

Supplementary Material Online for:

Plasma levels of apolipoprotein E, *APOE* genotype, and all-cause and cause-specific mortality in 105,949 individuals from a white general population cohort

Rasmussen KL, Tybjærg-Hansen A, Nordestgaard BG, Frikke-Schmidt R

Supplementary

Methods: Endpoints.

Supplementary

Methods: Statistical analysis.

Table S1: Egger plots for the weighted allele score groups and beta coefficients for a 1 mg/dL increase in plasma apoE in sensitivity analyses for instrumental variable analyses.

Figure S1: Distribution of plasma apolipoprotein E in women and men.

Figure S2: Overall survival as a function of age and *APOE* genotype in individuals in the general population.

Figure S3: Overall survival as a function of age and *APOE* genotype in women and men.

Figure S4: Risk of all-cause, cardiovascular, dementia-associated, and cancer mortality as a function of *APOE* genotype in women and men.

Figure S5: Multifactorially adjusted hazard ratios for all-cause mortality according to plasma levels of apolipoprotein E in women and men.

Figure S6: Multifactorially adjusted hazard ratios for cause-specific mortality according to plasma levels of apoE in women.

Figure S7: Multifactorially adjusted hazard ratios for cause-specific mortality according to plasma levels of apoE in men.

Figure S8: Plasma levels of apolipoprotein E and major lipids and lipoproteins as a function of *APOE* weighted and simple allele score group.

Figure S9: Multifactorially adjusted hazard ratios for all-cause and cause-specific mortality according to plasma levels of apolipoprotein E in analyses without adjustment for major lipids and lipoproteins.

Figure S10: Multifactorially adjusted hazard ratios for all-cause and cause-specific mortality for septiles of apolipoprotein E in individuals in the general population without adjustment for major lipids and lipoproteins.

- Figure S11:** Multifactorially adjusted hazard ratios for cause-specific mortality for septiles of apolipoprotein E in individuals in the general population for individuals with only one cause-specific event.
- Figure S12:** Multifactorially adjusted hazard ratios for all-cause and cause-specific mortality for tertiles, quartiles, quintiles and deciles of apolipoprotein E in individuals in the general population.
- Figure S13:** Multifactorially adjusted hazard ratios for all-cause and cause-specific mortality for septiles of apolipoprotein E in individuals before and after change of autoanalyzer in the Copenhagen General Population Study.
- Figure S14:** Risk of ischemic heart disease and cerebrovascular disease mortality as a function of *APOE* genotype.
- Figure S15:** Multifactorially adjusted hazard ratios for cause-specific mortality according to plasma levels of apoE in individuals in the general population adjusted for multiple hypothesis testing.
- Figure S16:** Risk of cause-specific mortality for a 1mg/dL change in observational and causal, genetically determined plasma apolipoprotein E level for the simple allele score.

Supplementary Methods: Endpoints.

The 100% complete Danish Civil Registration System records the date of death for all persons living in Denmark, while the national Danish Causes of Death Registry records ranked main causes of death as well as contributing causes of death, as reported by the attending physician in general practice, at a hospital, or in a forensic or pathology department. Causes of death were classified according to World Health Organization (WHO) International Classification of Diseases (ICD) 8th and 10th revision. Cardiovascular death was considered present if one of the top three ranked causes of death was cardiovascular (ICD-8 codes 390-458, ICD-10 codes I00-I99). Dementia-associated death was considered present if one of the top three ranked causes of death was all dementia (Alzheimer disease, vascular dementia and unspecified dementia (ICD8 290.10 and 290.18, and ICD10 F00, F01, F03, and G30)). Cancer associated death was considered if one of the top three ranked causes of death was a cancer diagnosis (ICD-8 codes 140-209, ICD-10 codes C00-C97). One individual could therefore contribute to more than one of the above-mentioned cause-specific deaths. Sensitivity analyses for plasma apoE and for *APOE* genotype excluding individuals with other cause-specific events (excluding all mixed events) gave similar results.

Supplementary Methods: Statistical analysis.

Kruskal-Wallis equality-of-populations rank test and Pearson's χ^2 -test were used to evaluate continuous and categorical variables. Missing data on categorical and continuous covariates (<0.3%) were imputed from age, sex, and population: multinomial logistic regression was applied for categorical variables and linear regression for continuous variables.

Multifactorially adjusted restricted cubic splines Cox regression using three knots were used to evaluate hazard ratios (HRs) between plasma apolipoprotein E and risk of all-cause and cause-specific mortality. Further, Cox regression was used to estimate HRs for septiles (seven equally sized groups) of plasma apoE as well as for *APOE* genotypes for otherwise identical models. Survival by *APOE* genotype was evaluated by Kaplan-Meier curves and log-rank trend tests.

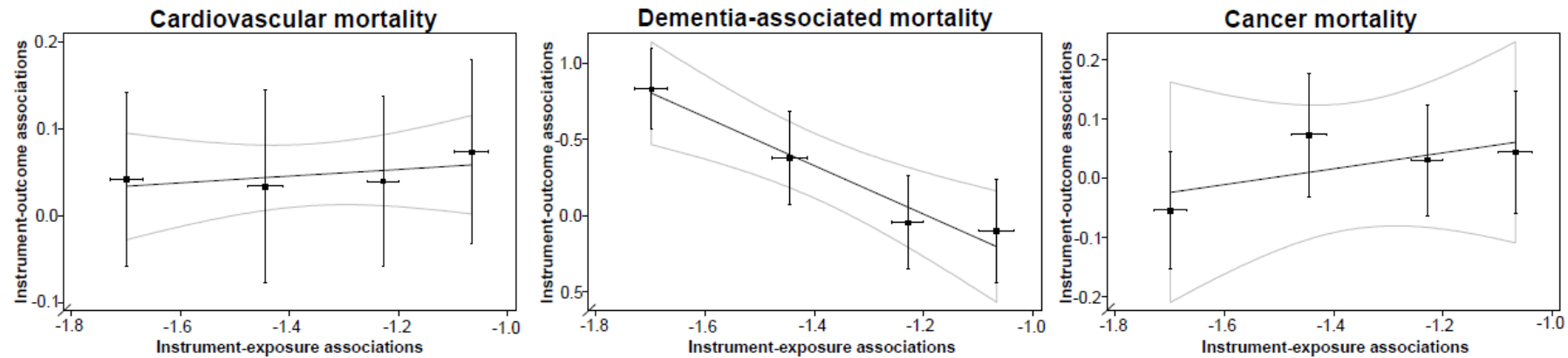
For Cox regression models, proportionality of hazards over time was assessed by plotting $-\ln(-\ln[\text{survival}])$ versus $\ln(\text{analysis time})$, and tested using Schoenfeld residuals. No major violations of the proportional hazards assumption were observed. Regression models were stratified in women and men and adjusted for known risk factors and markers of lifestyle: age (time scale), body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, hormonal replacement therapy, low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, and plasma triglycerides. For regression models for plasma apoE, further adjustment for *APOE* genotype was performed. Finally, sensitivity analyses only in *APOE* $\epsilon 33$ carriers were performed for plasma apoE to evaluate findings for individuals with wildtype isoform alone. Deviation from linearity on risk of all-cause and cause-specific mortality was evaluated by inclusion of apoE^2 as the interaction term in the Cox regression model, using a likelihood ratio test between models excluding and including the interaction term.

For instrumental variable analyses, we used a weighted allele score in five groups and an unweighted allele score in four groups based on five variants (rs449647, rs769446, rs405509, rs429358, and rs7412 in *APOE*), thus including 74,560 individuals in these analyses. Both promotor (rs449647, rs769446 and rs405509) and exonic (rs429358 and rs7412) *APOE* variants were common and therefore suited to serve as genetic instruments for plasma apoE levels. Combining all five genetic variants, we generated two different genetic instruments for plasma apoE. The first genetic instrument was calculated for each participant using a weighted sum of apoE changing alleles (coded as apoE lowering alleles), subsequently categorized into five groups of approximately equal size, named “weighted allele score group” and coded as 1 to 5 (5=lowest apoE). The weights correspond to the sum of the individual β -coefficients for apoE-lowering alleles in each individual obtained from a linear regression accounting for age, sex, cohort and the impact of all other four *APOE* variants. By doing so we ensured that both the strong $\epsilon 4$ and $\epsilon 2$ effect and the independent contributions of the three *APOE* promotor variants were captured. Because all five genetic variants were entered simultaneously in the linear regression linkage disequilibrium among variants was adjusted for. The

β -coefficient thus represent the independent effect on plasma apoE levels of each specific variant. The second genetic instrument was a simple counting of the number of apoE-changing alleles in each individual (coded as apoE lowering alleles), subsequently categorized into four groups of approximately equal size and named “simple allele score group” and coded 1 to 4 (4=lowest apoE). Laboratory methods for genotyping of rs449647, rs769446 and rs405509 and further details on statistical methods are described in reference 24 (Rasmussen et al, 2018, Alzheimers Dement) in the main manuscript. Sensitivity analyses for instrumental variable analyses were performed using *mrrobust* and *cv_regress* packages for Stata.

Table S1: Egger plots for the weighted allele score groups and beta coefficients for a 1 mg/dL increase in plasma apoE in sensitivity analyses for instrumental variable analyses.

