)	82.2
Cooked vegetable intake	19.1	0.87 (0.75, 1.01)	6.1	1.09 (0.94, 1.27)	68.3
All CVD mortality					
Total vegetable intake	58.2	0.63 (0.55, 0.73)	8.0	0.83 (0.71, 0.96)	86.3
Raw vegetable intake	63.8	0.74 (0.65, 0.84)	8.2	0.85 (0.74, 0.97)	87.2
Cooked vegetable intake	53.9	0.67 (0.57, 0.78)	6.3	0.96 (0.80, 1.13)	88.4
All-cause mortality					
Total vegetable intake	298.7	0.61 (0.57, 0.65)	57.9	0.80 (0.74, 0.85)	80.6
Raw vegetable intake	352.7	0.69 (0.65, 0.73)	57.3	0.82 (0.77, 0.87)	83.8
Cooked vegetable intake	347.8	0.57 (0.53, 0.61)	72.0	0.81 (0.75, 0.87)	79.3

Improvement in the prediction of relative risk (LR $\chi 2$ statistic) by the addition of the given vegetable intake term to the basic model (in which the relative risk depends only on age, sex, ethnicity, and region), and to the fully-adjusted models with all major potential confounders, including mutual adjustment for raw and cooked vegetables, as described in the Methods. *Hazard ratio for highest versus lowest vegetable intake group. †Proportional reduction in $\chi 2$ for the improvement in model fit relative to the basic model, equivalent to the proportion of the association attenuated by the potential confounders.