Supplementary Information

A global analysis of dairy consumption and incident cardiovascular disease

Table of Contents

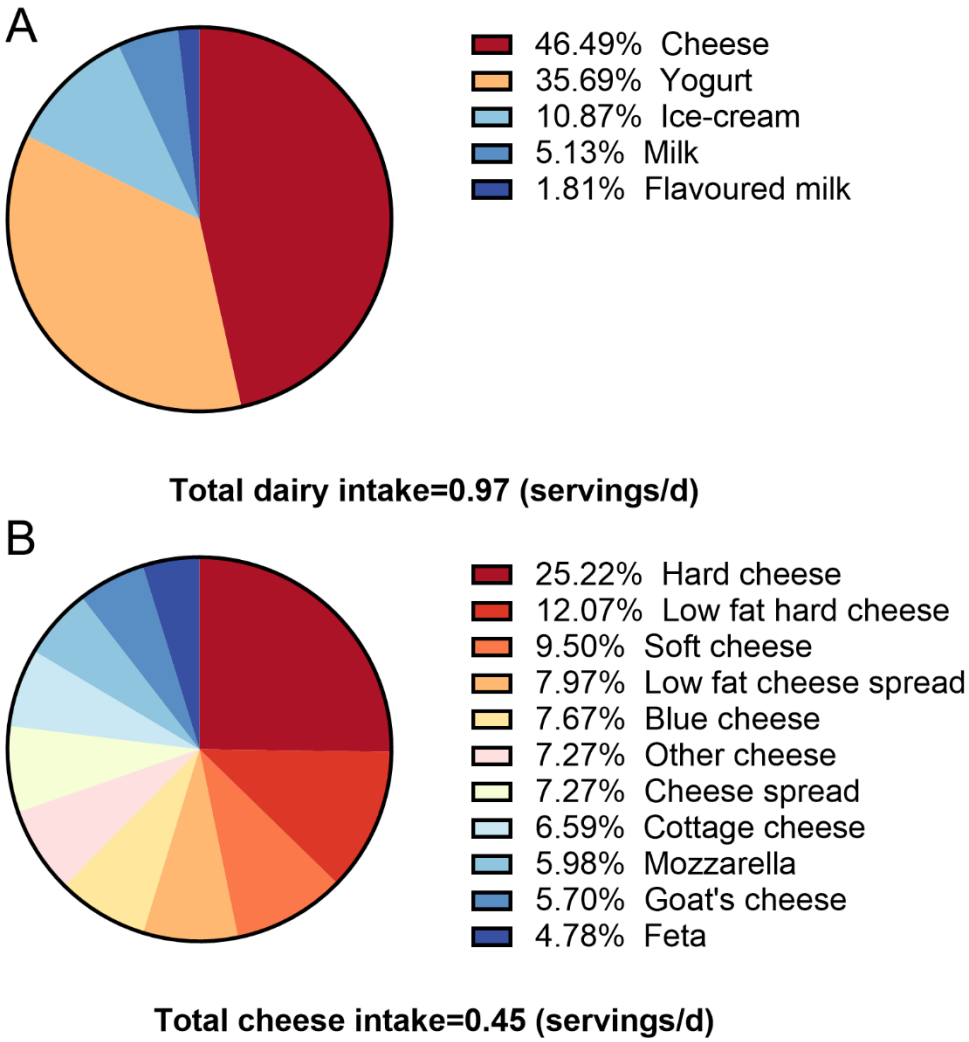
Supplementary Methods

Description of meta-analysis methodology

We performed a systematic review and updated meta-analysis including UKB and CKB studies as well as previous prospective cohort studies which explored the relationship of dairy product intake with CVD risk in the general population. This meta-analysis was registered on the international prospective register (PROSPERO: CRD42021283876) and followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement.[1] The systematic search of the PubMed, EMBASE databases, and Web of Science was conducted and updated to 12 November 2023. The Supplementary Table 2 shows the search strategy. Besides, reference lists of included studies were also screened to find potentially relevant studies. Inclusion criteria were shown as follows: observational studies with prospective design; assessing the association between dairy products consumption (total dairy, milk, yogurt, cheese, cream, or butter) and incidence of CVD (CHD, stroke, or total CVD); providing risk estimates and 95% CIs for the final models; available description of covariables in the statistical models.

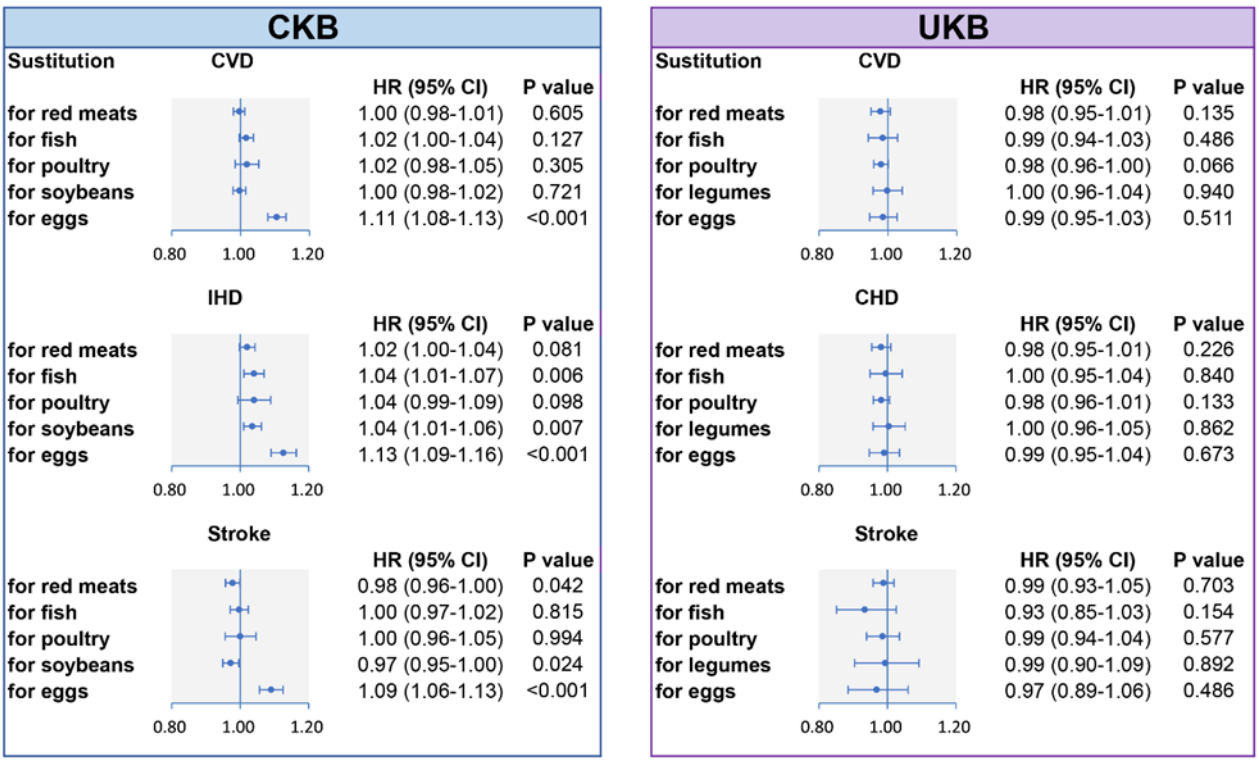
Two reviewers independently extracted the data from the eligible articles, including cohort name, first author, year of publication, location, follow-up duration, number of participants, age range, dairy types, methods for dietary assessment, outcome, approach for outcome ascertainment, number of cases, dairy product intake categories, relative risks (RRs) (95% CIs) and potential confounders in the fully adjusted models. Disagreements were resolved by a third reviewer who checked and consolidated the data. The quality of included studies was assessed by the Newcastle-Ottawa scale.[2] For each type of dairy product, a meta-analysis was conducted if more than 3 studies were selected. Low-fat dairy included low-fat milk, low-fat yogurt, low-fat cheese and low-fat ice cream. High-fat dairy included high-fat milk, high-fat yogurt, high-fat cheese, and cream or butter, which was

consistent with previous studies[3-5]. As the intake of low-fat cheese cannot be ignored in the UKB (Supplementary Fig. 1), we separately analyzed low-fat and high-fat cheese, categorizing them accordingly in analyses of low-fat and high-fat dairy products in meta-analysis. A random-effect model based on the DerSimonian and Laird method[6] was applied to calculate summary RRs and 95% CIs comparing the highest with the lowest category of intake. If no significant heterogeneity was found, we also conducted a fixed-effect model to calculate summary HRs and 95% CIs comparing the highest with the lowest category. For dose-response meta-analysis, standard serving sizes were applied to each type of dairy product according to the definition in primary studies. In terms of dairy consumption, the median or midpoint of the lower and upper bounds were assigned to each category. If the lower or upper bound was not reported, we estimated it by multiplying the upper bound by 0.5 for the lowest category and the lower bound by 1.75 for the highest category.[7] When studies only reported RRs and 95% CIs for CHD and stroke but not total CVD, we pooled it by using fixed effect meta-analysis. If studies reported results separately for sex or different cohorts, they were included as separate studies. The extent of heterogeneity was assessed by I^2 (ranging from 0% to 100%, >50% indicates heterogeneity among studies, >80% indicates severe heterogeneity among studies) and Cochran's Q statistic test (significant at $P<0.10$)[8,9]. Potential publication bias was assessed using Begg's test, Egger's test, and visual funnel plot asymmetry. We assessed the confidence of evidence using the GRADE approach, categorizing it into four levels: very low, low, moderate, and high[10-12]. All statistical analyses for the meta-analysis were conducted using Stata version 16.0 (StataCorp). Absolute risk values were calculated with GRADEpro software.



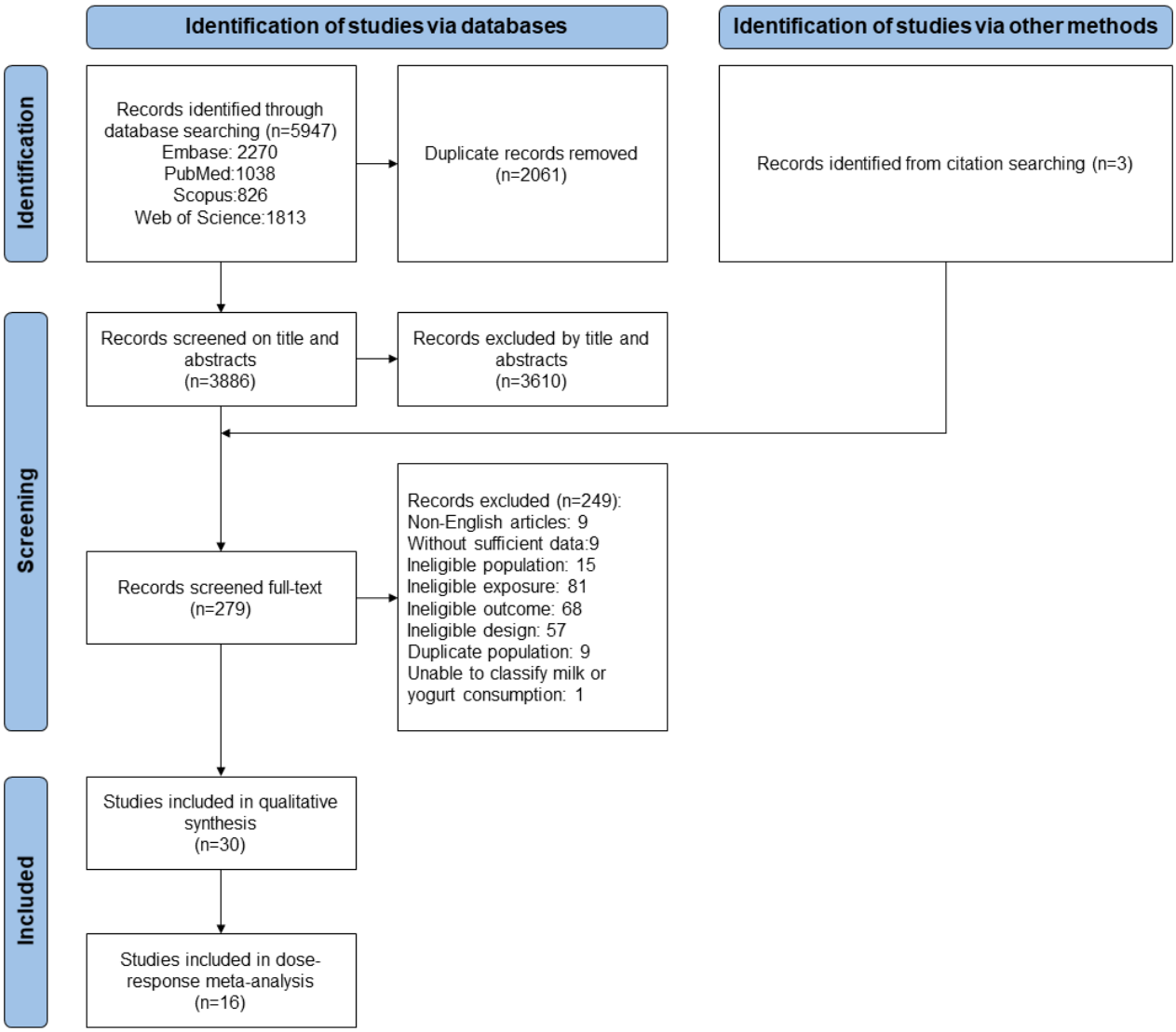
Supplementary Figure 1. Consumption of different types of dairy products/cheese in the UK Biobank.

(A) Total dairy. (B) Cheese. Source data are provided as a Source Data file.

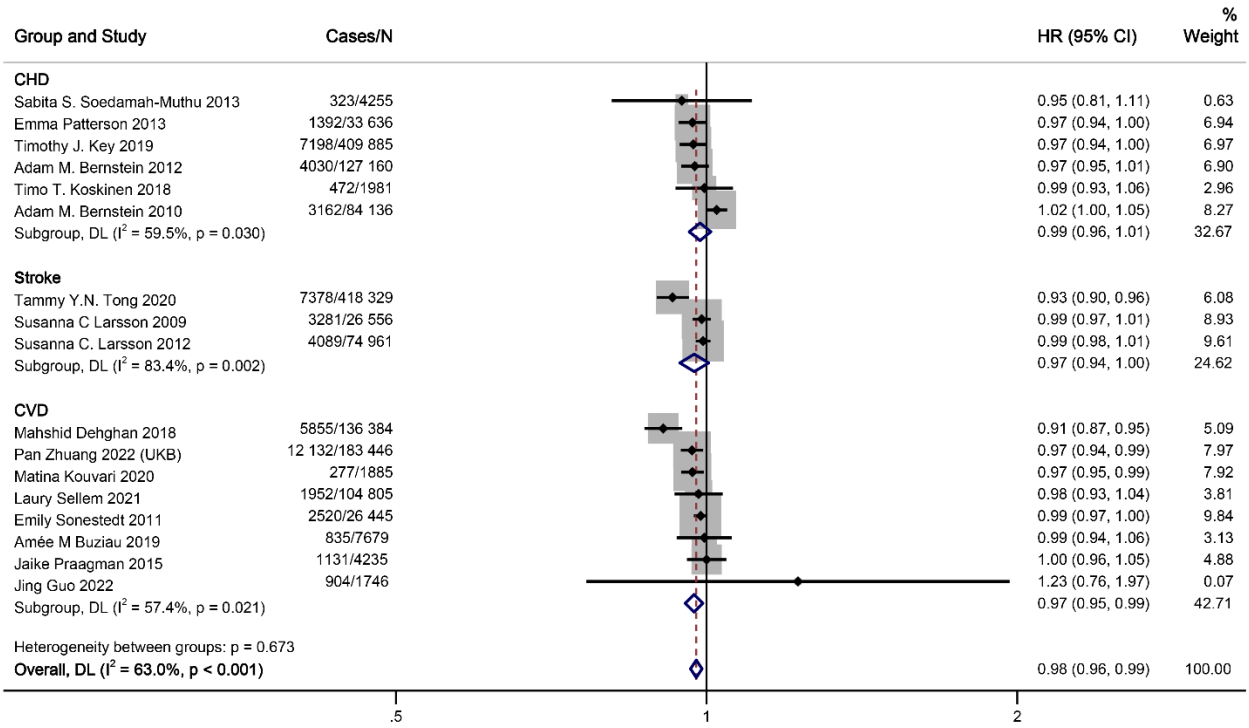


Supplementary Figure 2. Statistical model-based hazard ratios and 95% confidence intervals for incident cardiovascular disease, coronary heart diseases, stroke associated with replacement of one serving per day of other major protein sources with one serving per day of dairy products in China Kadoorie Biobank and UK Biobank.

Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Multi-variable Cox proportional hazard model was used. The multi-variable model was adjusted for age, sex, study area, survey season, BMI, education, income, physical activity, smoking, alcohol drinking, history of hypertension, diabetes, family history of CVD, aspirin use, vitamins use, minerals use, and intake of red meat, fish, poultry, vegetables, fruits, and eggs in the China Kadoorie Biobank. The multi-variable model was adjusted for age, sex, centers, survey season, BMI, education, household income, physical activity, smoking, alcohol drinking, history of hypertension, diabetes, family history of CVD, aspirin use, vitamins use, minerals use, and intake of red meat, processed red meat, oily fish, non-oily fish, poultry, vegetables, fruits, and eggs in the UK Biobank. CHD, coronary heart disease; CI, confidence interval; CKB, China Kadoorie Biobank; CVD, cardiovascular disease; HR, hazard ratio; UKB, UK Biobank. Source data are provided as a Source Data file.

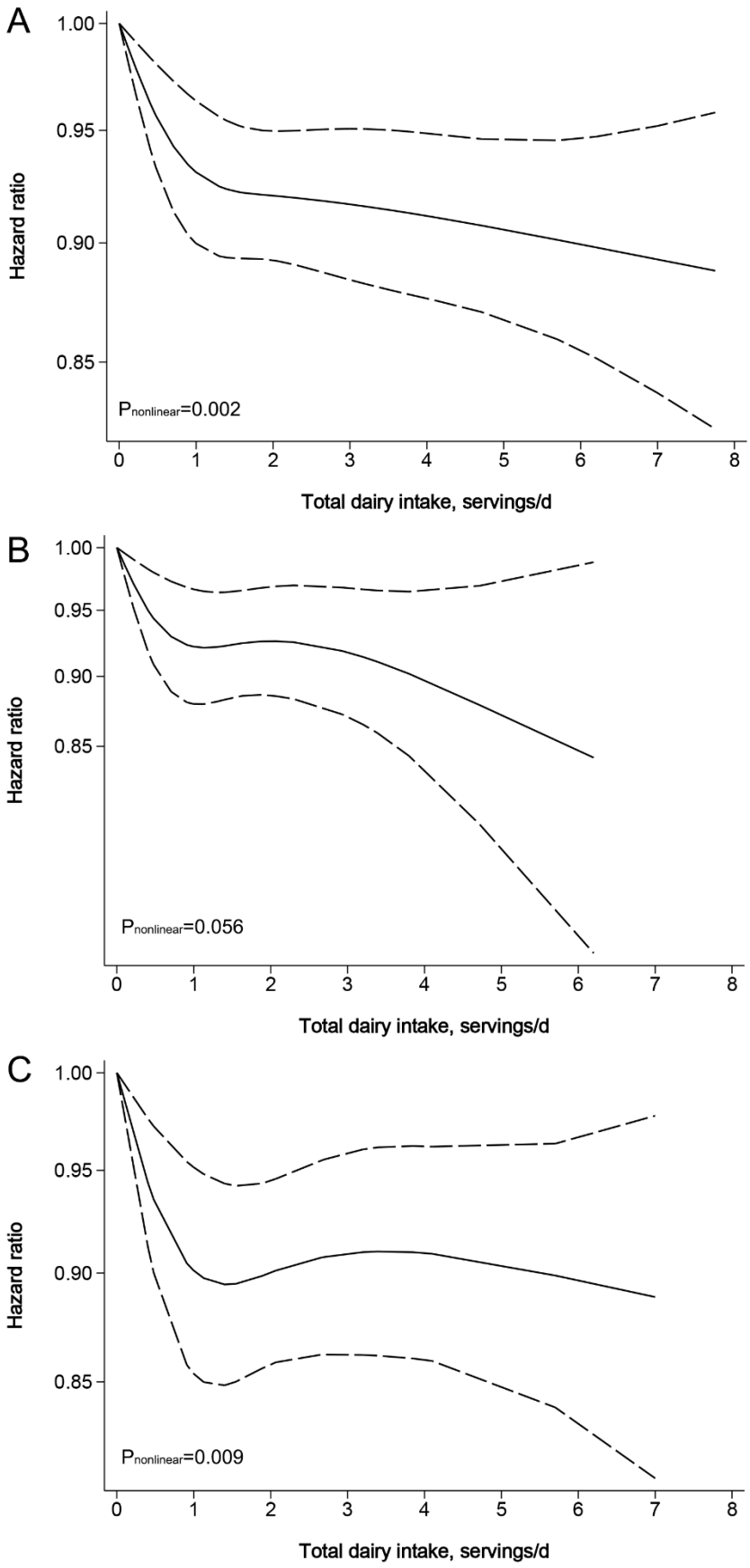


Supplementary Figure 3. Flow chart for participated study selection in the meta-analysis.

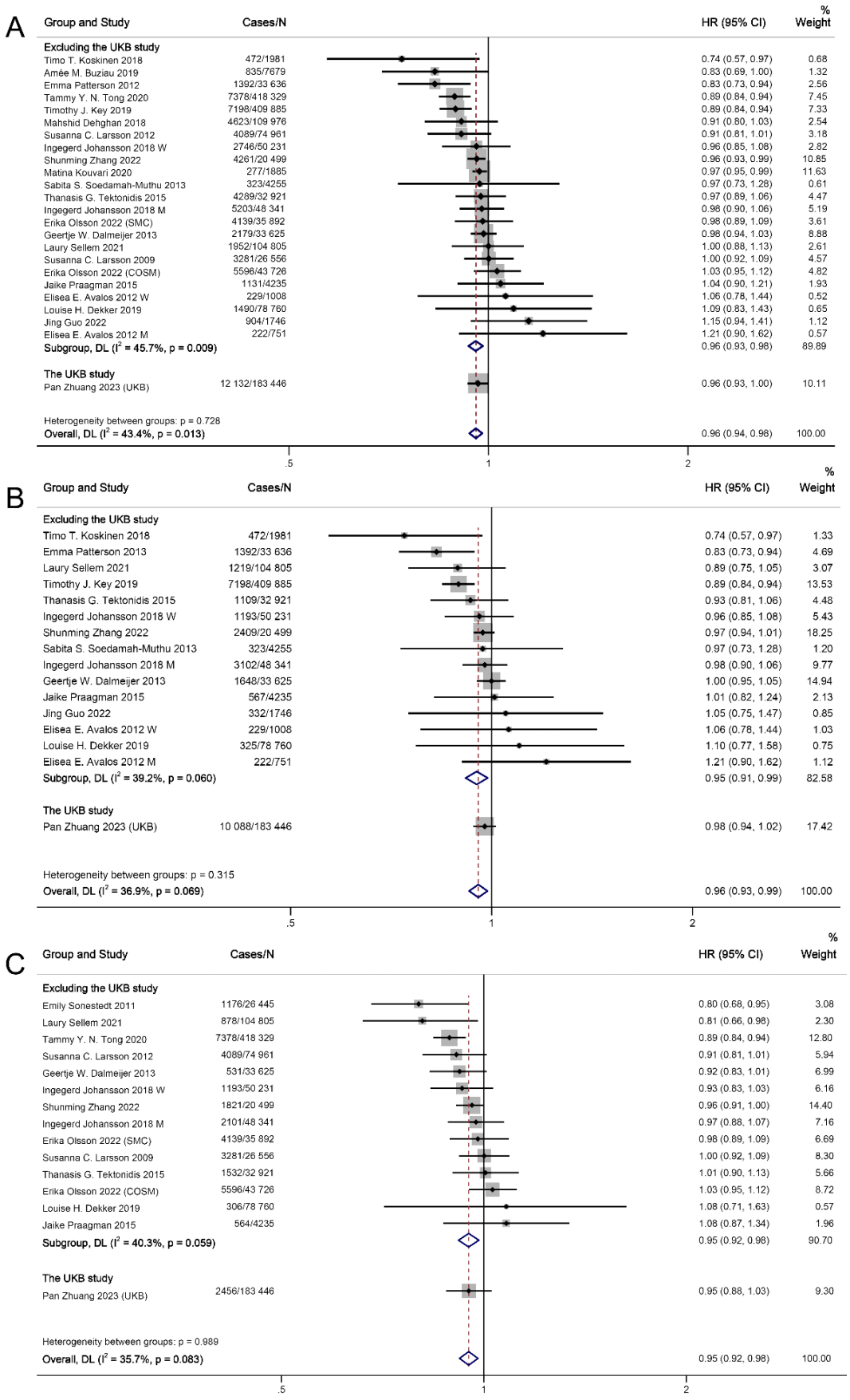


Supplementary Figure 4. Association of total dairy product consumption with CVD risk for 1 serving per day increase using random-effects meta-analysis.

Meta-analysis pooling of aggregate data using the random-effects inverse-variance model with DerSimonian-Laird estimate of tau². Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Squares represent study-specific HRs. Horizontal lines denote 95% CIs. Gray square areas are proportional to the individual study weight for the overall meta-analysis. The red dotted line represents risk ratio of pooled meta-analysis. The blue hollow diamonds represent the results of the meta-analysis for each group, with the center indicating the risk ratio and the width representing the 95% CI. I^2 refers to the proportion of heterogeneity among studies. UKB, UK Biobank. Source data are provided as a Source Data file.

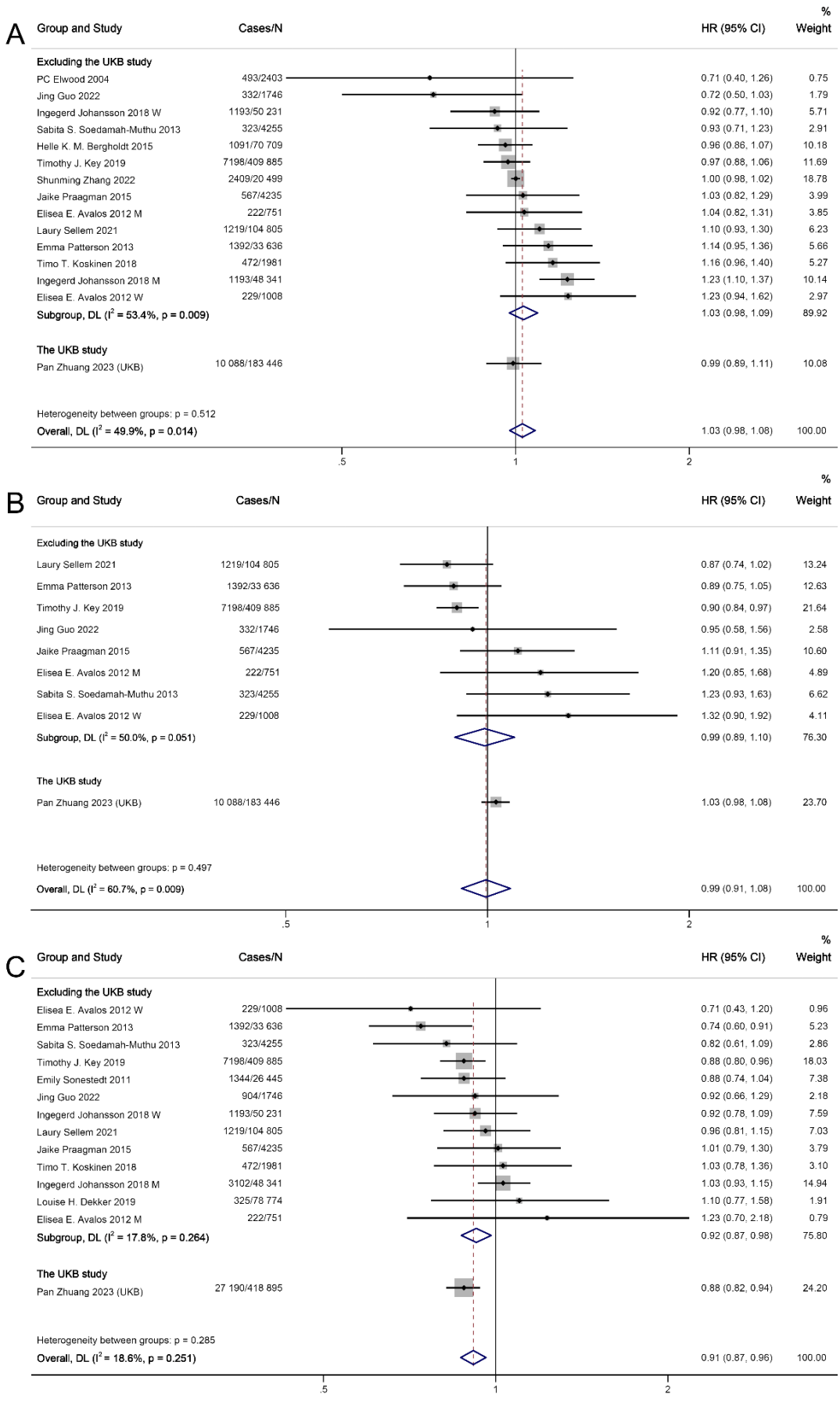


Supplementary Figure 5. Dose-response of total dairy product consumption with CVD risk. (A) Cardiovascular disease. (B) Coronary heart disease. (C) Stroke. Solid line represents non-linear dose response and dotted lines represent 95% confidence interval. Source data are provided as a Source Data file.



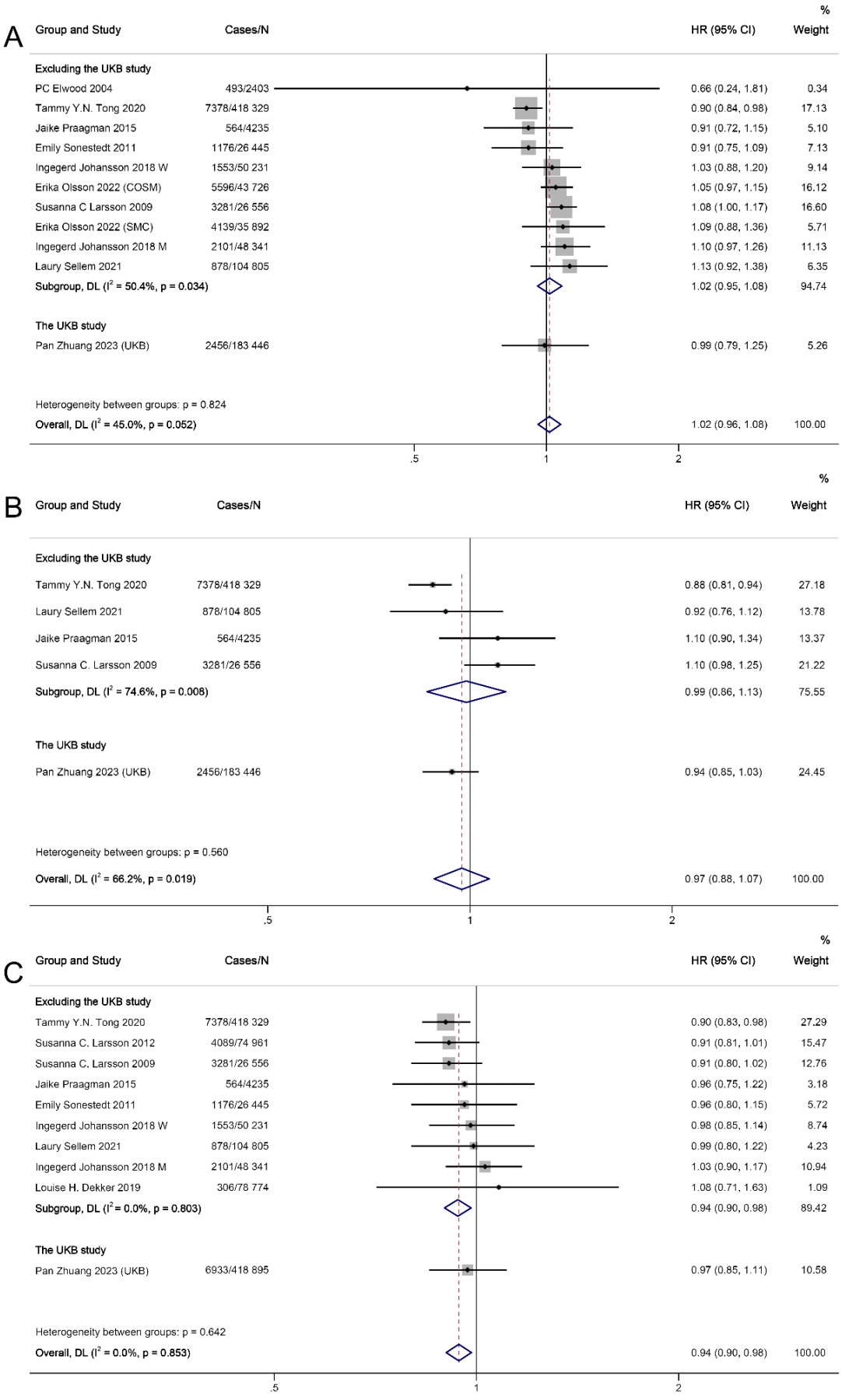
Supplementary Figure 6. Association of fermented dairy consumption with cardiovascular disease, coronary heart disease, and stroke risk for high compared with low intake using random-effects meta-analysis.