A)



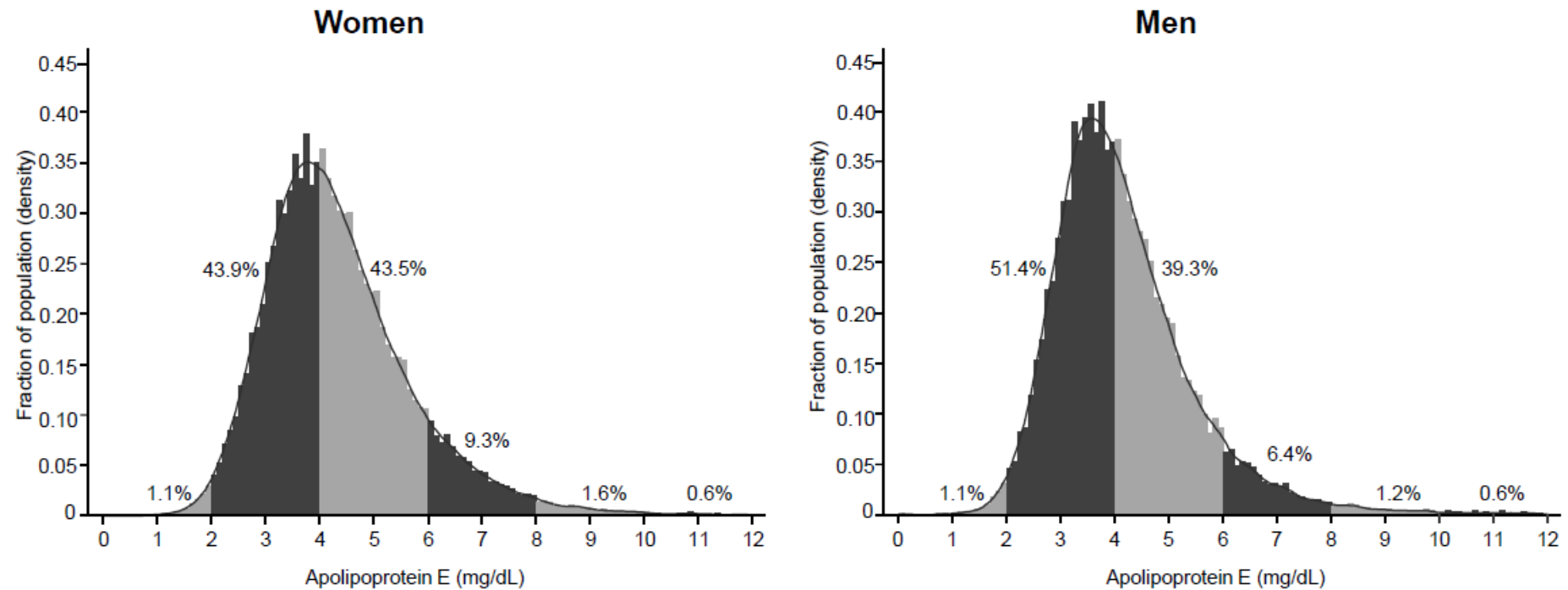
B)

Method	Beta coefficients for a 1 mg/dL increase in plasma apoE		
	Cardiovascular mortality	Dementia-associated mortality	Cancer mortality
Univariable Poisson instrumental variable analysis	-0.026 (p=0.35)	-0.53 (p=4x10 ⁻⁶)	0.0058 (p=0.83)
Multivariable Poisson instrumental variable analysis	0.15 (p=0.68)	-0.58 (p=0.02)	-0.037 (p=0.80)
MR-Egger, multiplicative standard errors	0.040 (p=0.72)	-1.60 (p=5x10 ⁻⁷)	0.13 (p=0.21)
MR-Egger, fixed effect standard errors	0.040 (p=0.72)	-1.60 (p= 5x10 ⁻⁷)	0.13 (p=0.21)
Inverse-variance weighted	-0.032 (p=0.09)	-0.27 (p=0.04)	-0.011 (p=0.58)
Weighted median	-0.026 (p=0.23)	-0.32 (p=7x10 ⁻⁵)	-0.025 (p=0.31)
Leave-one-out	0.065 (p=0.55)	-1.50 (p=3x10 ⁻⁶)	0.0039 (p=0.97)

A) Egger plots for groups of weighted allele score, excluding the reference group, for beta coefficients for instrument-outcome associations as a function of instrument-exposure associations with 95% confidence intervals for beta coefficients (black vertical and horizontal lines) and for MR-Egger (grey lines).

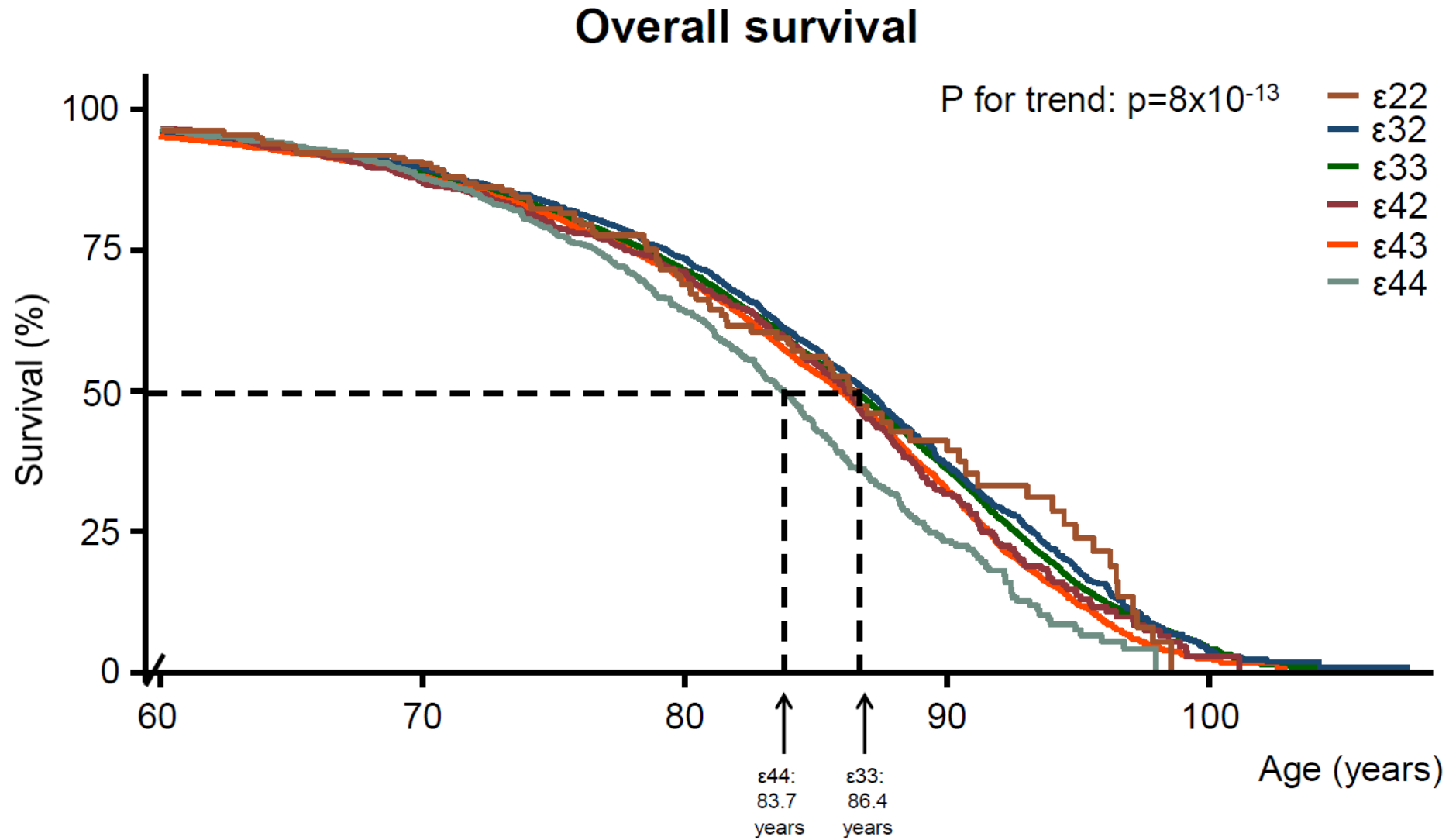
B) Analyses were performed for beta coefficients for the weighted allele score groups for instrument-outcome associations as a function of instrument-exposure associations. Leave-one-out estimations were performed for a MR-Egger regression using the leave-one-out predictions from a corresponding linear regression model and with standard errors for instrument-outcome associations from the original MR-Egger regression. P-values for beta coefficients are given in brackets.

Figure S1: Distribution of plasma apolipoprotein E in women and men.



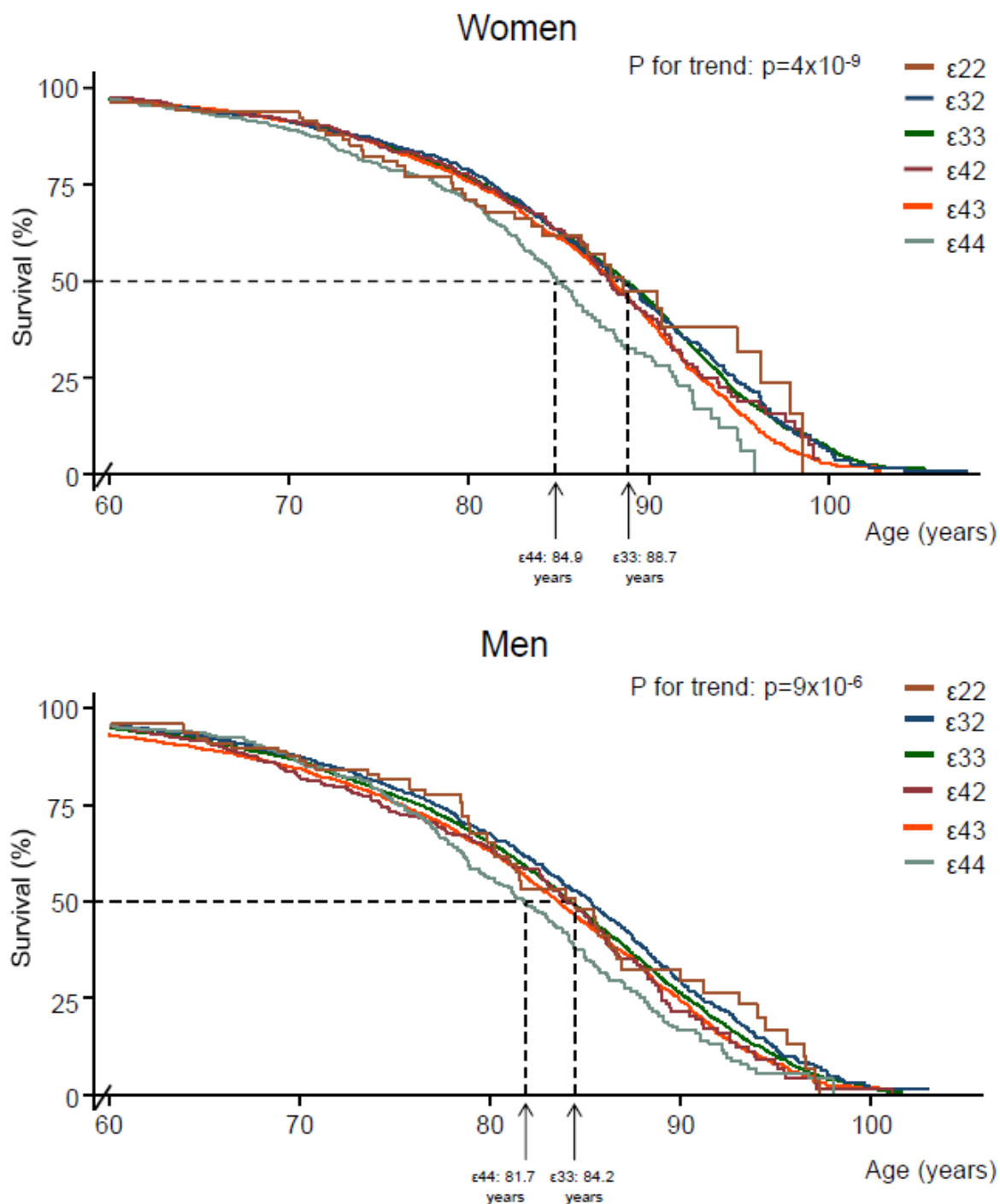
Kernel density estimates are superimposed both distributions. Graphs are truncated at 12 mg/dL, range is 0-27.0 mg/dL for women and 0-33.2 mg/dL for men. Percentages shown are percentages of men and women with plasma apolipoprotein E ≤ 2.0 , $2.0 < apoE \leq 4.0$, $4.0 < apoE \leq 6.0$, $6.0 < apoE \leq 8.0$, $8.0 < apoE \leq 10.0$, and $apoE > 10.0$ mg/dL. Mean values of apolipoprotein E are 4.4 mg/dL for women and 4.2 for men, and the mean for $\epsilon 33$ is 4.2 mg/dL, whereas the overall mean value for the population is 4.3 mg/dL. apoE=plasma apolipoprotein E.