(A) Cardiovascular disease. (B) Coronary heart disease. (C) Stroke. Meta-analysis pooling of aggregate data using the random-effects inverse-variance model with DerSimonian-Laird estimate of tau^2. Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Squares represent study-specific HRs. Horizontal lines denote 95% CIs. Gray square areas are proportional to the individual study weight for the overall meta-analysis. The red dotted line represents risk ratio of pooled meta-analysis. The blue hollow diamonds represent the results of the meta-analysis for each group, with the center indicating the risk ratio and the width representing the 95% CI. I^2 refers to the proportion of heterogeneity among studies. M, men; W, women; UKB, UK Biobank. Source data are provided as a Source Data file.



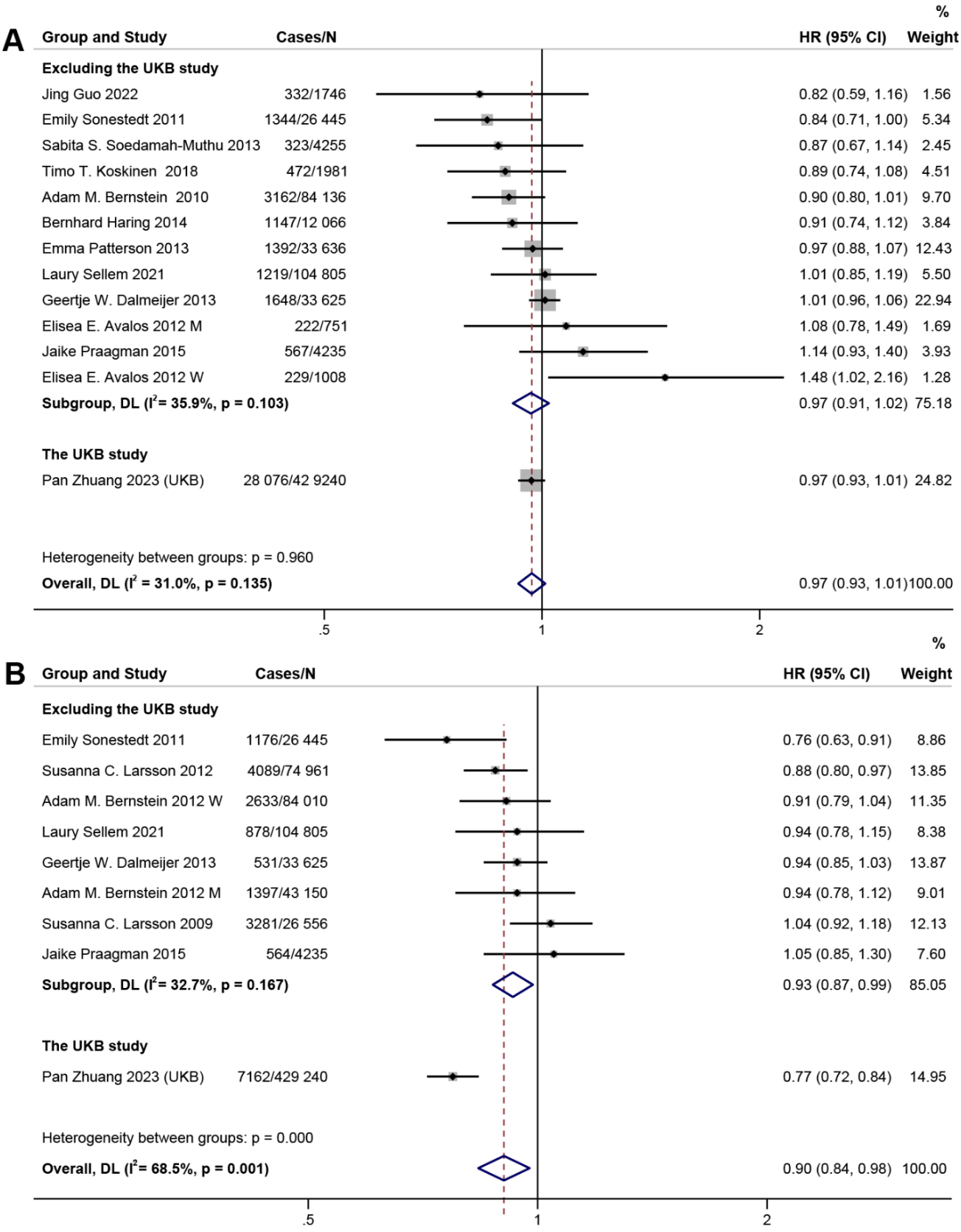
Supplementary Figure 7. Association of milk, yogurt, cheese consumption with coronary heart disease risk for high compared with low intake using random-effects meta-analysis.

(A) Milk. (B) Yogurt. (C) Cheese. Meta-analysis pooling of aggregate data using the random-effects inverse-variance model with DerSimonian-Laird estimate of tau^2. Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Squares represent study-specific HRs. Horizontal lines denote 95% CIs. Gray square areas are proportional to the individual study weight for the overall meta-analysis. The red dotted line represents risk ratio of pooled meta-analysis. The blue hollow diamonds represent the results of the meta-analysis for each group, with the center indicating the risk ratio and the width representing the 95% CI. I^2 refers to the proportion of heterogeneity among studies. M, men; W, women; UKB, UK Biobank. Source data are provided as a Source Data file.



Supplementary Figure 8. Association of milk, yogurt, cheese consumption with stroke risk for high compared with low intake using random-effects meta-analysis.

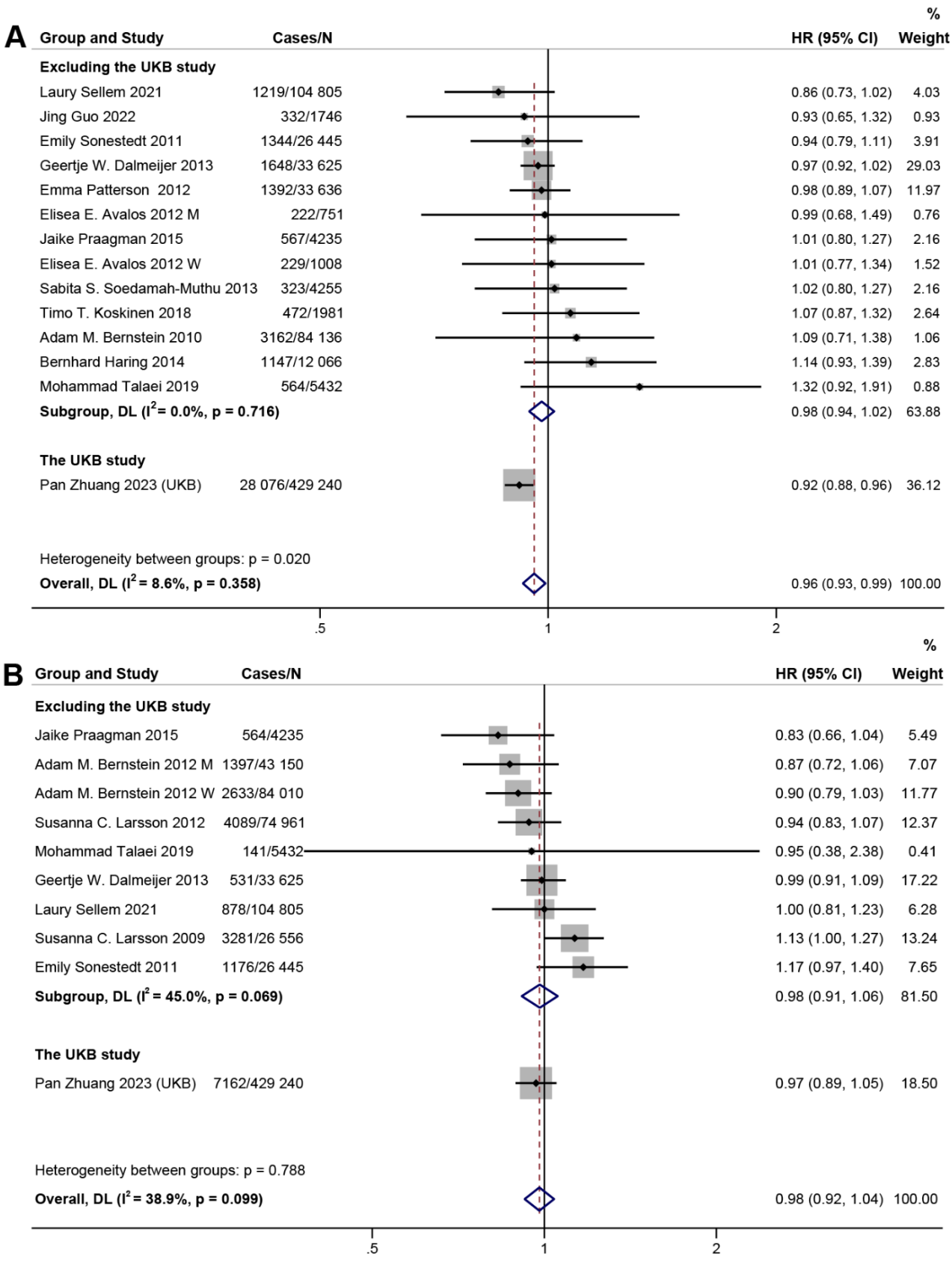
(A) Milk. (B) Yogurt. (C) Cheese. Meta-analysis pooling of aggregate data using the random-effects inverse-variance model with DerSimonian-Laird estimate of tau^2. Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Squares represent study-specific HRs. Horizontal lines denote 95% CIs. Gray square areas are proportional to the individual study weight for the overall meta-analysis. The red dotted line represents risk ratio of pooled meta-analysis. The blue hollow diamonds represent the results of the meta-analysis for each group, with the center indicating the risk ratio and the width representing the 95% CI. I^2 refers to the proportion of heterogeneity among studies. M, men; W, women; UKB, UK Biobank. Source data are provided as a Source Data file.



Supplementary Figure 9. Association of low-fat dairy consumption with coronary heart disease and stroke risk for high compared with low intake using random-effects meta-analysis.

(A) Coronary heart disease. (B) Stroke. Meta-analysis pooling of aggregate data using the random-effects inverse-variance model with DerSimonian-Laird estimate of tau². Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Squares represent study-specific HRs. Horizontal lines denote 95% CIs. Gray square areas

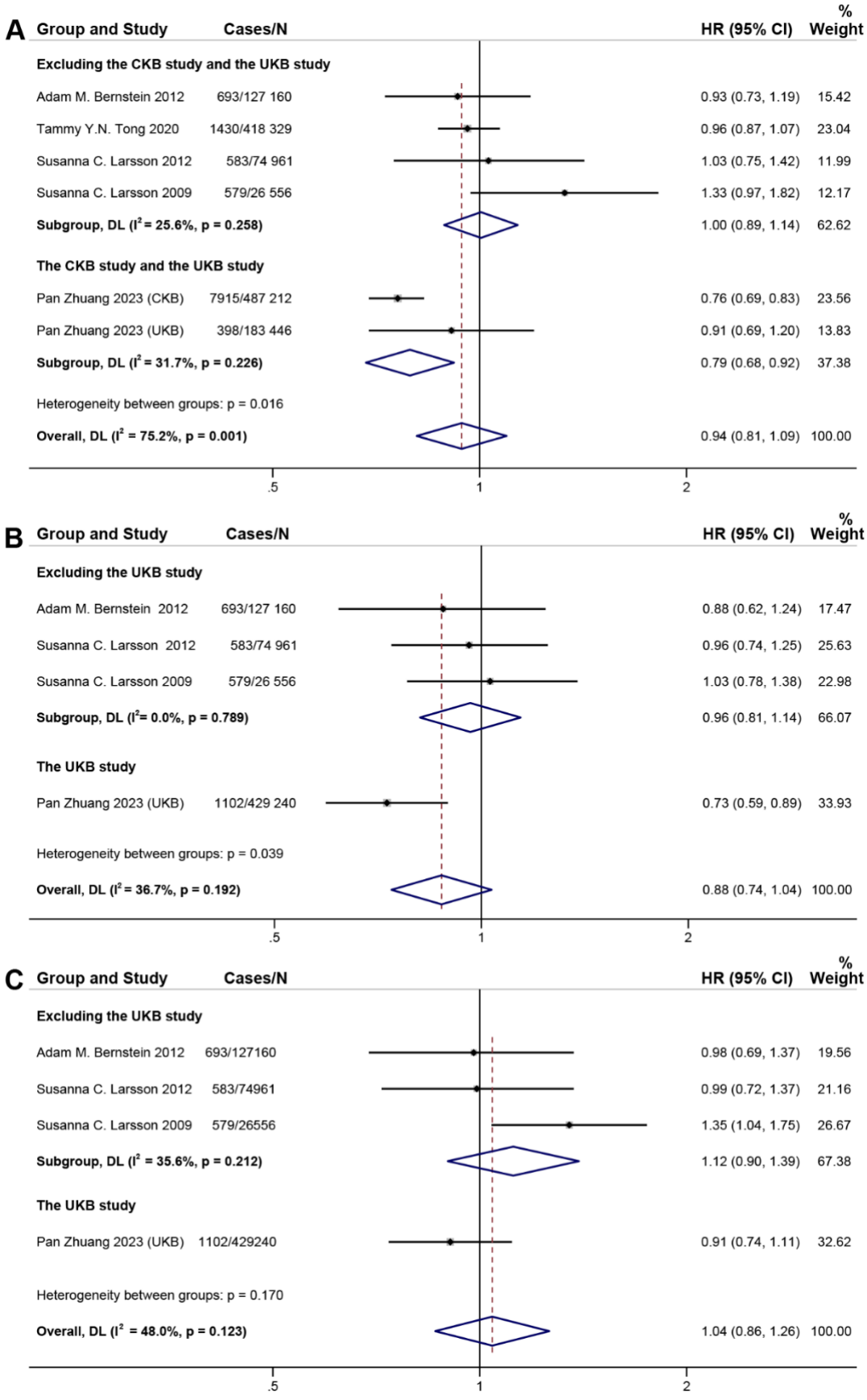
are proportional to the individual study weight for the overall meta-analysis. The red dotted line represents risk ratio of pooled meta-analysis. The blue hollow diamonds represent the results of the meta-analysis for each group, with the center indicating the risk ratio and the width representing the 95% CI. I^2 refers to the proportion of heterogeneity among studies. M, men; W, women; UKB, UK Biobank. Source data are provided as a Source Data file.



Supplementary Figure 10. Association of high-fat dairy consumption with coronary heart disease and stroke risk for high compared with low intake using random-effects meta-analysis.

(A) Coronary heart disease. (B) Stroke. Meta-analysis pooling of aggregate data using the random-effects inverse-variance model with DerSimonian-Laird estimate of tau². Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Squares represent study-specific HRs. Horizontal lines denote 95% CIs. Gray square areas

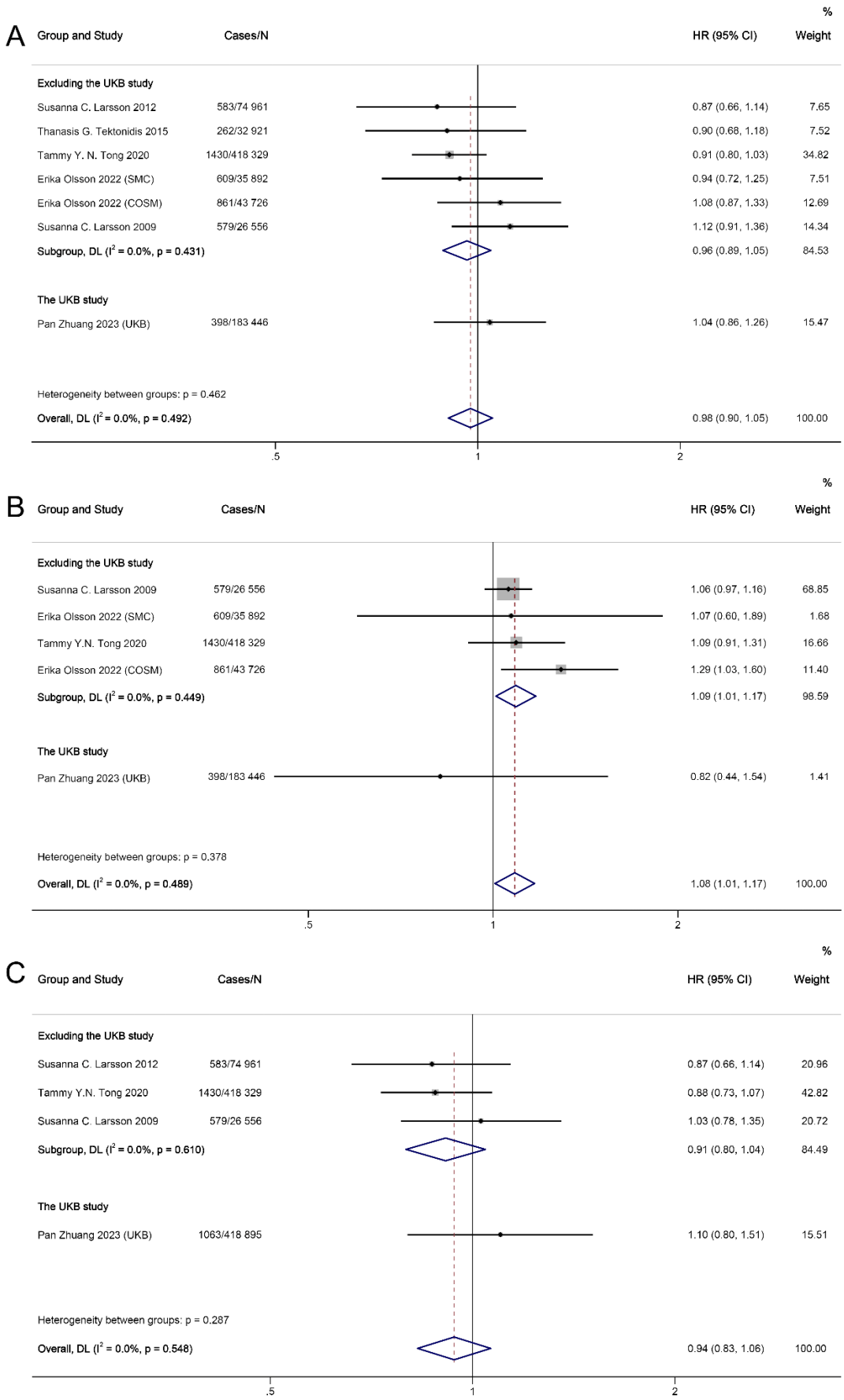
are proportional to the individual study weight for the overall meta-analysis. The red dotted line represents risk ratio of pooled meta-analysis. The blue hollow diamonds represent the results of the meta-analysis for each group, with the center indicating the risk ratio and the width representing the 95% CI. I^2 refers to the proportion of heterogeneity among studies. M, men; W, women; UKB, UK Biobank. Source data are provided as a Source Data file.



Supplementary Figure 11. Association of total, low-fat, high-fat dairy consumption with hemorrhagic stroke risk for high compared with low intake using random-effects meta-analysis.

(A) Total dairy. (B) Low-fat dairy. (C) High-fat dairy. Meta-analysis pooling of aggregate data using the random-effects

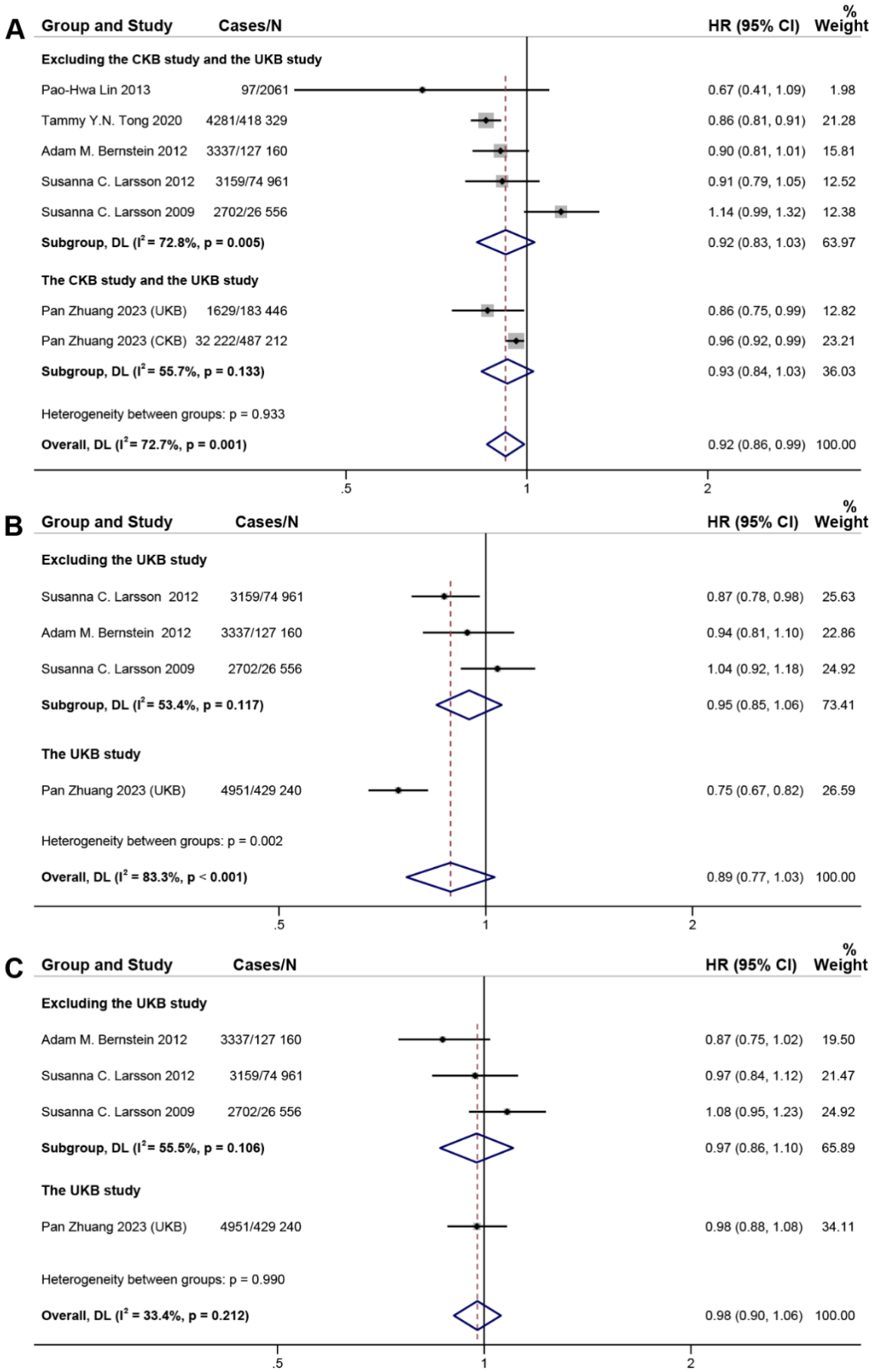
inverse-variance model with DerSimonian-Laird estimate of tau^2. Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Squares represent study-specific HRs. Horizontal lines denote 95% CIs. Gray square areas are proportional to the individual study weight for the overall meta-analysis. The red dotted line represents risk ratio of pooled meta-analysis. The blue hollow diamonds represent the results of the meta-analysis for each group, with the center indicating the risk ratio and the width representing the 95% CI. I^2 refers to the proportion of heterogeneity among studies. M, men; W, women; UKB, UK Biobank. Source data are provided as a Source Data file.



Supplementary Figure 12. Association of fermented dairy, milk, cheese consumption with hemorrhagic stroke risk for high compared with low intake using random-effects meta-analysis.

(A) Fermented dairy. (B) Milk. (C) Cheese. Meta-analysis pooling of aggregate data using the random-effects inverse-

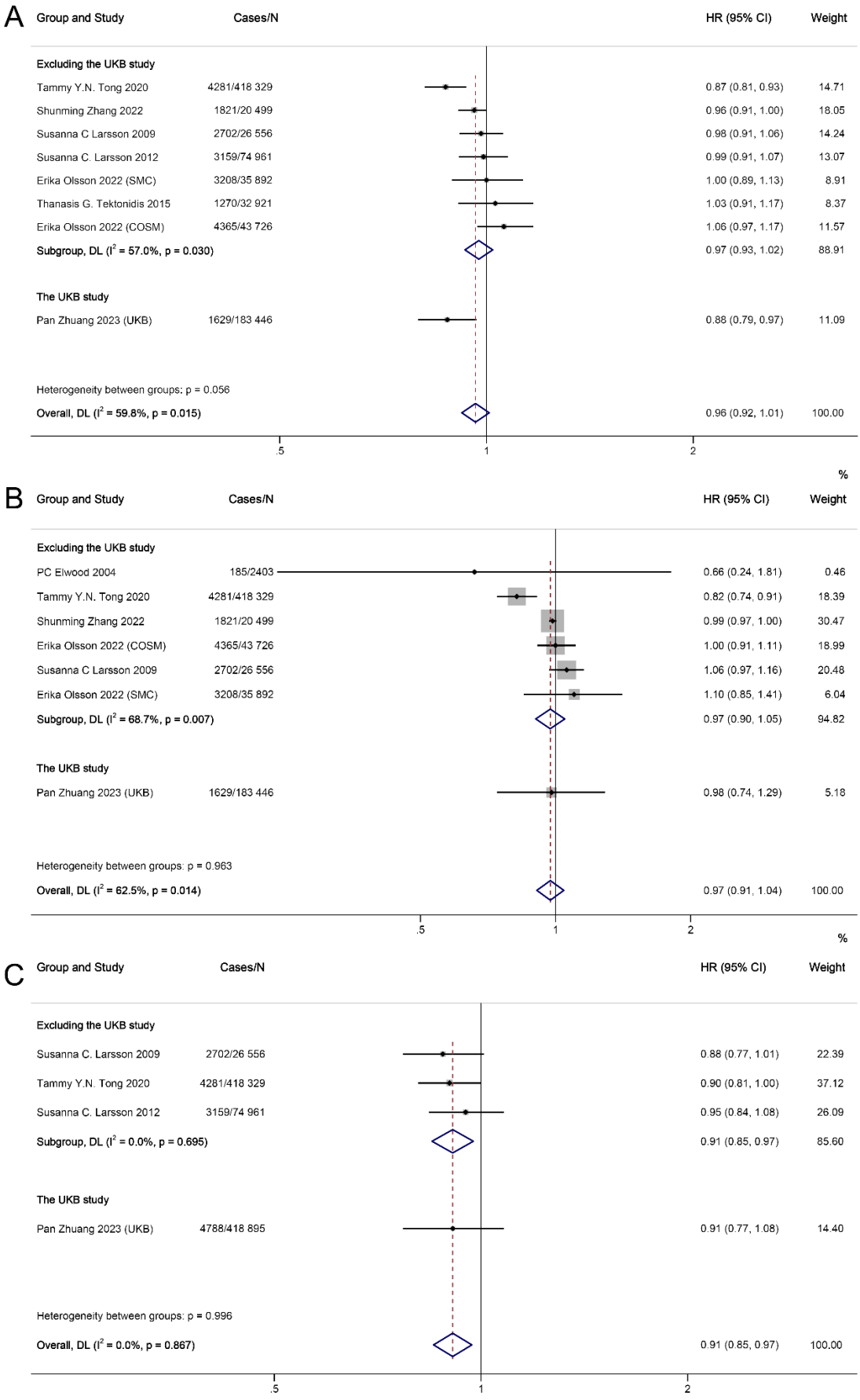
variance model with DerSimonian-Laird estimate of tau^2. Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Squares represent study-specific HRs. Horizontal lines denote 95% CIs. Gray square areas are proportional to the individual study weight for the overall meta-analysis. The red dotted line represents risk ratio of pooled meta-analysis. The blue hollow diamonds represent the results of the meta-analysis for each group, with the center indicating the risk ratio and the width representing the 95% CI. I^2 refers to the proportion of heterogeneity among studies. M, men; W, women; UKB, UK Biobank. Source data are provided as a Source Data file.



Supplementary Figure 13. Association of total, low-fat, high-fat dairy consumption with ischemic stroke risk for high compared with low intake using random-effects meta-analysis.

(A) Total dairy. (B) Low-fat dairy. (C) High-fat dairy. Meta-analysis pooling of aggregate data using the random-effects

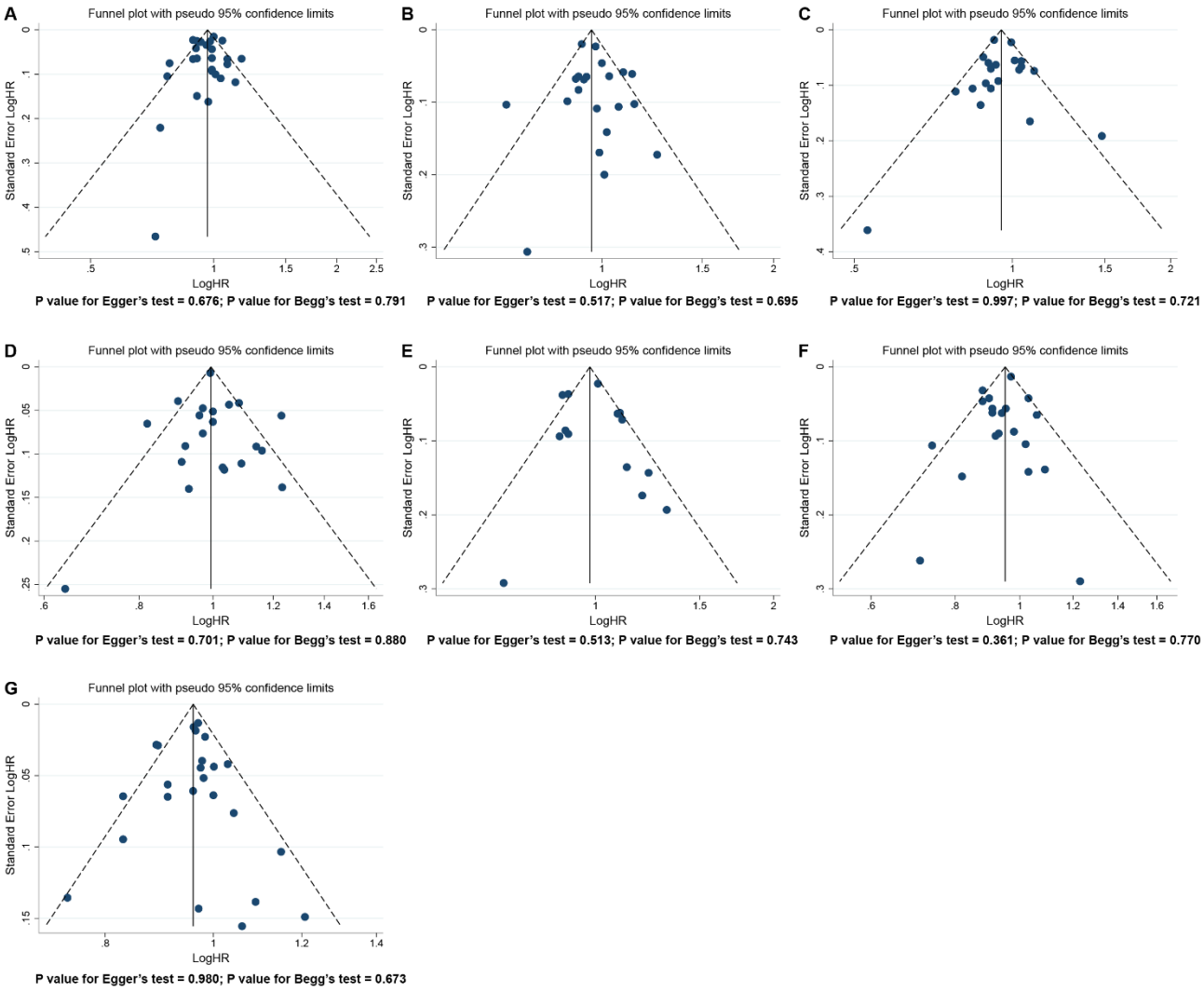
inverse-variance model with DerSimonian-Laird estimate of tau^2. Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Squares represent study-specific HRs. Horizontal lines denote 95% CIs. Gray square areas are proportional to the individual study weight for the overall meta-analysis. The red dotted line represents risk ratio of pooled meta-analysis. The blue hollow diamonds represent the results of the meta-analysis for each group, with the center indicating the risk ratio and the width representing the 95% CI. I^2 refers to the proportion of heterogeneity among studies. UKB, UK Biobank. Source data are provided as a Source Data file.