Figure S2: Overall survival as a function of age and APOE genotype in individuals in the general population.



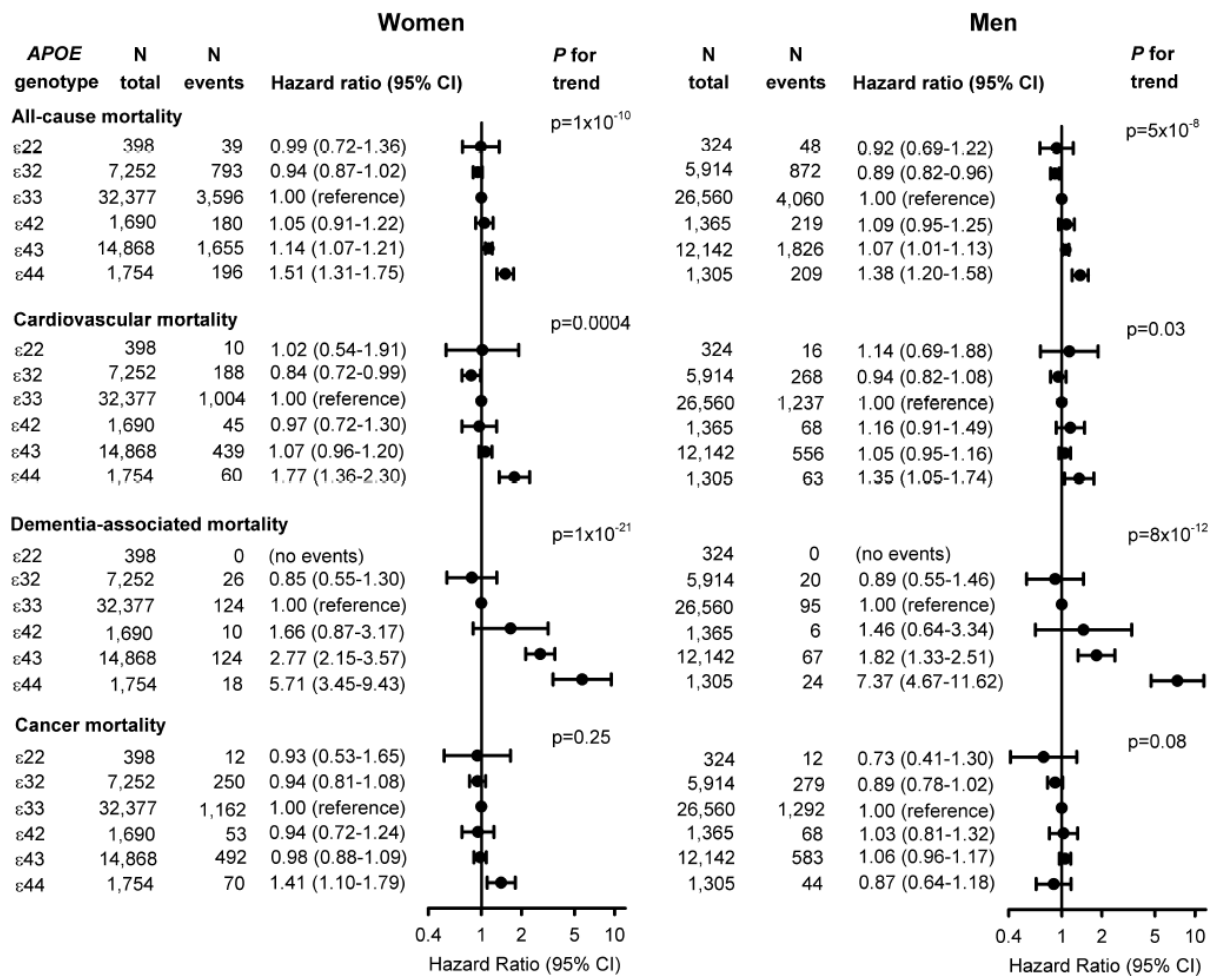
The Kaplan-Meier survival curves for the six common APOE genotypes in the Copenhagen General Population Study and the Copenhagen City Heart Study combined is shown. Dashed lines indicate median survival for ε44 and ε33. P for trend is by log-rank trend test.

Figure S3: Overall survival as a function of age and APOE genotype in women and men.



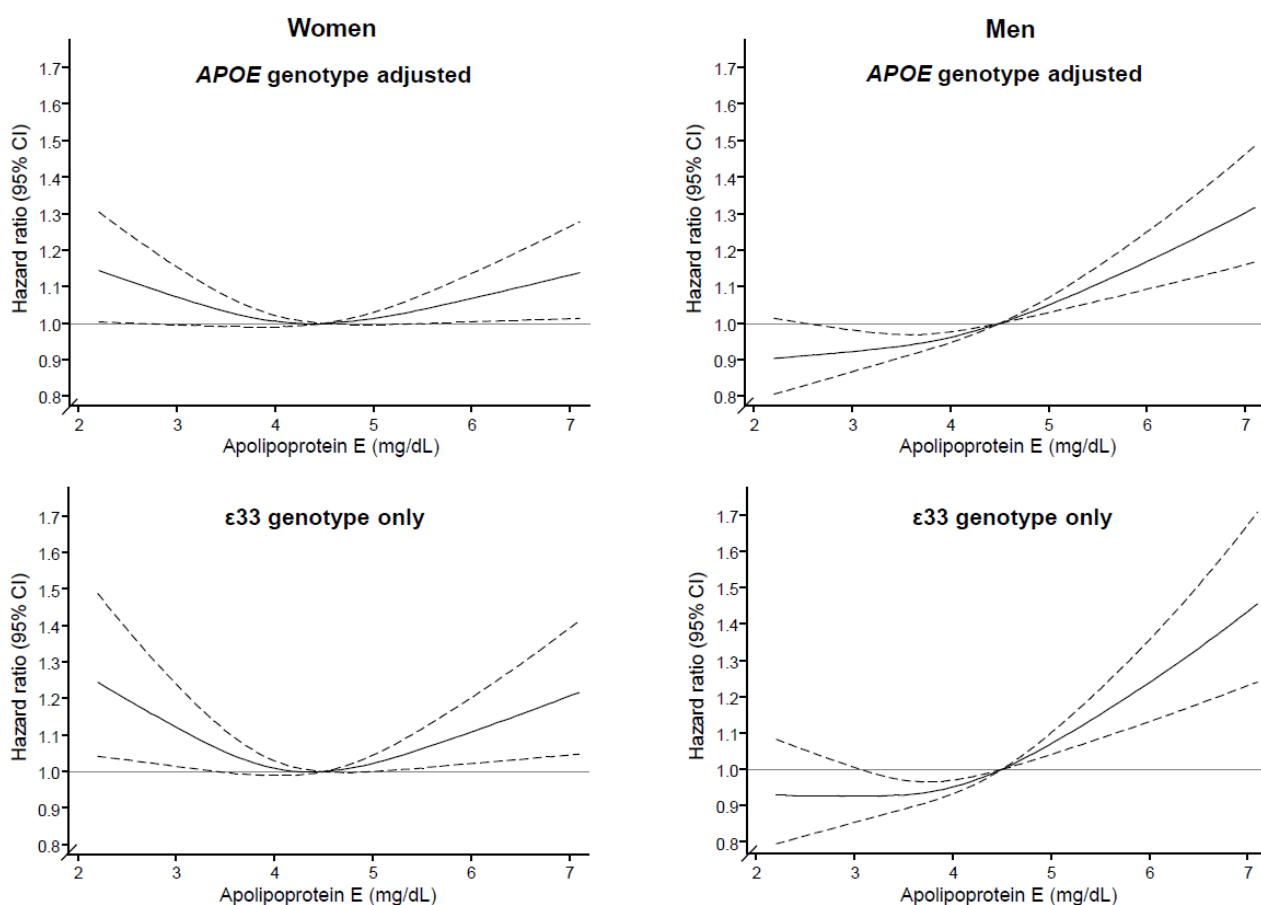
The Kaplan-Meier survival curves for the six common APOE genotypes stratified in women (above) and men (below) in the Copenhagen General Population Study and the Copenhagen City Heart Study combined is shown. Dashed lines indicate median survival for ϵ_{44} and ϵ_{33} . P for trend is by log-rank trend test. Median survival for did decrease from 88.7 to 88.5 to 88.0 to 87.8 to 87.7 to 84.9 years for ϵ_{33} to ϵ_{32} to ϵ_{43} to ϵ_{22} to ϵ_{42} to ϵ_{44} in women, and from 85.3 to 84.2 to 84.0 to 83.9 to 83.5 to 81.7 years for ϵ_{32} to ϵ_{33} to ϵ_{22} to ϵ_{42} to ϵ_{43} to ϵ_{44} for men.