Supplementary Figure 14. Association of fermented dairy, milk, cheese consumption with ischemic stroke risk for high compared with low intake using random-effects meta-analysis.

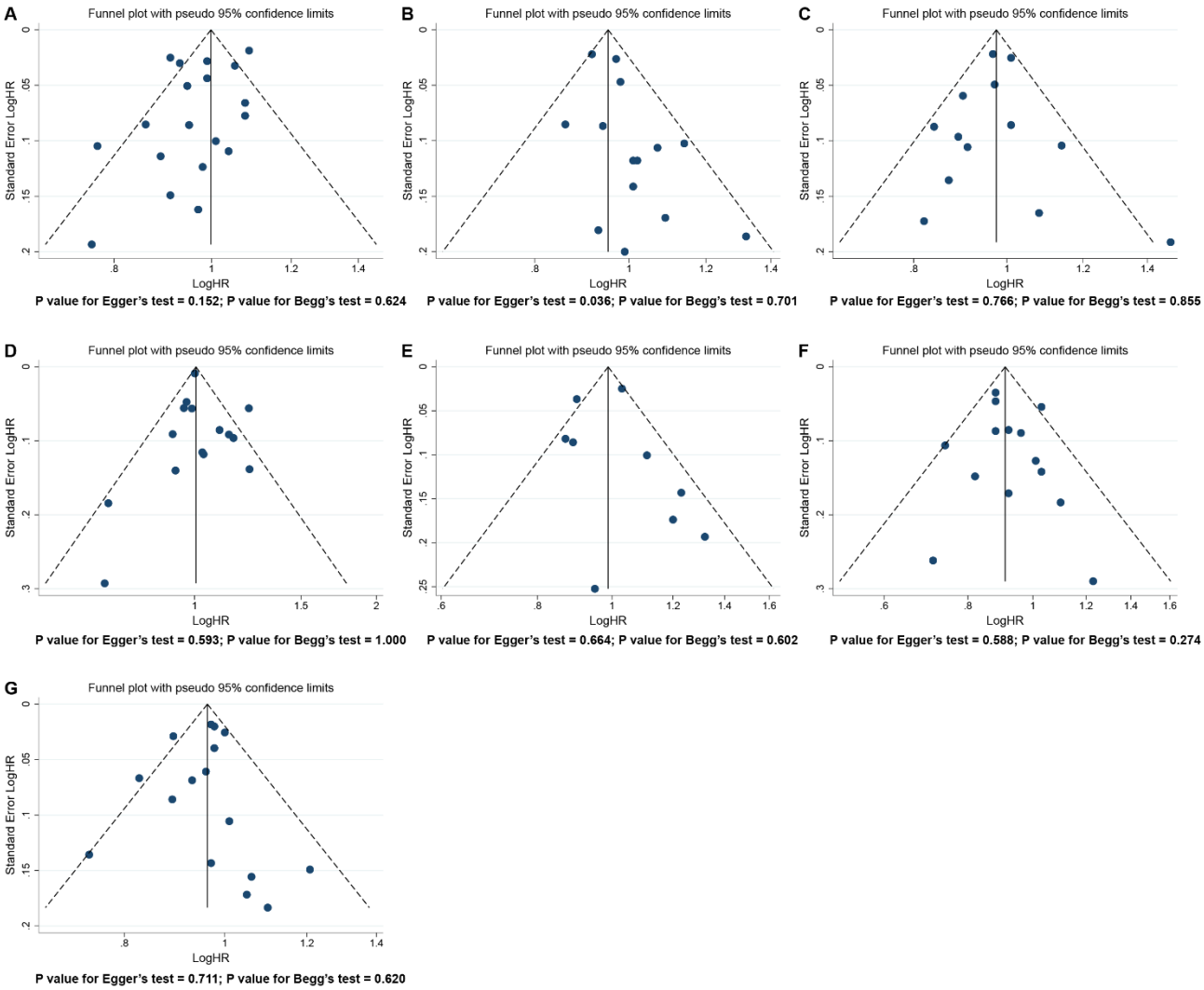
(A) Fermented dairy. (B) Milk. (C) Cheese. Meta-analysis pooling of aggregate data using the random-effects inverse-

variance model with DerSimonian-Laird estimate of tau^2. Data are presented as hazard ratios (HRs) and 95% confidence intervals (CIs). Squares represent study-specific HRs. Horizontal lines denote 95% CIs. Gray square areas are proportional to the individual study weight for the overall meta-analysis. The red dotted line represents risk ratio of pooled meta-analysis. The blue hollow diamonds represent the results of the meta-analysis for each group, with the center indicating the risk ratio and the width representing the 95% CI. I^2 refers to the proportion of heterogeneity among studies. M, men; W, women; UKB, UK Biobank. Source data are provided as a Source Data file.



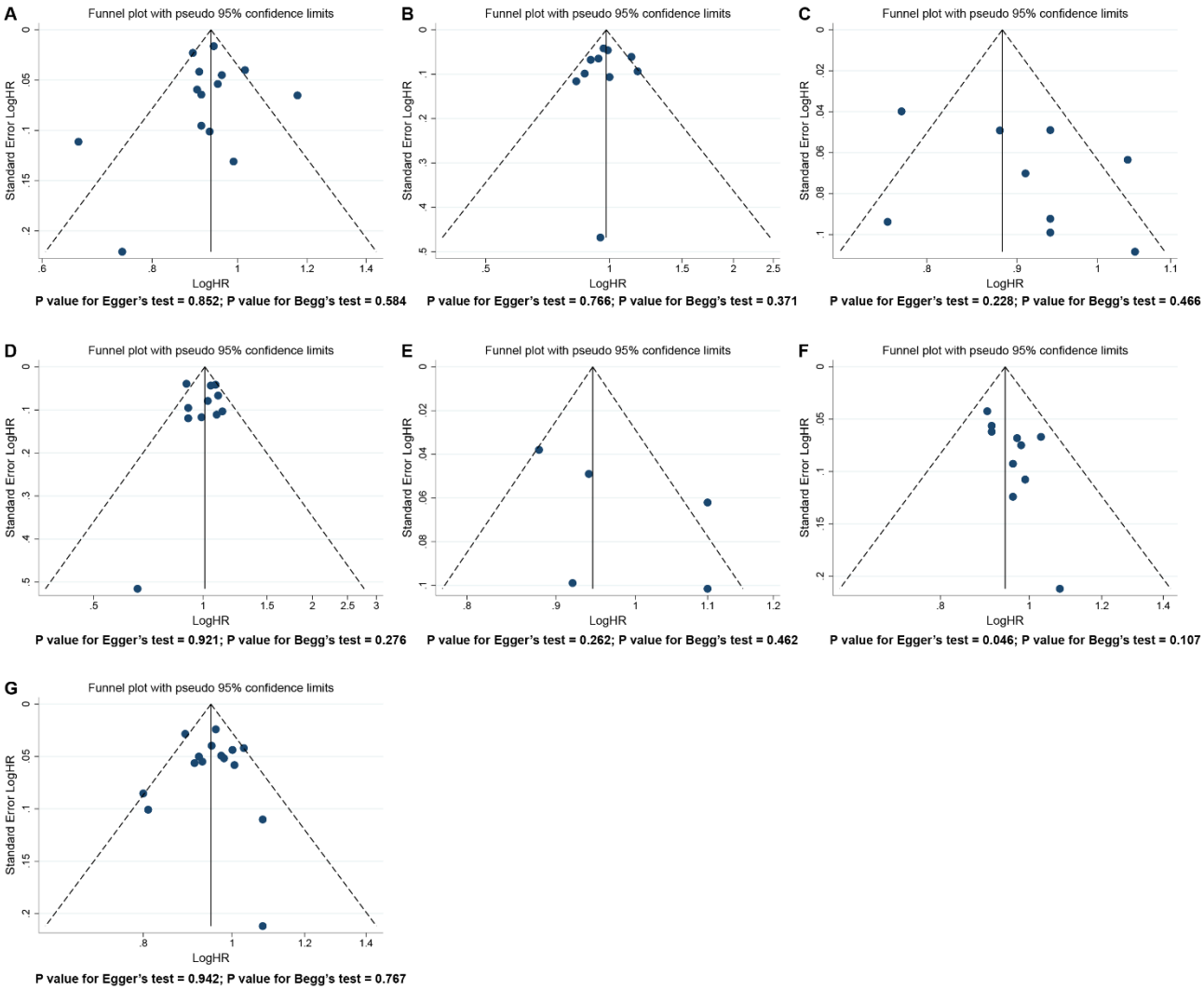
Supplementary Figure 15. Funnel plot for assessment of publication bias for the association between total dairy and subtypes of dairy consumption and risk of cardiovascular disease.

(A) Total dairy. (B) High-fat dairy. (C) Low-fat-dairy. (D) Milk. (E) yogurt. (F) Cheese. (G) Fermented dairy. Begg's Test and Egger's test was used for assessment of publication bias. The blue dots represent each study included in the analysis. Source data are provided as a Source Data file.



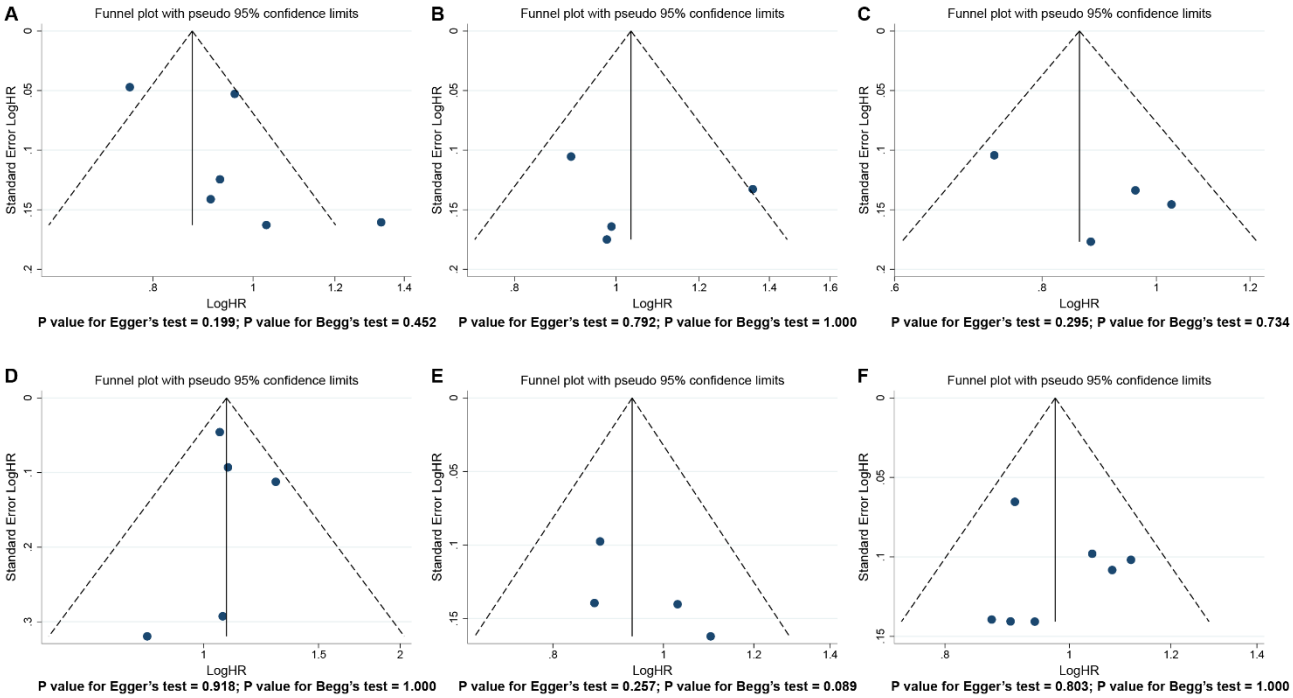
Supplementary Figure 16. Funnel plot for assessment of publication bias for the association between total dairy and subtypes of dairy consumption and risk of coronary heart disease.

(A) Total dairy. (B) High-fat dairy. (C) Low-fat-dairy. (D) Milk. (E) Yogurt. (F) Cheese. (G) Fermented dairy. Begg's Test and Egger's test was used for assessment of publication bias. The blue dots represent each study included in the analysis. Source data are provided as a Source Data file.



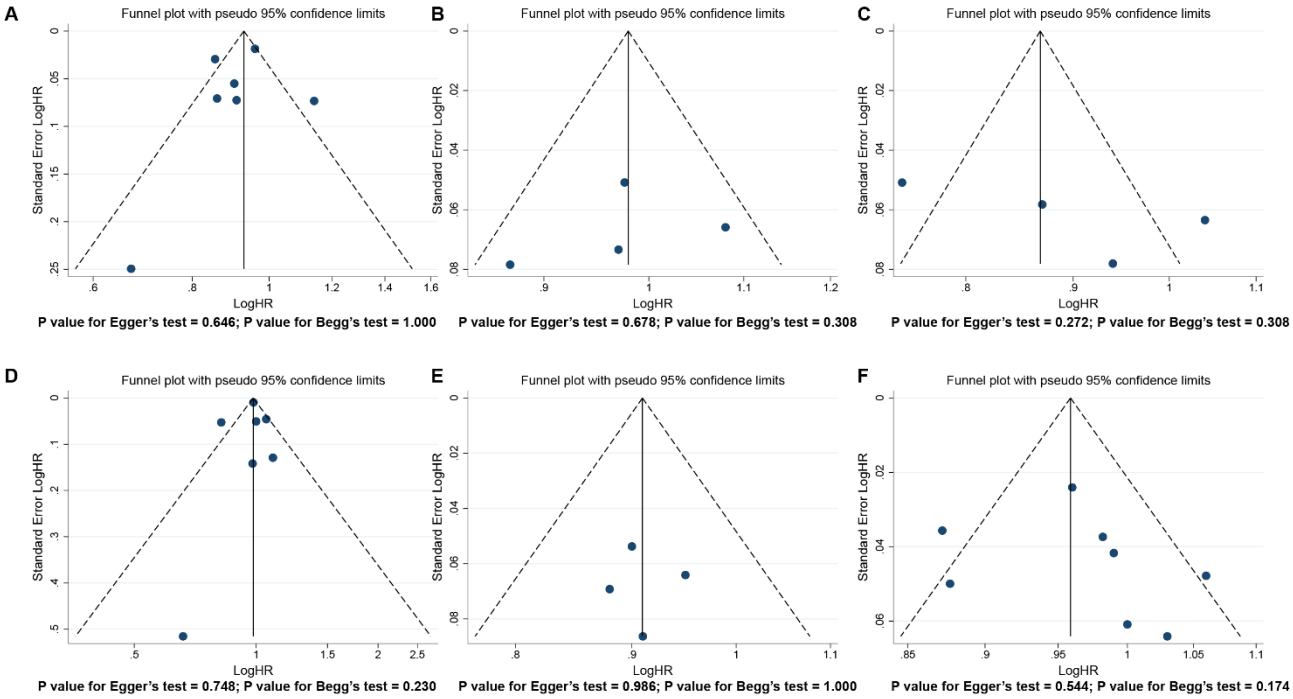
Supplementary Figure 17. Funnel plot for assessment of publication bias for the association between total dairy and subtypes of dairy consumption and risk of stroke.

(A) Total dairy. (B) High-fat dairy. (C) Low-fat-dairy. (D) Milk. (E) yogurt. (F) Cheese. (G) Fermented dairy. Begg's Test and Egger's test was used for assessment of publication bias. The blue dots represent each study included in the analysis. Source data are provided as a Source Data file.



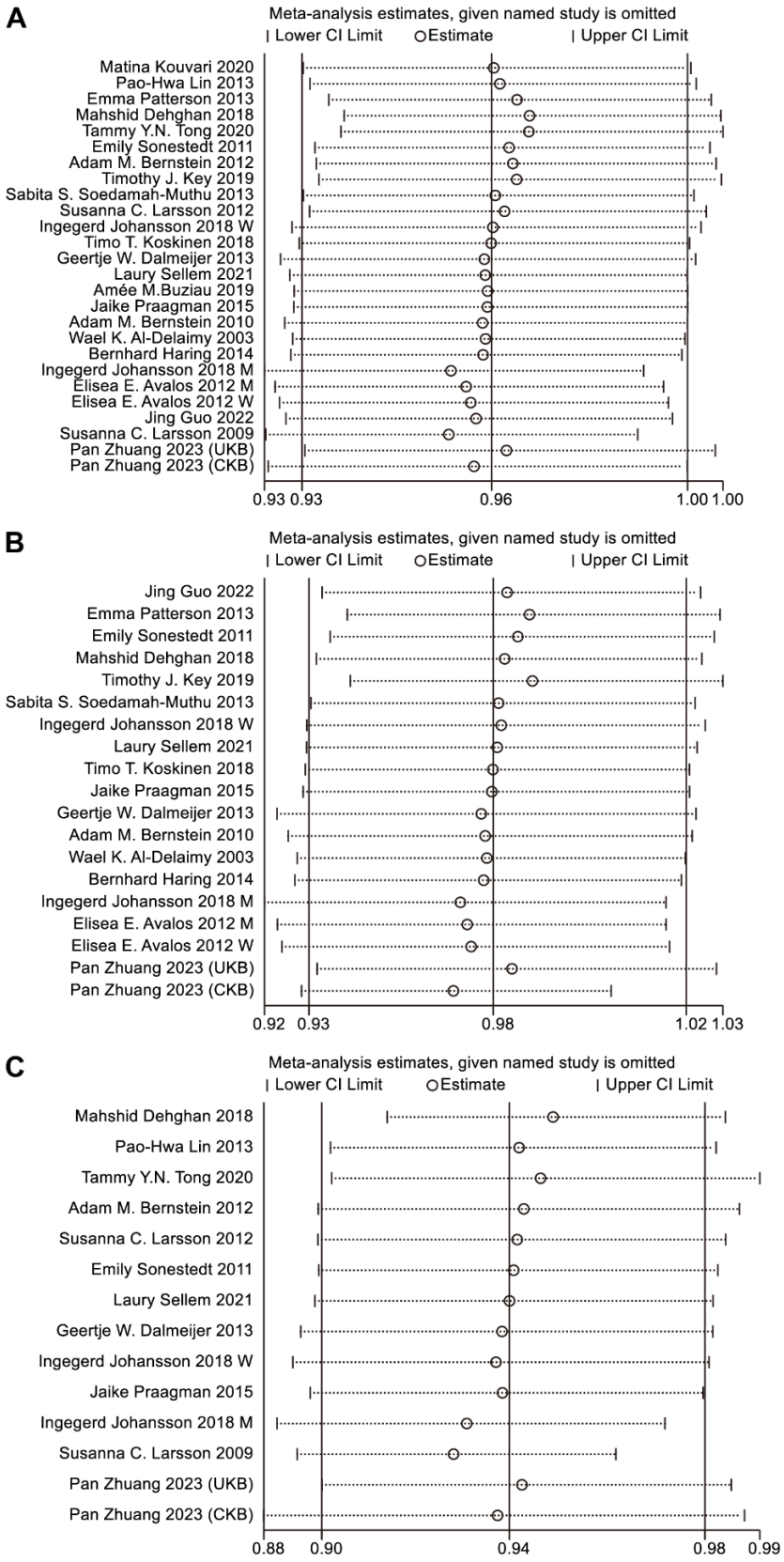
Supplementary Figure 18. Funnel plot for assessment of publication bias for the association between total dairy and subtypes of dairy consumption and risk of hemorrhagic stroke.

(A) Total dairy. (B) High-fat dairy. (C) Low-fat-dairy. (D) Milk. (E) Cheese. (F) Fermented dairy**.** Begg's Test and Egger's test was used for assessment of publication bias. The blue dots represent each study included in the analysis. Source data are provided as a Source Data file.



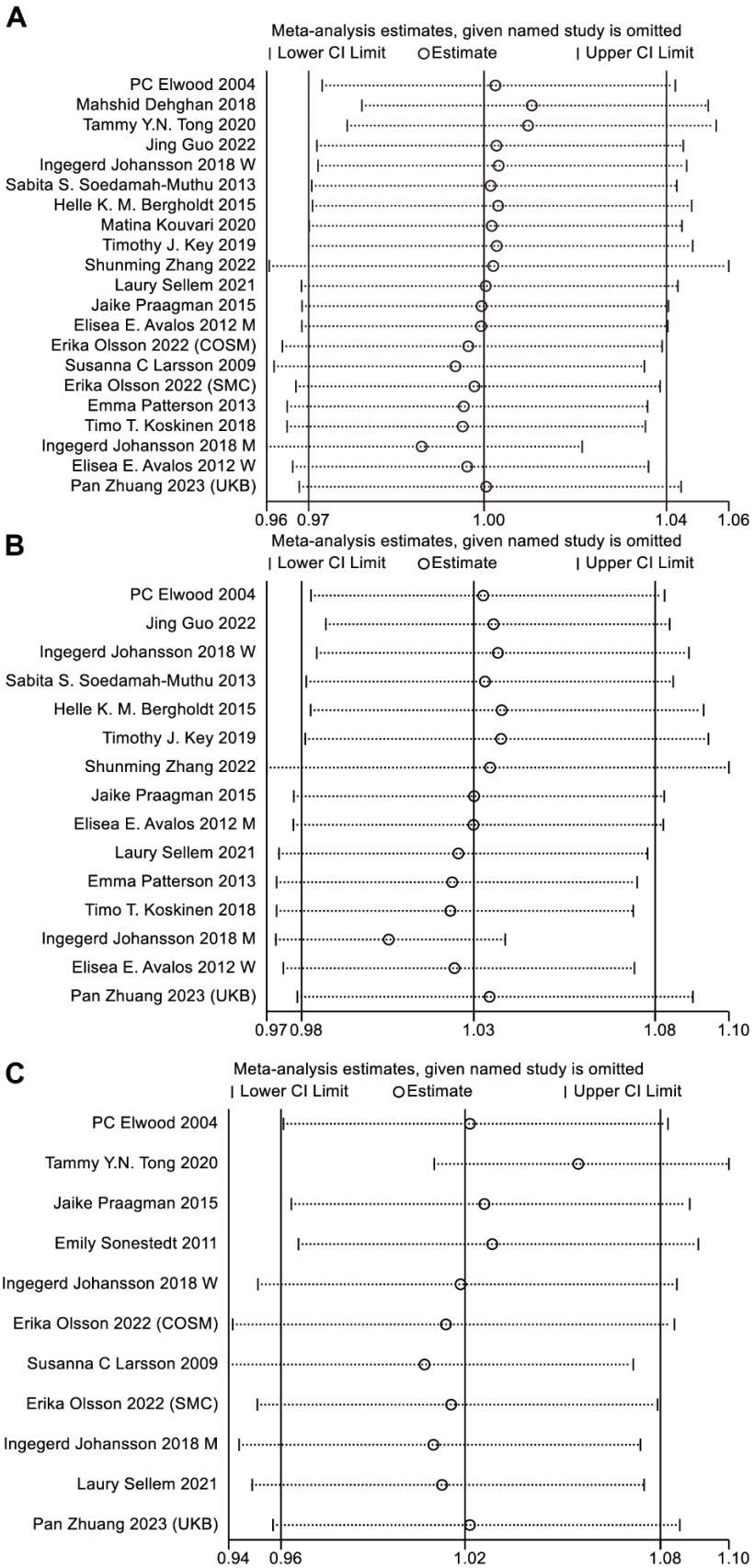
Supplementary Figure 19. Funnel plot for assessment of publication bias for the association between total dairy and subtypes of dairy consumption and risk of ischemic stroke.

(A) Total dairy. (B) High-fat dairy. (C) Low-fat-dairy. (D) Milk. (E) Cheese. (F) Fermented dairy**.** Begg's Test and Egger's test was used for assessment of publication bias. The blue dots represent each study included in the analysis. Source data are provided as a Source Data file.



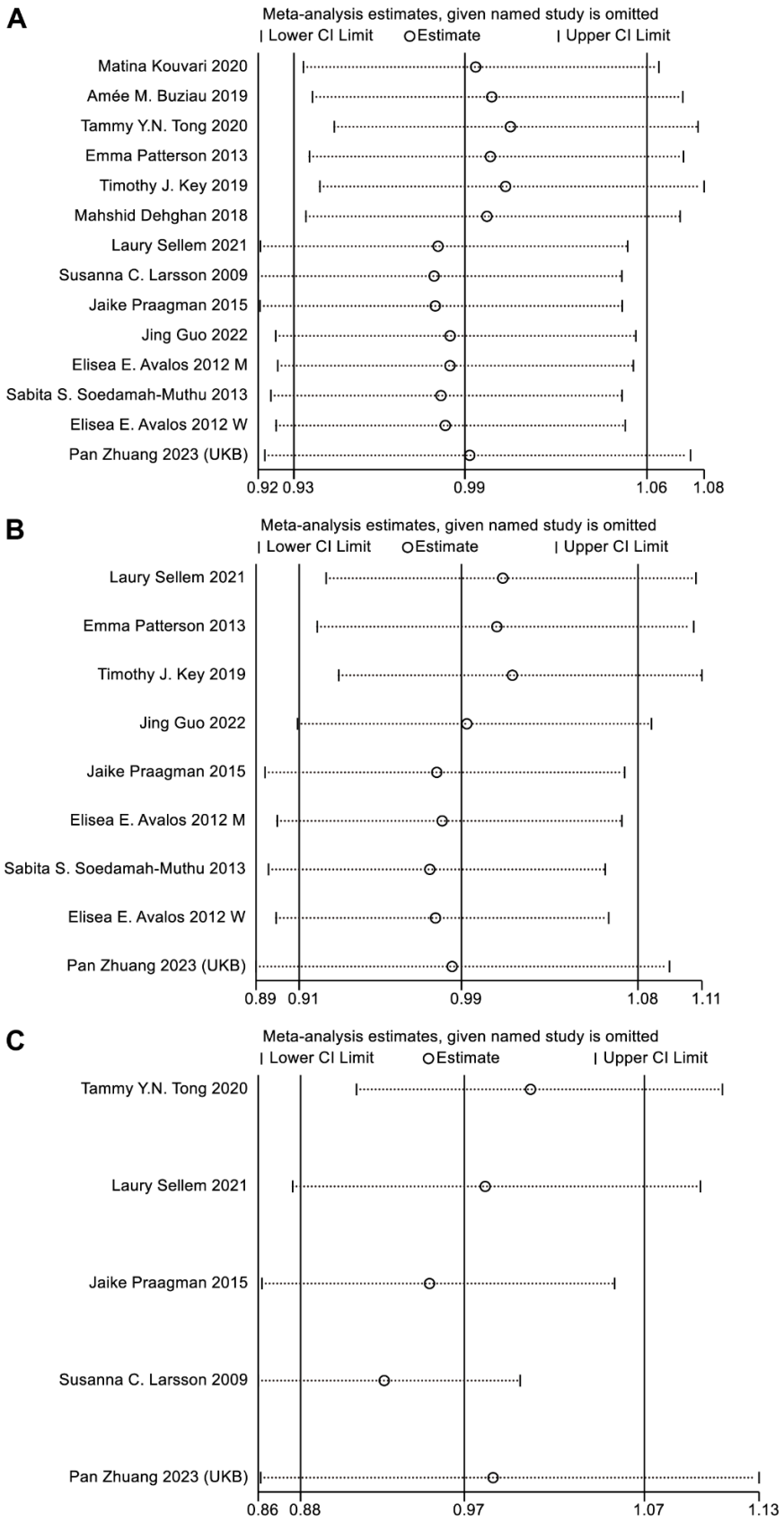
Supplementary Figure 20. Forest plot of influence analysis for the association between total dairy product consumption and risk of cardiovascular disease, coronary heart disease, and stroke.

(A) Cardiovascular disease. (B) coronary heart disease. (C) Stroke. Each dot and horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



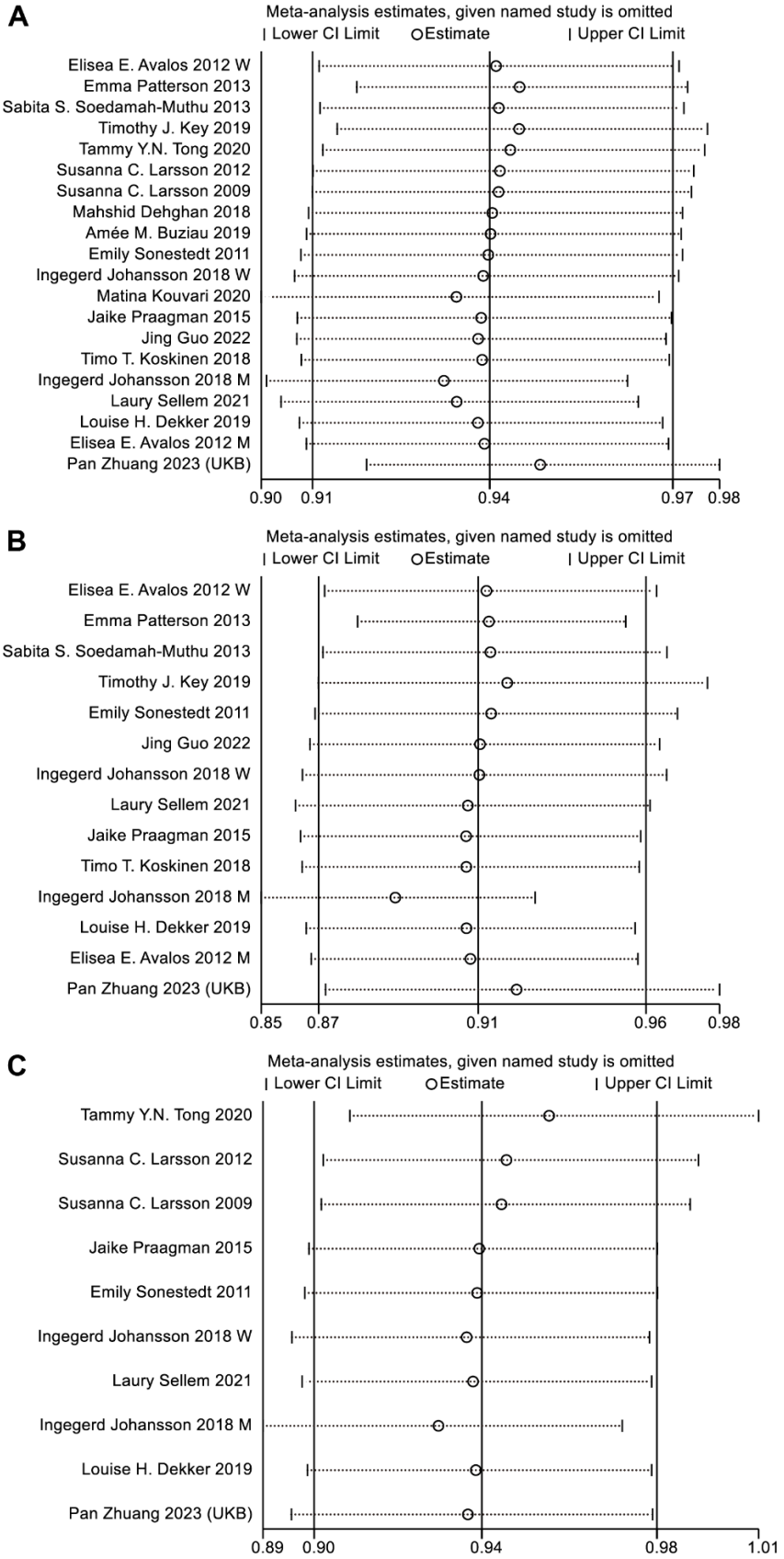
Supplementary Figure 21. Forest plot of influence analysis for the association between milk consumption and risk of cardiovascular disease, coronary heart disease, and stroke.