Figure S4: Risk of all-cause, cardiovascular, dementia-associated, and cancer mortality as a function of APOE genotype in women and men.



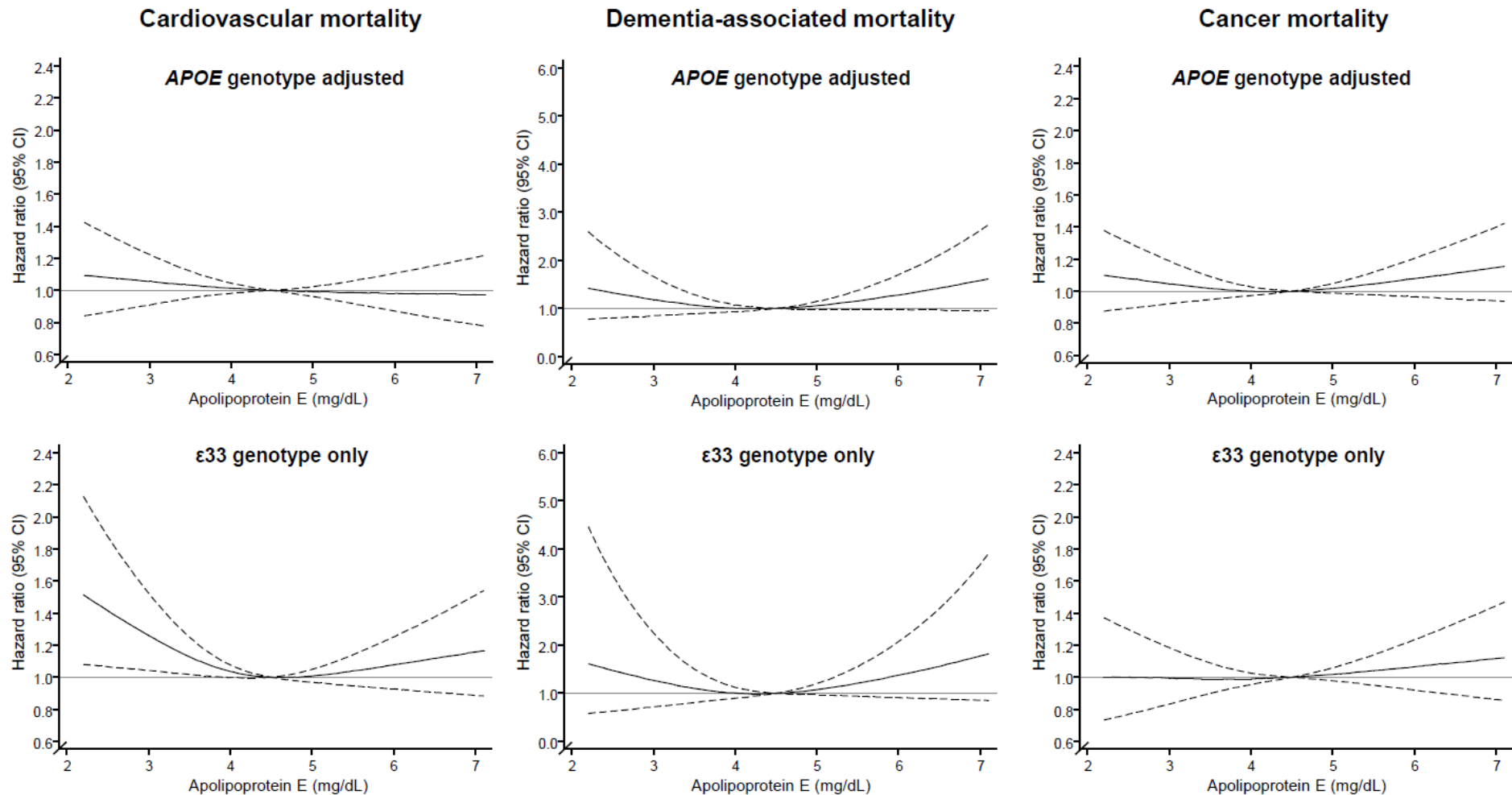
Cox regression models were stratified in women (left) and men (right) and adjusted for age (time scale), body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, hormonal replacement therapy, LDL cholesterol, HDL cholesterol and triglycerides. APOE = APOE ε2/ε3/ε4 genotype. CI=95% confidence interval.

Figure S5: Multifactorially adjusted hazard ratios for all-cause mortality according to plasma levels of apolipoprotein E in women and men.



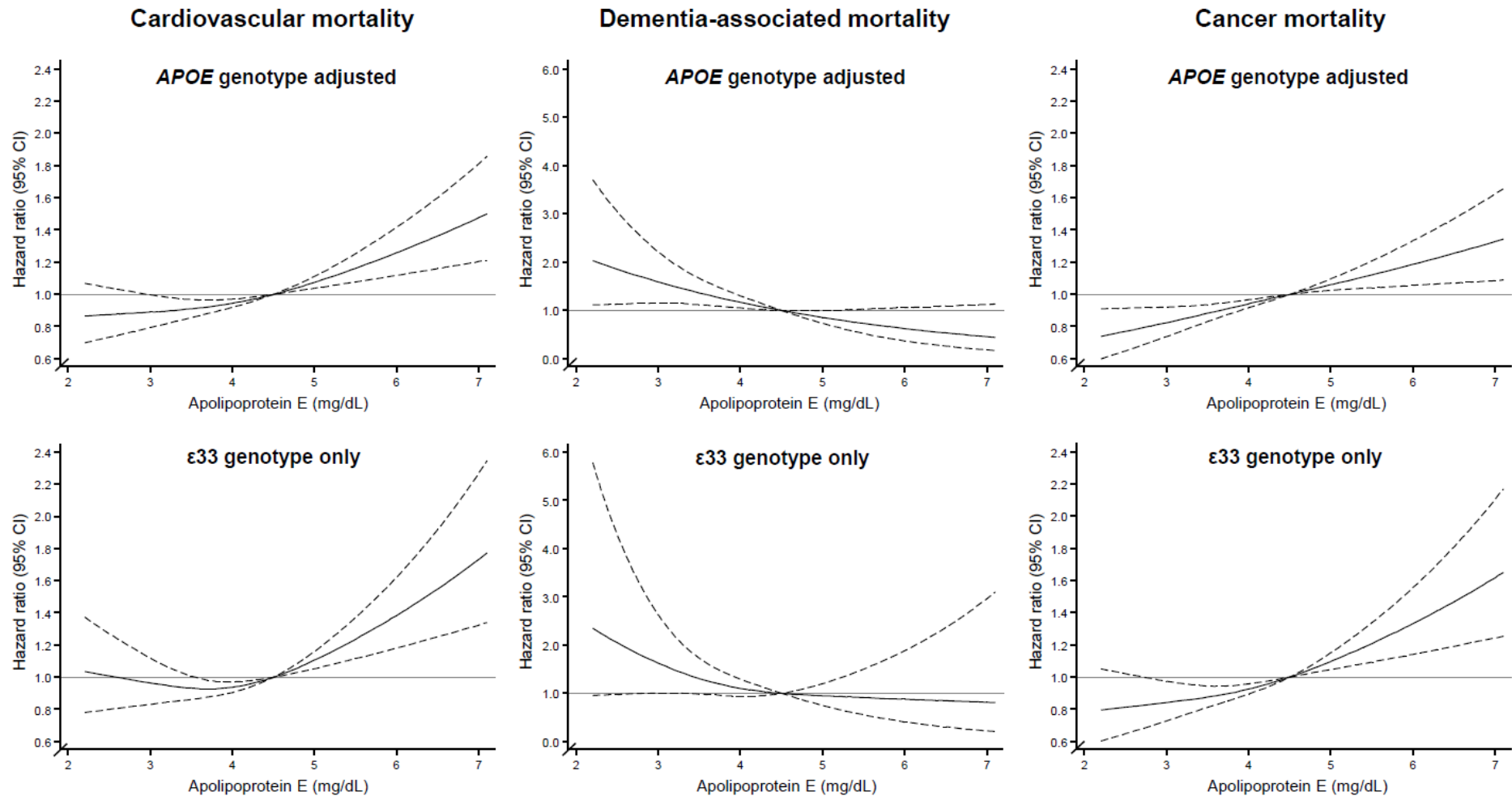
Solid lines are multifactorially adjusted hazard ratios, whereas dashed lines indicate 95% confidence intervals derived from restricted cubic splines with three knots and with the reference defined as the plasma level of apolipoprotein E with lowest overall mortality (4.5 mg/dL in both men and women). Graphs are truncated at the level of 2.1 mg/dL and 7.1 mg/dL, due to statistically unstable estimates at extreme low and high levels thus only including 53,928 women and 44,422 men from the Copenhagen General Population Study and the Copenhagen City Heart Study in these analyses (30,894 women and 25,370 men in the $\epsilon 33$ genotype stratified analyses). Cox regression models were stratified in women (left) and men (right) and adjusted for age (time scale), body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, hormonal replacement therapy, LDL cholesterol, HDL cholesterol, triglycerides (all panels) and further adjusted for *APOE* $\epsilon 2/\epsilon 3/\epsilon 4$ genotype (upper panels) and in analyses restricted to *APOE* $\epsilon 33$ carriers (bottom panels). 95% CI=95% confidence interval. *APOE*=*APOE* $\epsilon 2/\epsilon 3/\epsilon 4$ genotype. $\epsilon 33$ =*APOE* wildtype.

Figure S6: Multifactorially adjusted hazard ratios for cause-specific mortality according to plasma levels of apoE in women.



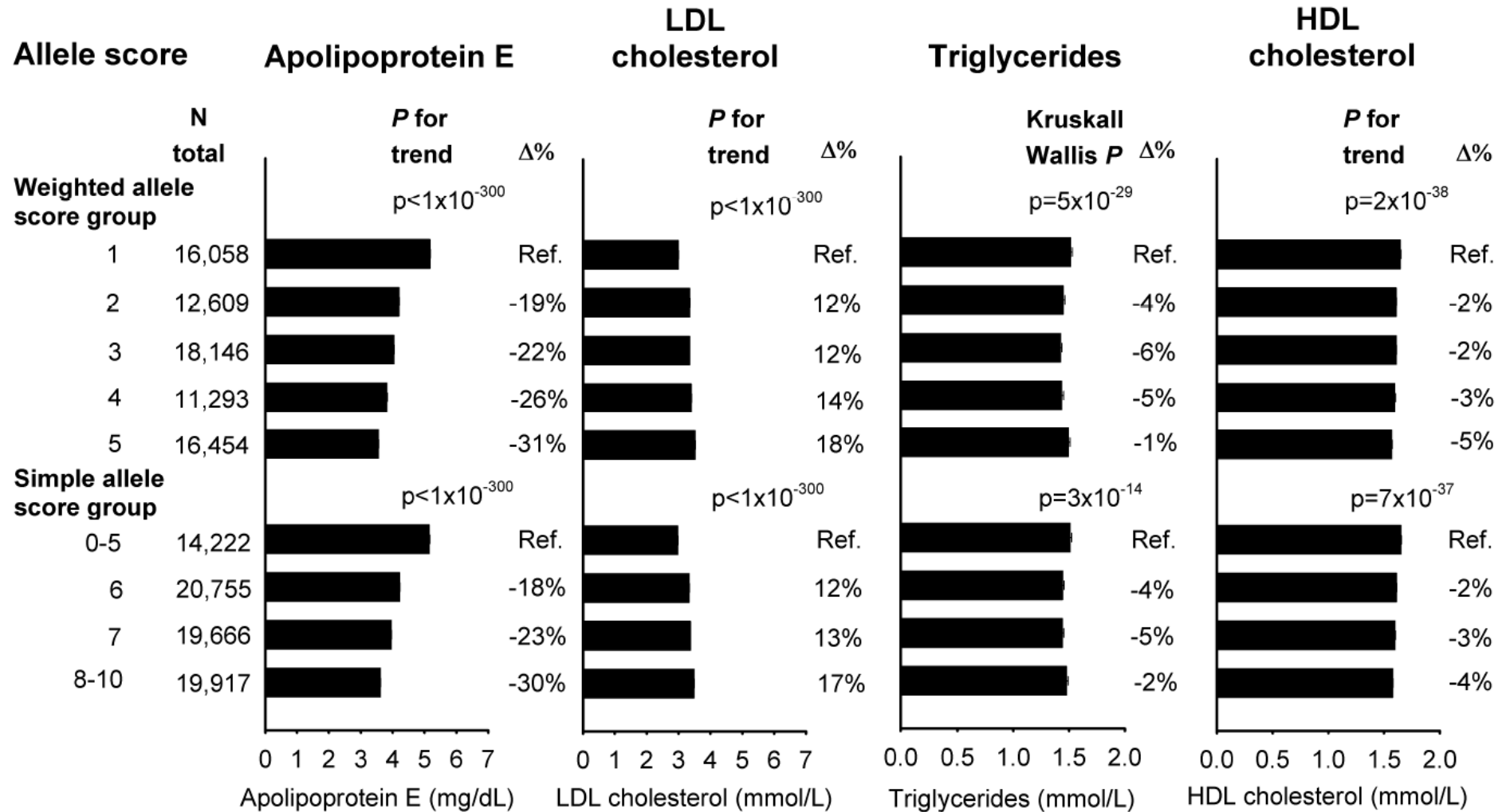
Solid lines are hazard ratios, whereas dashed lines indicate 95% confidence intervals derived from restricted cubic spline regressions models similar to Figure 3 with respect to reference (4.5 mg/dL), truncation (2.1-7.1 mg/dL) and Cox regression models were adjusted for age (time scale), body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, hormonal replacement therapy, LDL cholesterol, HDL cholesterol, triglycerides (all panels) and further adjusted for *APOE* genotype (upper panels, N=53,928) and in analyses restricted to ε33 carriers (bottom panels, N=30,894). 95% CI=95% confidence interval. *APOE*=*APOE* ε2/ε3/ε4 genotype. ε33=*APOE* wildtype.

Figure S7: Multifactorially adjusted hazard ratios for cause-specific mortality according to plasma levels of apoE in men.



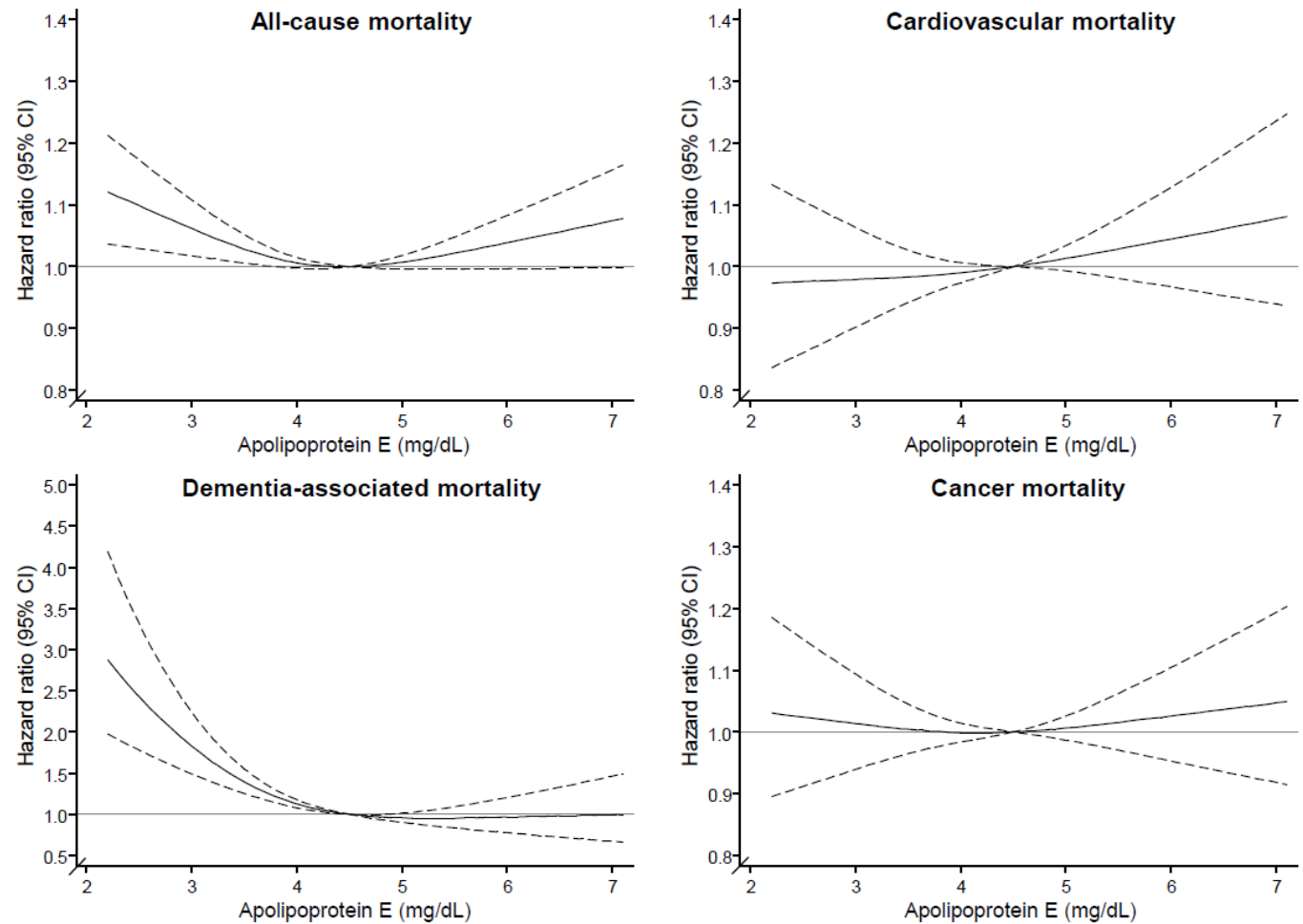
Solid lines are hazard ratios, whereas dashed lines indicate 95% confidence intervals derived from restricted cubic spline regressions models similar to Figure 3 with respect to reference (4.5 mg/dL), truncation (2.1-7.1 mg/dL) and Cox regression models were adjusted for age (time scale), body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, hormonal replacement therapy, LDL cholesterol, HDL cholesterol, triglycerides (all panels) and further adjusted for *APOE* genotype (upper panels, N=44,422) and in analyses restricted to ε33 carriers (bottom panels, N=25,370). 95% CI=95% confidence interval. *APOE*=*APOE* ε2/ε3/ε4 genotype. ε33=*APOE* wildtype.

Figure S8: Plasma levels of apolipoprotein E and major lipids and lipoproteins as a function of APOE weighted and simple allele score group.