(A) Cardiovascular disease. (B) coronary heart disease. (C) Stroke. Each dot horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



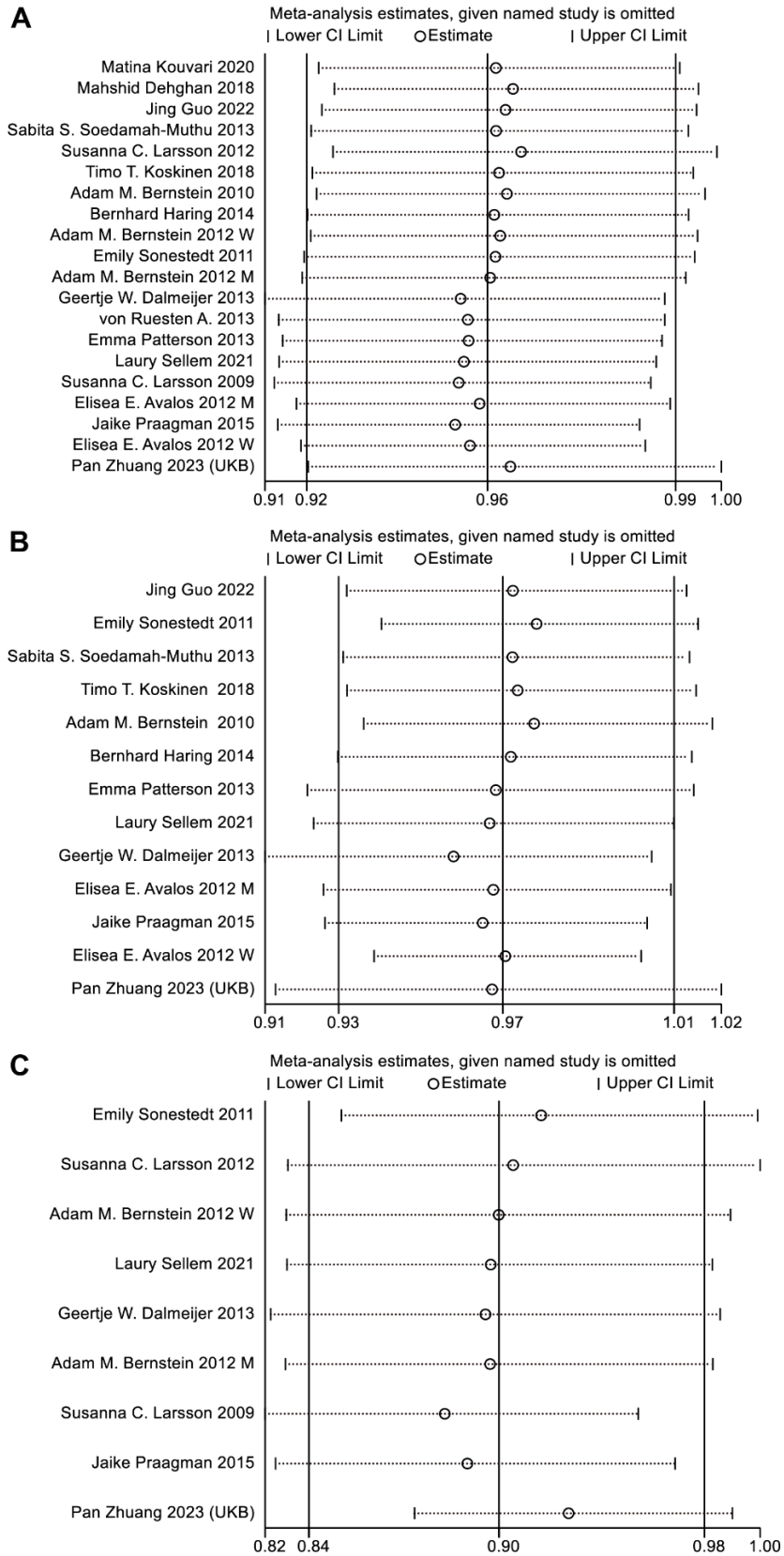
Supplementary Figure 22. Forest plot of influence analysis for the association between yogurt consumption and risk of cardiovascular disease, coronary heart disease, and stroke.

(A) Cardiovascular disease. (B) coronary heart disease. (C) Stroke. Each dot horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



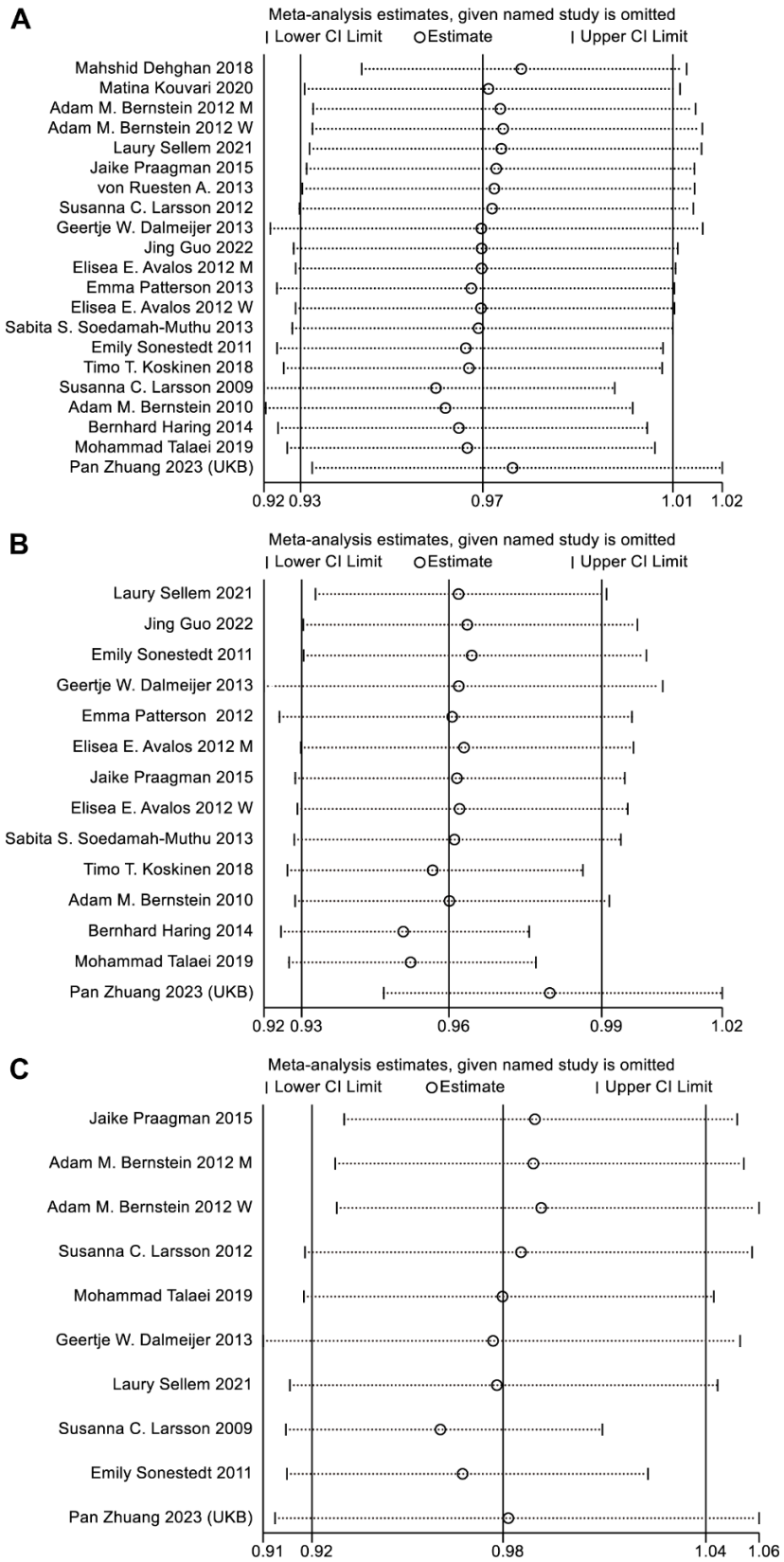
Supplementary Figure 23. Forest plot of influence analysis for the association between cheese consumption and risk of cardiovascular disease, coronary heart disease, and stroke.

(A) Cardiovascular disease. (B) coronary heart disease. (C) Stroke. Each dot horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



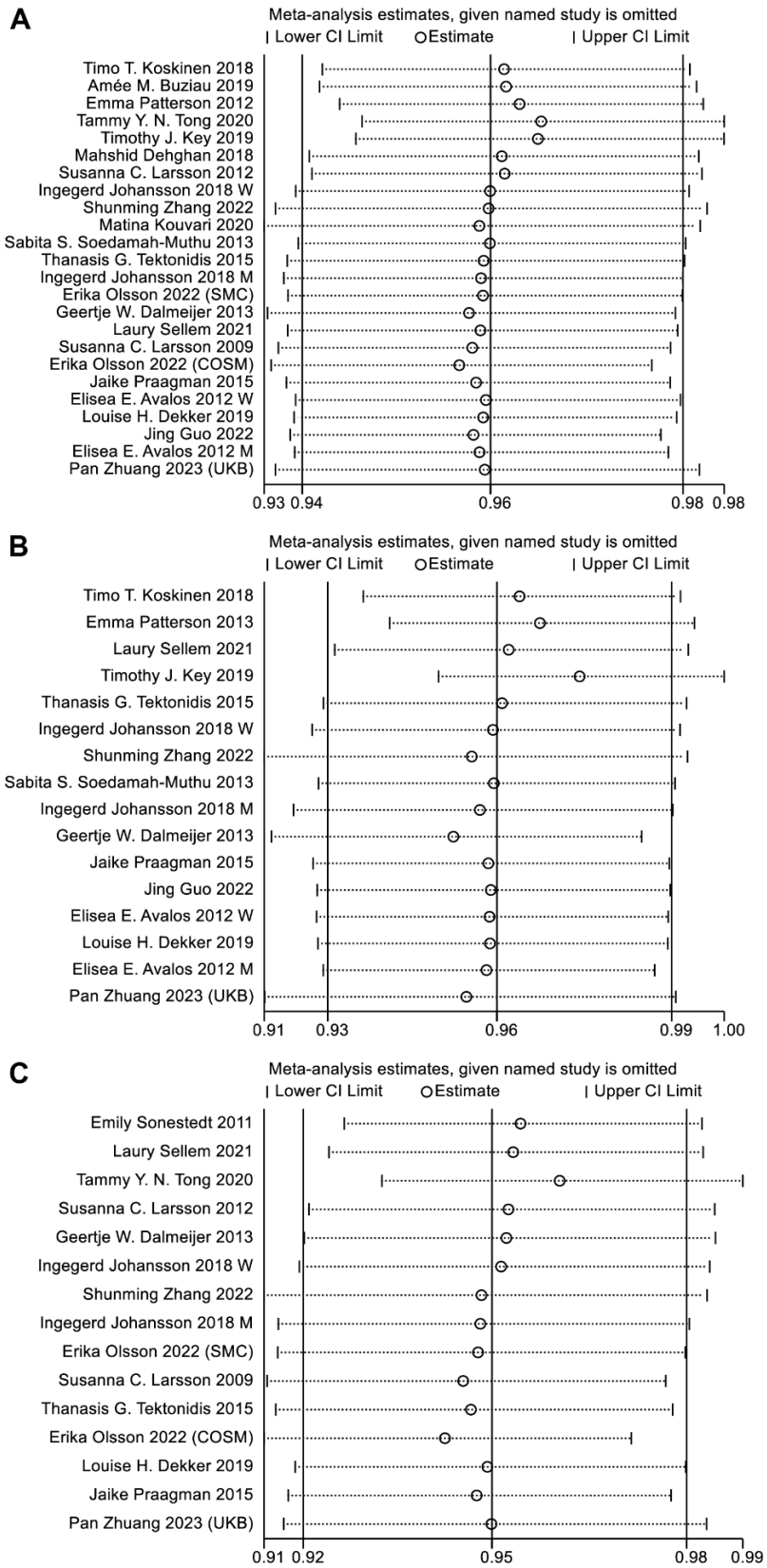
Supplementary Figure 24. Forest plot of influence analysis for the association between low-fat dairy products consumption and risk of cardiovascular disease, coronary heart disease, and stroke.

(A) Cardiovascular disease. (B) coronary heart disease. (C) Stroke. Each dot horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



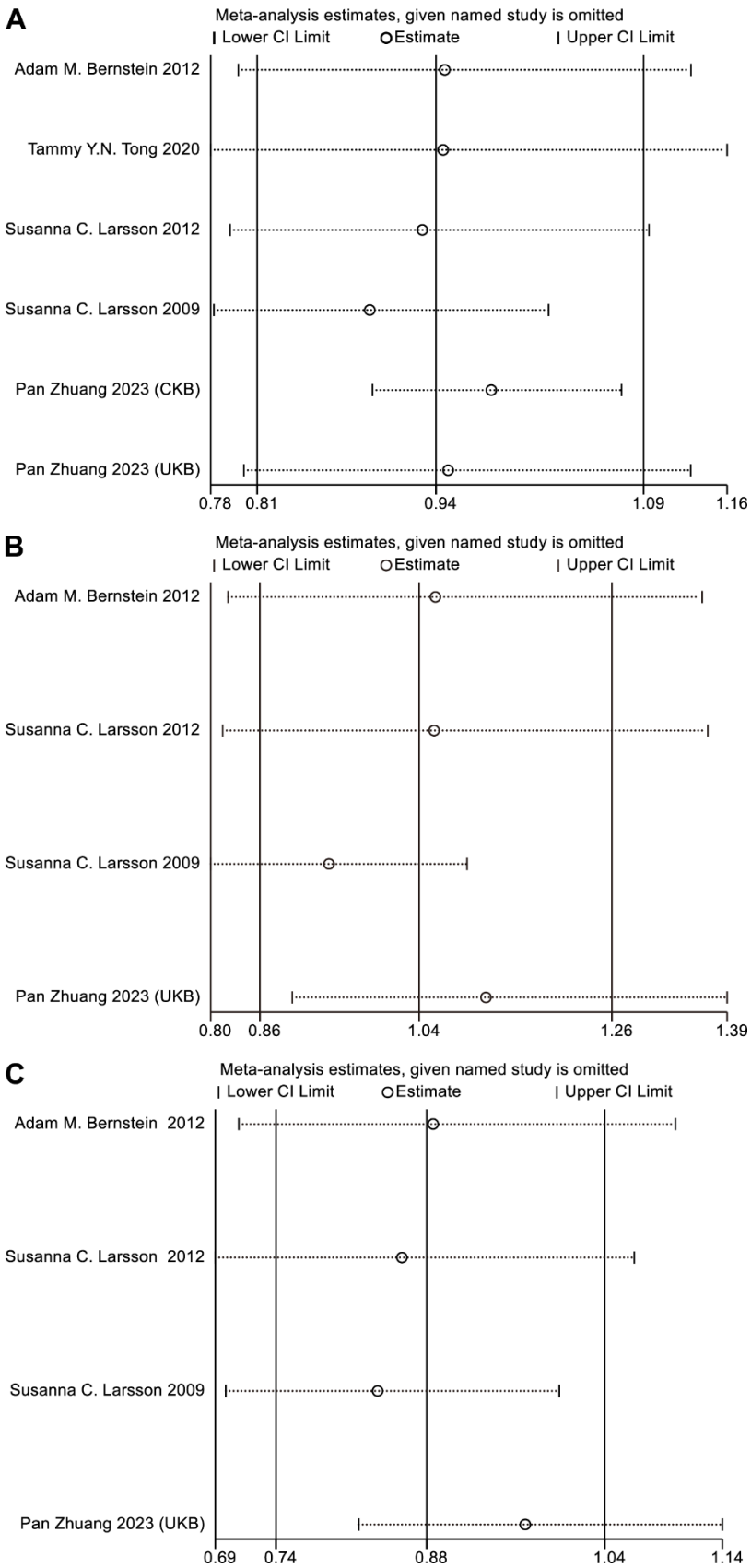
Supplementary Figure 25. Forest plot of influence analysis for the association between high-fat dairy products consumption and risk of cardiovascular disease, coronary heart disease, and stroke.

(A) Cardiovascular disease. (B) coronary heart disease. (C) Stroke. Each dot horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



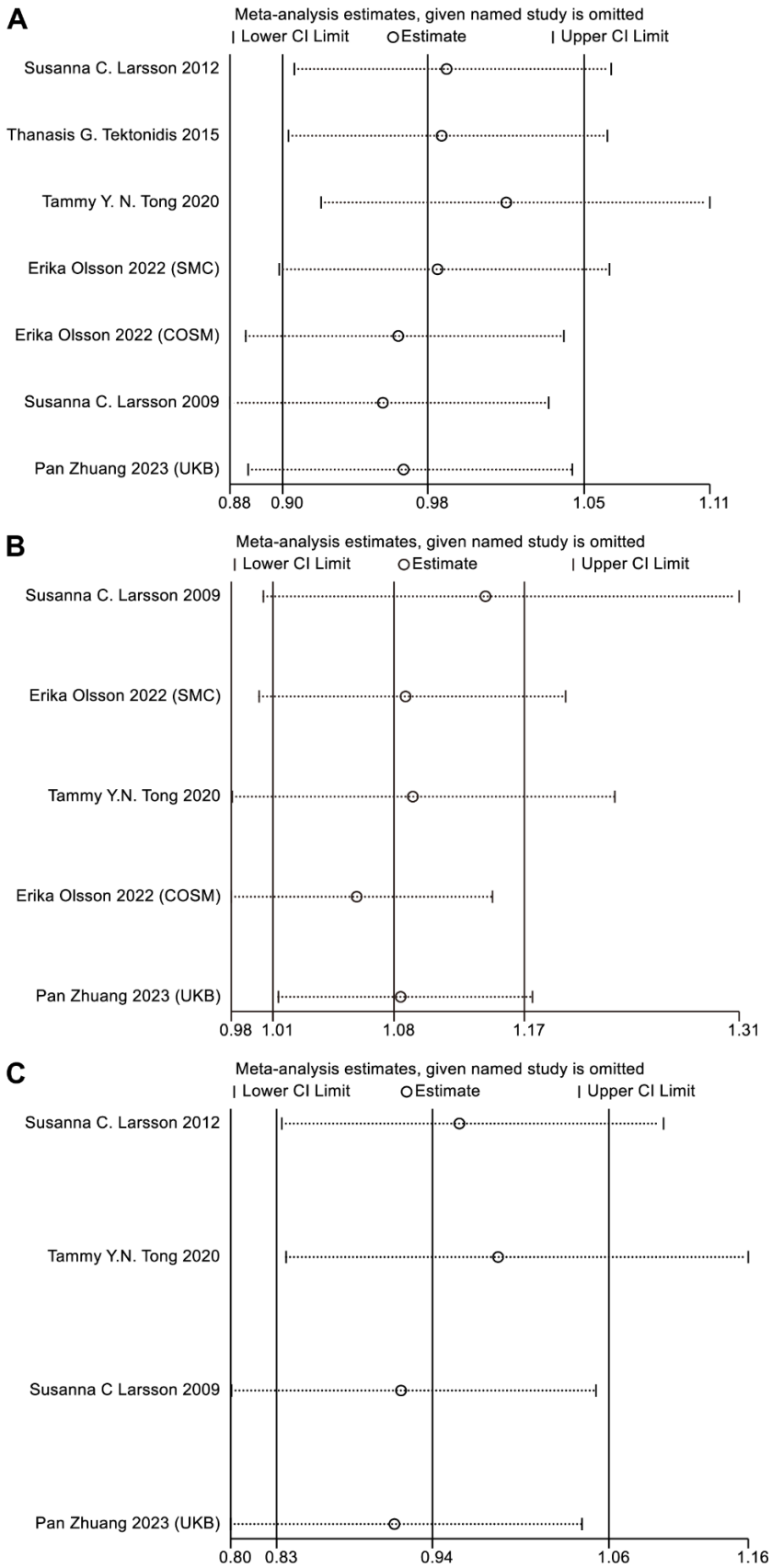
Supplementary Figure 26. Forest plot of influence analysis for the association between fermented dairy products consumption and risk of cardiovascular disease, coronary heart disease, and stroke.

(A) Cardiovascular disease. (B) coronary heart disease. (C) Stroke. Each dot horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



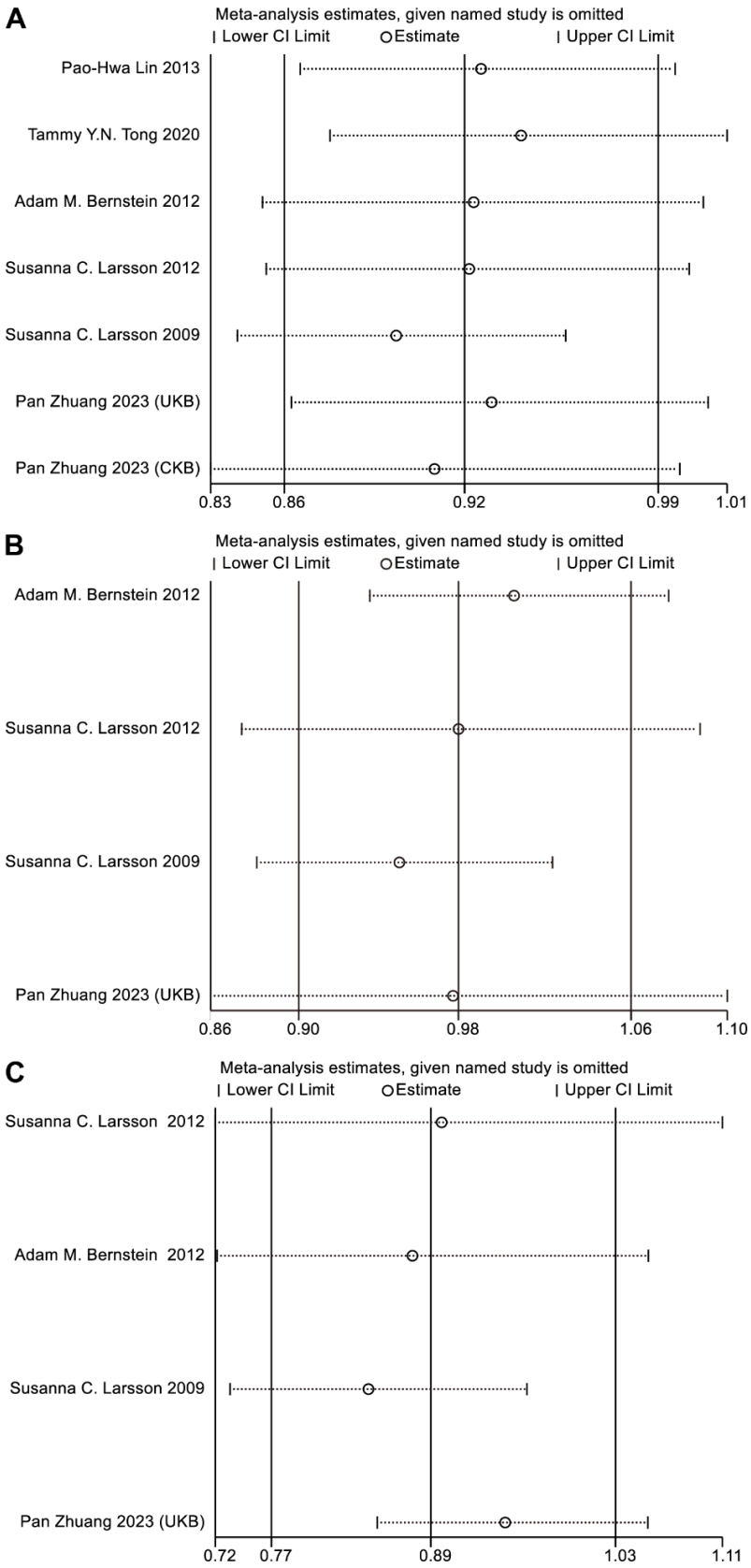
Supplementary Figure 27. Forest plot of influence analysis for the association between total, high-fat, and low-fat dairy consumption and risk of hemorrhagic stroke.

(A) Total dairy. (B) High-fat dairy. (C) Low-fat dairy. Each dot horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



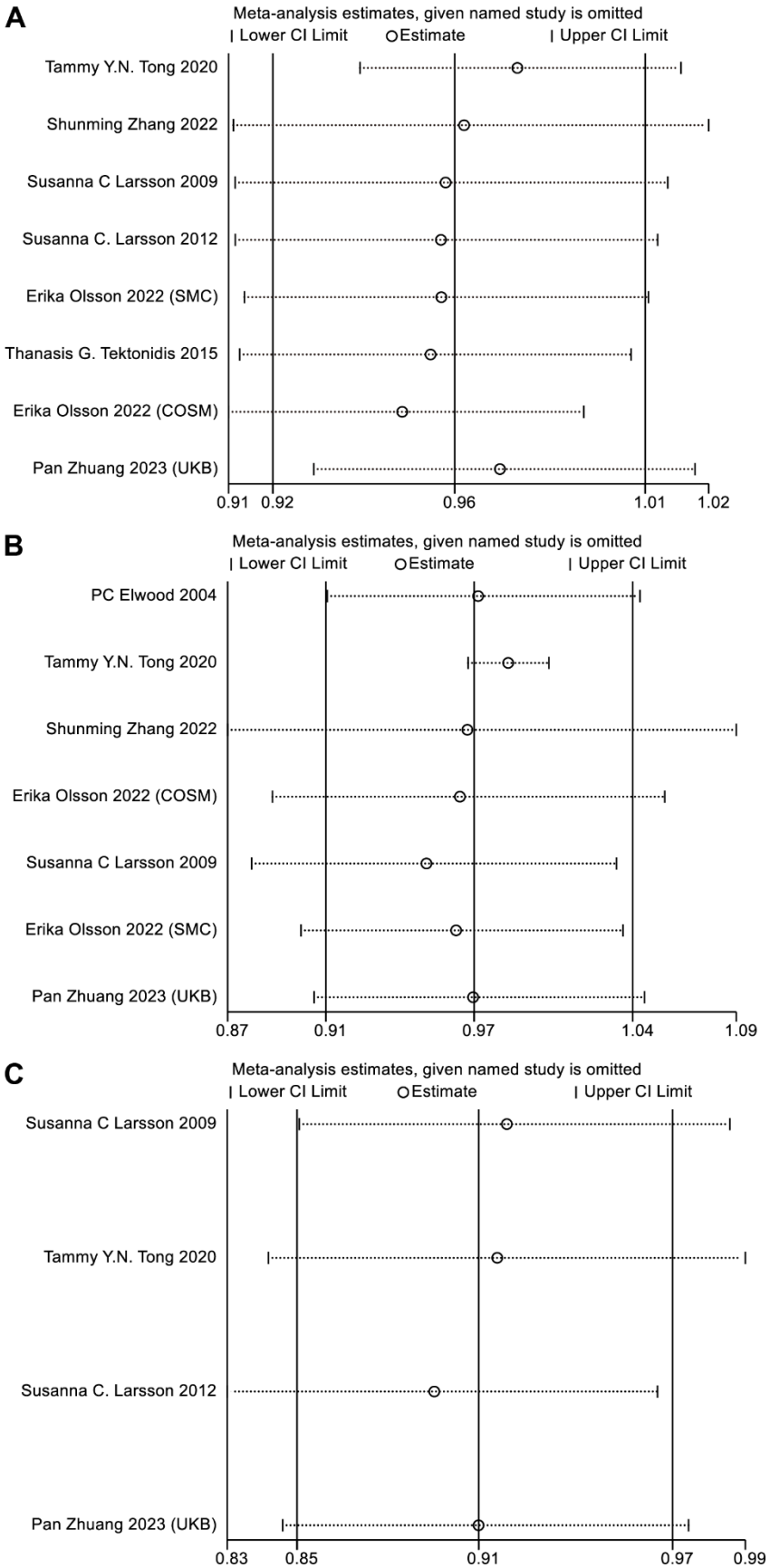
Supplementary Figure 28. Forest plot of influence analysis for the association between fermented dairy, milk, and cheese consumption and risk of hemorrhagic stroke.

(A) Fermented dairy. (B) Milk. (C) Cheese. Each dot horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



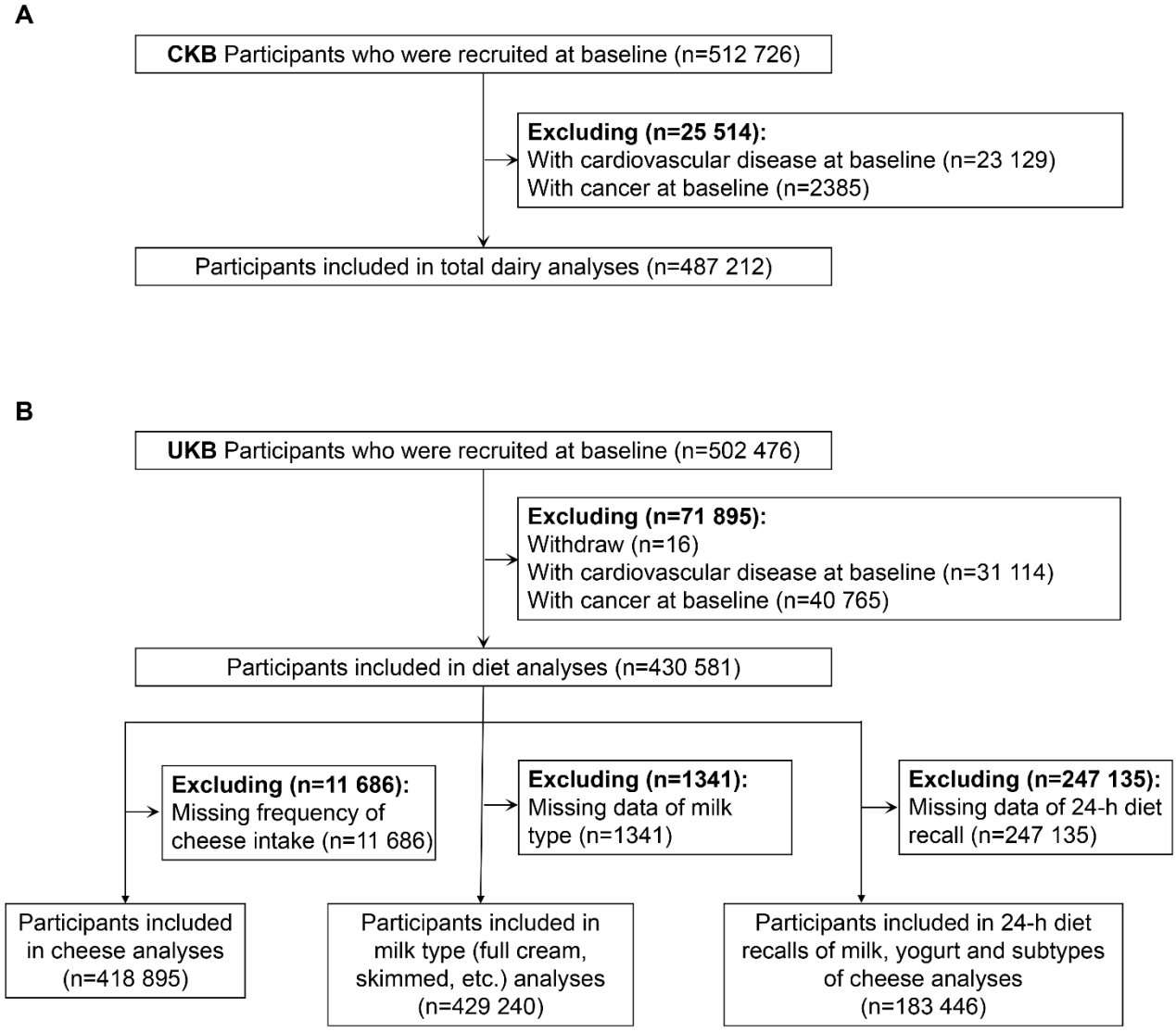
Supplementary Figure 29. Forest plot of influence analysis for the association between total dairy, high-fat dairy, and low-fat dairy consumption and risk of ischemic stroke.

(A) Total dairy. (B) High-fat dairy. (C) Low-fat dairy. Each dot horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



Supplementary Figure 30. Forest plot of influence analysis for the association between fermented dairy, milk, and cheese consumption and risk of ischemic stroke.

(A) Fermented dairy. (B) Milk. (C) Cheese. Each dot horizontal line represent the pooled RR (95% CI) following the exclusion of the study listed on the left using random-effects meta-analysis. Source data are provided as a Source Data file.



Supplementary Figure 31. Flow chart for participants.
(A) China Kadoorie Biobank. (B) UK Biobank.

References

1. Page, M.J., *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).
2. Wells GASB, O.C.D., Peterson J, Welch V, Losos M, Tugwell P. The Newcastle–Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses. (2009).
3. Bernstein, A.M., *et al.* Dietary protein sources and the risk of stroke in men and women. *Stroke* **43**, 637-644 (2012).
4. Bernstein, A.M., *et al.* Major dietary protein sources and risk of coronary heart disease in women. *Circulation* **122**, 876-883 (2010).
5. Dehghan, M., *et al.* Association of dairy intake with cardiovascular disease and mortality in 21 countries from five continents (PURE): a prospective cohort study. *Lancet* **392**, 2288-2297 (2018).
6. Jackson, D., White, I.R. & Thompson, S.G. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Stat. Med.* **29**, 1282-1297 (2010).
7. Drouin-Chartier, J.P., *et al.* Egg consumption and risk of cardiovascular disease: three large prospective US cohort studies, systematic review, and updated meta-analysis. *BMJ* **368**, m513 (2020).
8. Hedges, L.V. & Olkin, I. Statistical Methods for Meta-Analysis. (1987).
9. Higgins, J.P., Thompson, S.G., Deeks, J.J. & Altman, D.G. Measuring inconsistency in meta-analyses. *BMJ* **327**, 557-560 (2003).
10. Guyatt, G.H., *et al.* GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* **336**, 924-926 (2008).
11. Guyatt, G., *et al.* GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J. Clin. Epidemiol.* **64**, 383-394 (2011).
12. Guyatt, G.H., Oxman, A.D., Schünemann, H.J., Tugwell, P. & Knottnerus, A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J. Clin. Epidemiol.* **64**, 380-382 (2011).