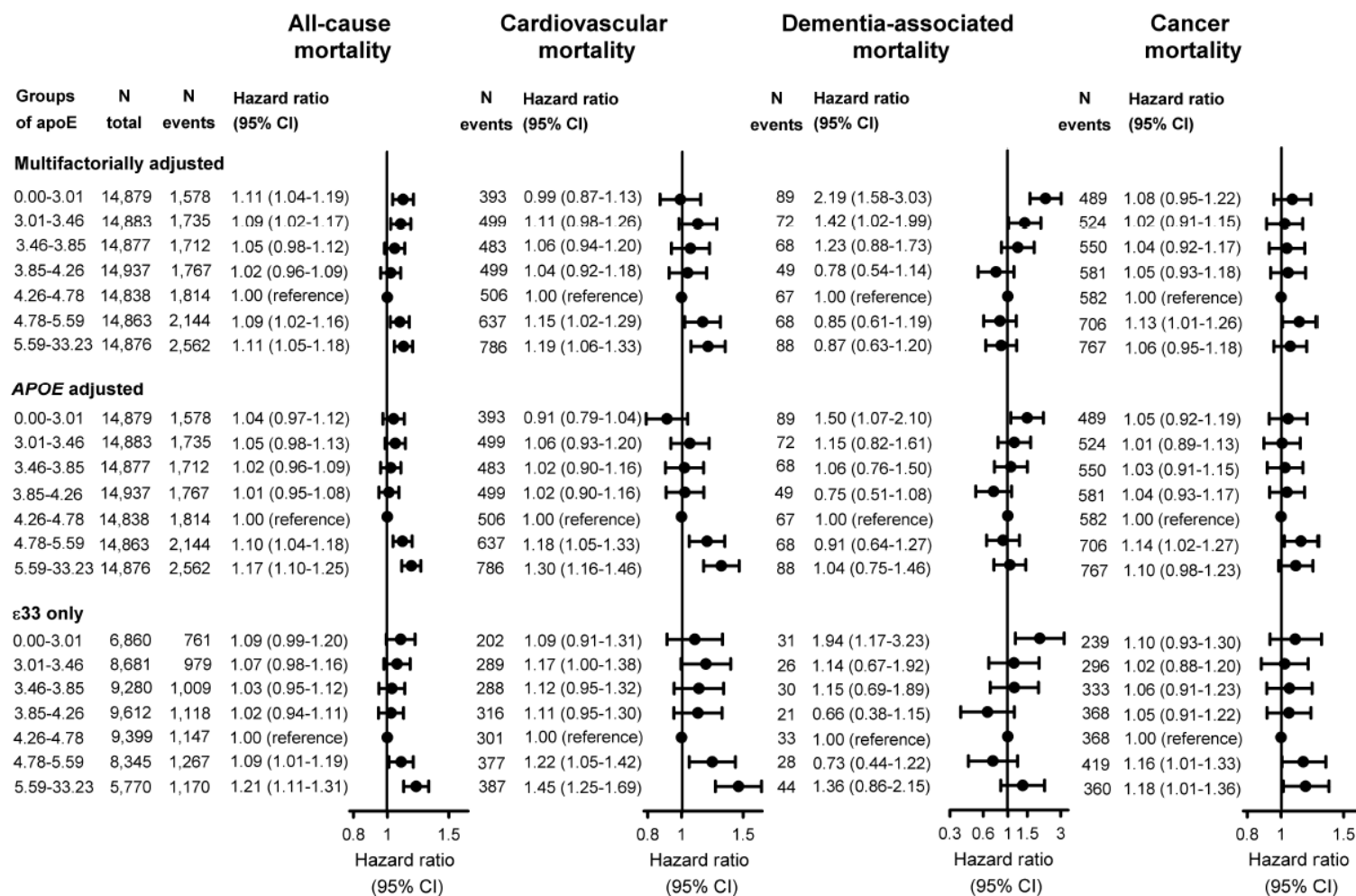
Values for plasma levels of apolipoprotein E and triglycerides are geometric mean \pm standard error of the mean, values for LDL cholesterol and HDL cholesterol are arithmetic mean \pm standard error of the mean. HDL=high-density lipoprotein. LDL=low-density lipoprotein. Ref.=reference.

Figure S9: Multifactorially adjusted hazard ratios for all-cause and cause-specific mortality according to plasma levels of apolipoprotein E in analyses without adjustment for major lipids and lipoproteins.



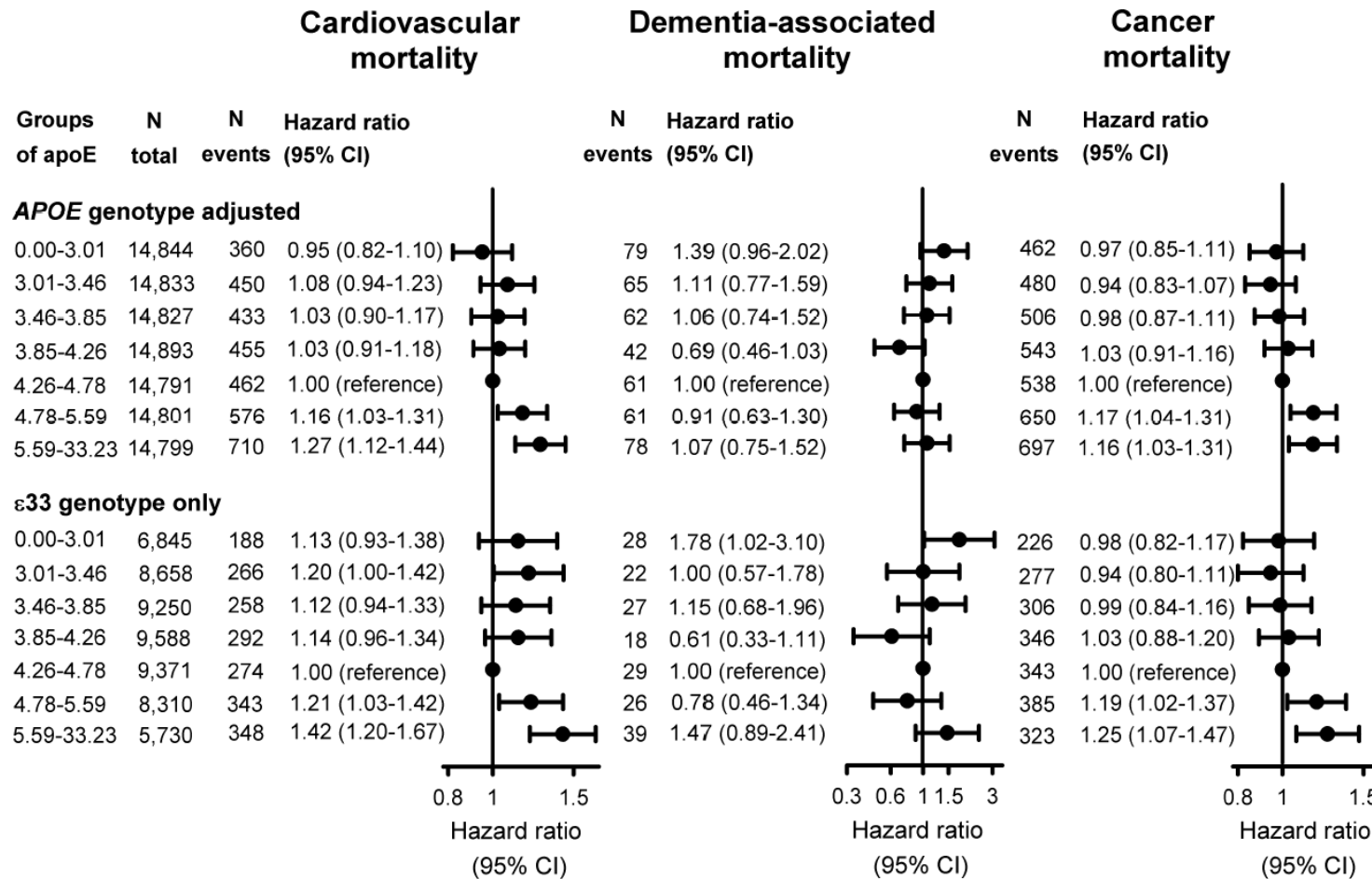
Solid lines are hazard ratios, whereas dashed lines indicate 95% confidence intervals derived from restricted cubic spline regression models similar to Figure 3 with respect to reference (4.5 mg/dL) and truncation (2.1-7.1 mg/dL). Cox regression models were adjusted for age (time scale), sex, body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, and hormonal replacement therapy. 95% CI=95% confidence interval.

Figure S10: Multifactorially adjusted hazard ratios for all-cause and cause-specific mortality for septiles of apolipoprotein E in individuals in the general population without adjustment for major lipids and lipoproteins.



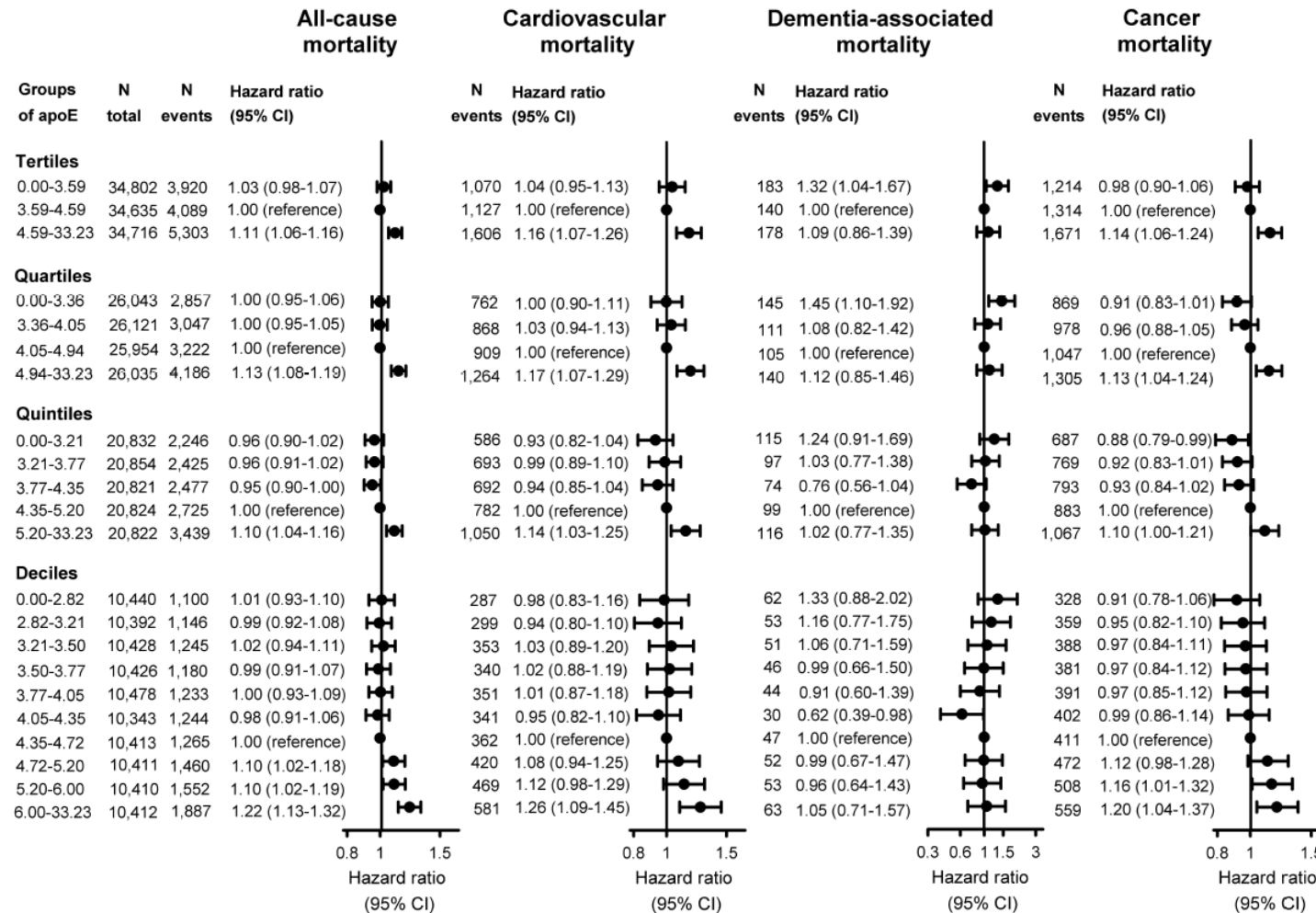
Cox regression models were multifactorially adjusted for age (time scale), sex, body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, and hormonal replacement therapy (upper panels), in analyses restricted to ε33 carriers (bottom panels), and further adjusted for APOE genotype (middle panels). 95% CI=95% confidence interval. apoE=plasma apolipoprotein E level. APOE=APOE ε2/ε3/ε4 genotype. ε33=APOE wildtype carriers.

Figure S11: Multifactorially adjusted hazard ratios for cause-specific mortality for septiles of apolipoprotein E in individuals in the general population for individuals with only one cause-specific event.



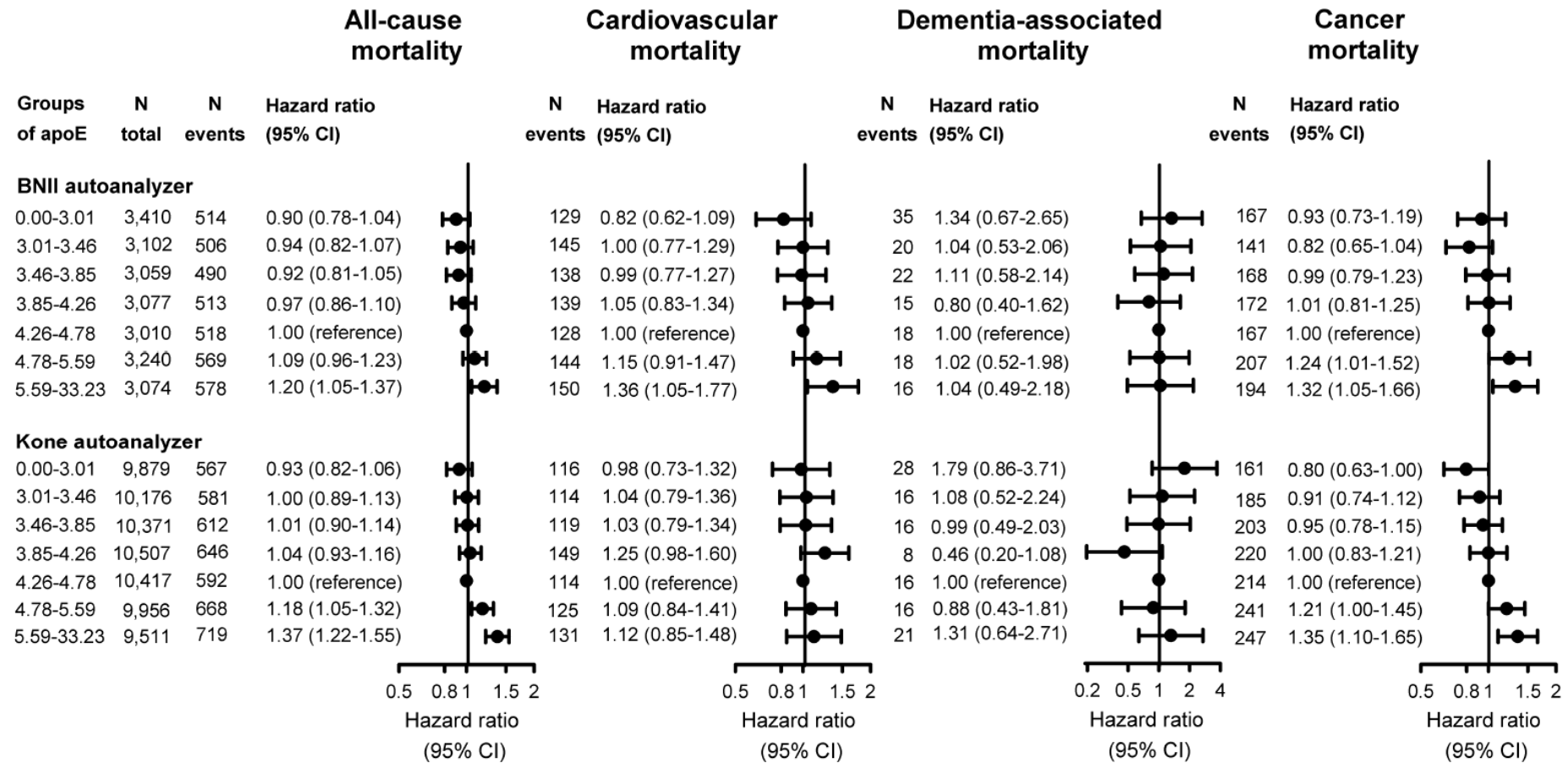
365 individuals contributing with events in two or three of the cause-specific categories (cardiovascular mortality, dementia-associated mortality and/or cancer mortality) were excluded, but otherwise analyses were similar to the analyses shown in Figure 5. Cox regression models were multifactorially adjusted for age (time scale), sex, body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, hormonal replacement therapy, LDL cholesterol, HDL cholesterol, triglycerides (all panels), with further adjustment for *APOE* genotype (upper panels) and in analyses restricted to ε33 carriers (bottom panels). 95% CI=95% confidence interval. apoE=plasma apolipoprotein E level. *APOE*=*APOE* ε2/ε3/ε4 genotype. ε33=*APOE* wildtype carriers.

Figure S12: Multifactorially adjusted hazard ratios for all-cause and cause-specific mortality for tertiles, quartiles, quintiles and deciles of apolipoprotein E in individuals in the general population.



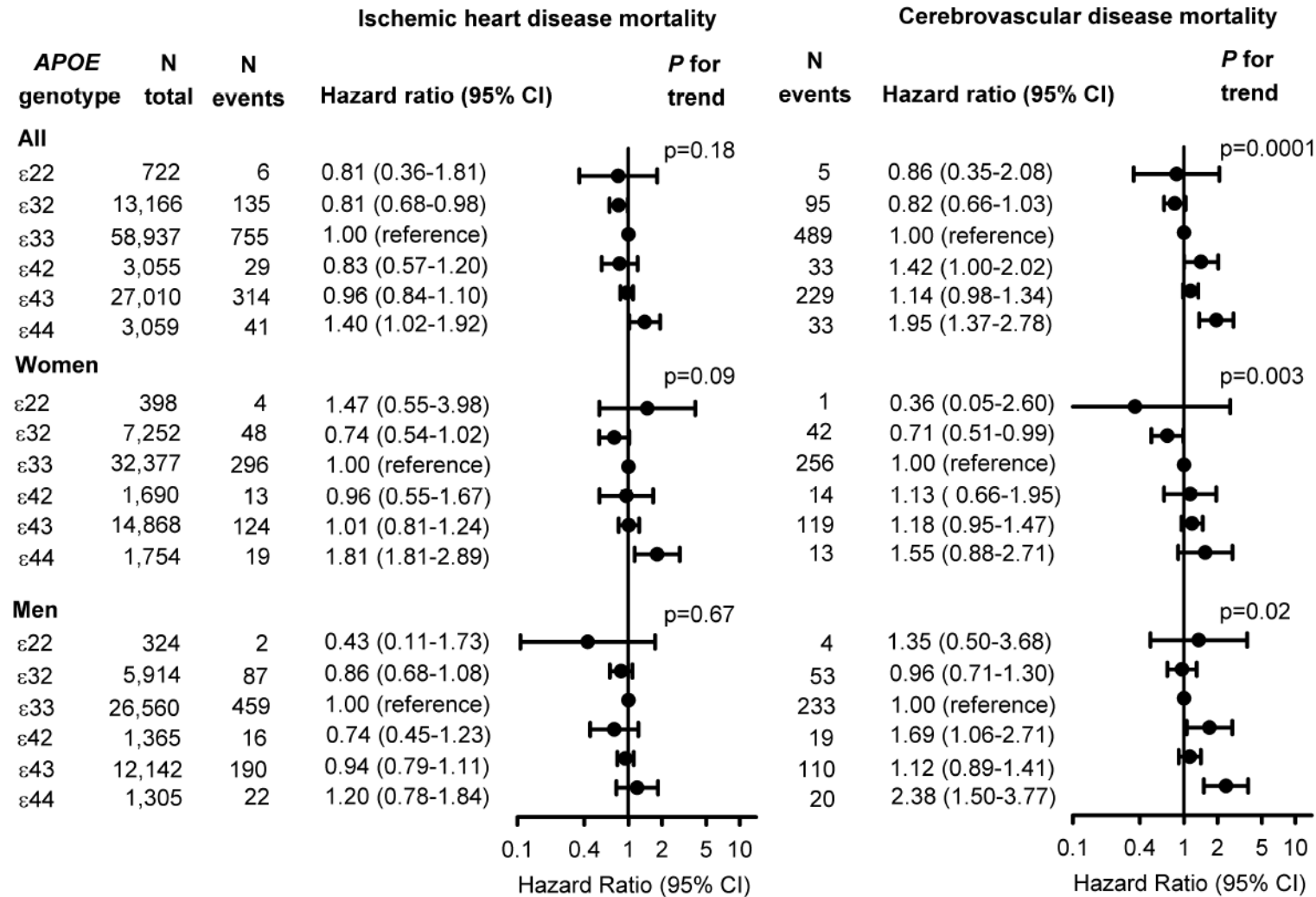
Cox regression models were multifactorially adjusted for age (time scale), sex, body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, hormonal replacement therapy, LDL cholesterol, HDL cholesterol, triglycerides and *APOE* genotype (all panels). 95% CI=95% confidence interval. apoE=plasma apolipoprotein E level.

Figure S13: Multifactorially adjusted hazard ratios for all-cause and cause-specific mortality for septiles of apolipoprotein E in individuals before and after change of autoanalyzer in the Copenhagen General Population Study.



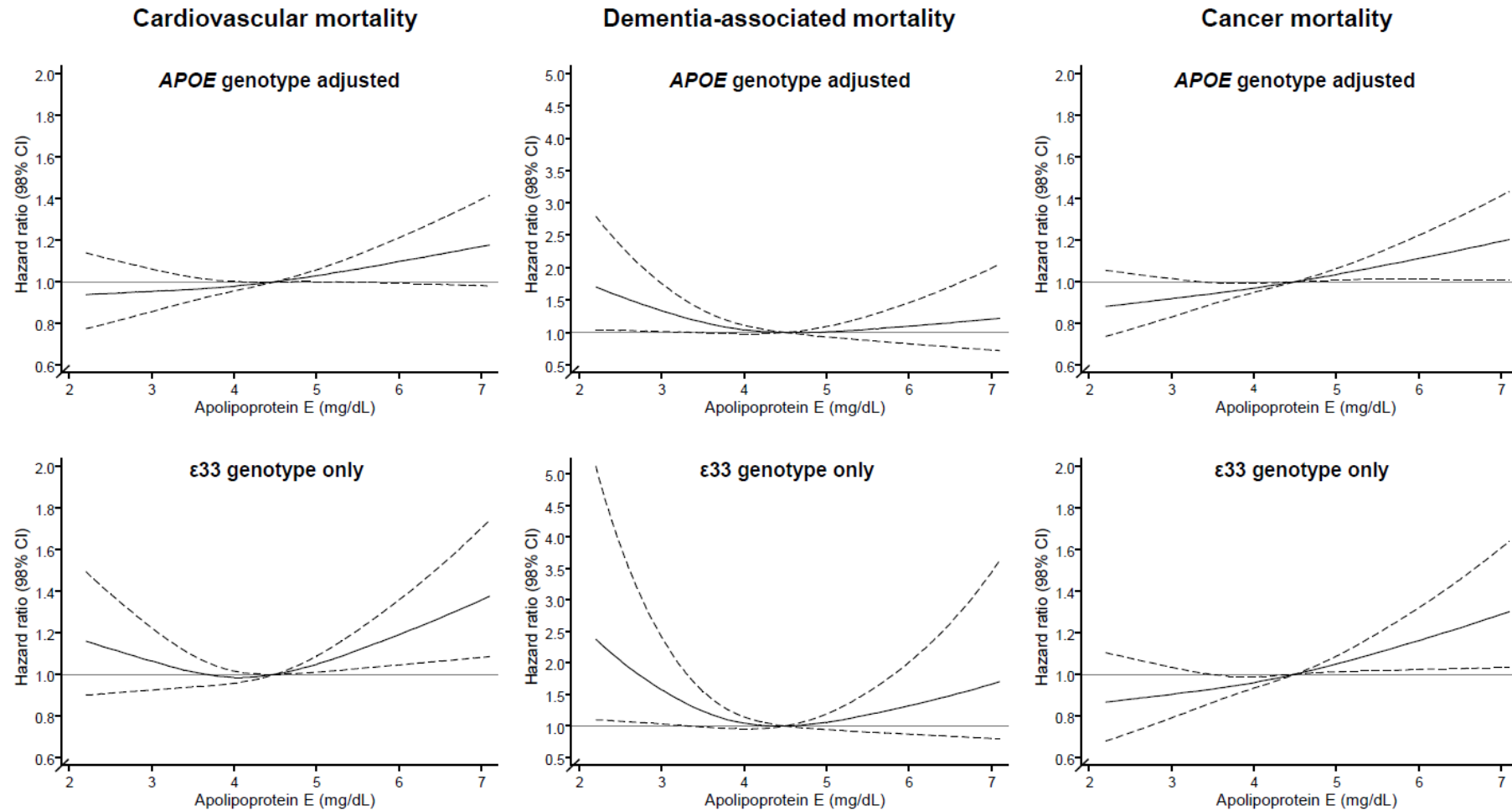
Cox regression models were multifactorially adjusted for age (time scale), sex, body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, hormonal replacement therapy, LDL cholesterol, HDL cholesterol, triglycerides and *APOE* genotype (all panels). 21,972 individuals with apoE measurements using the BNII autoanalyzer and 70,817 individuals with apoE measurements using the Kone autoanalyzer were included, thus excluding individuals from the transitional period. Methods for analyzing plasma apoE are detailed in *Rasmussen et al. Plasma levels of apolipoprotein E and risk of dementia in the general population. Ann Neurol. 2015;77:301-311*. 95% CI=95% confidence interval. apoE=plasma apolipoprotein E level.

Figure S14: Risk of ischemic heart disease and cerebrovascular disease mortality as a function of APOE genotype.



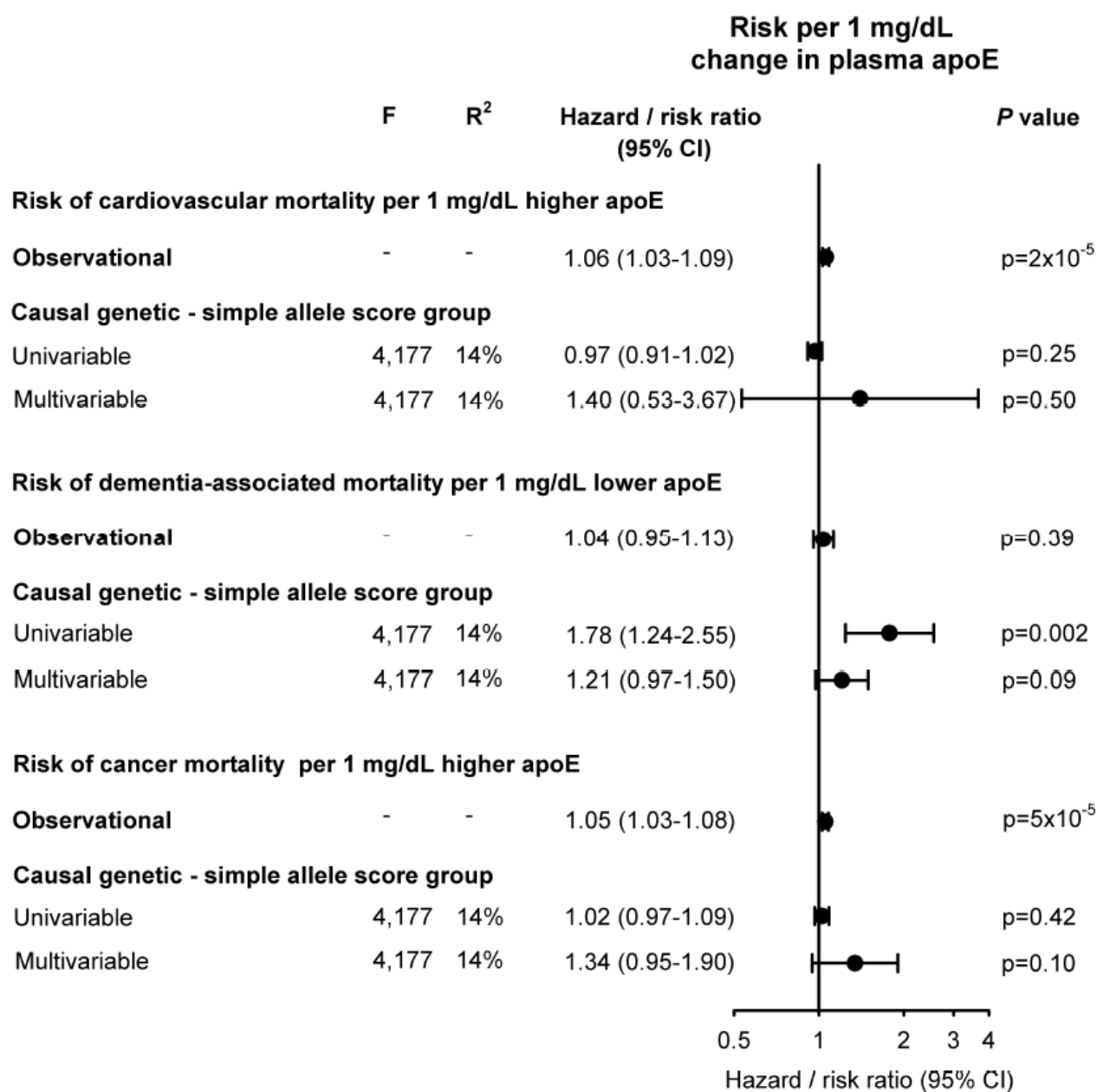
Cox regression models are shown for all (upper panel) and were further stratified in women (middle panel) and men (lower panel) and adjusted for age (time scale), body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, hormonal replacement therapy, LDL cholesterol, HDL cholesterol and triglycerides. APOE = APOE ε2/ε3/ε4 genotype. CI=95% confidence interval.

Figure S15: Multifactorially adjusted hazard ratios for cause-specific mortality according to plasma levels of apoE in individuals in the general population adjusted for multiple hypothesis testing.



For cause-specific mortality, there are three parallel analyses: when Bonferroni-correction is applied ($p=0.05/3$), therefore we have here applied the significance level $(1-0.05/3)*100\%=98\%$ for the confidence intervals. Solid lines are hazard ratios, whereas dashed lines indicate 98% confidence intervals derived from restricted cubic spline regression models otherwise similar to Figure 3 and 4 with respect to reference (4.5 mg/dL) and truncation (2.1-7.1 mg/dL) and multifactorial adjustment with adjustment for *APOE* genotype (upper panels) and in analyses restricted to $\epsilon 33$ carriers (bottom panels). 98% CI=98% confidence interval. *APOE*=*APOE* $\epsilon 2/\epsilon 3/\epsilon 4$ genotype. $\epsilon 33$ =*APOE* wildtype carriers.

Figure S16: Risk of cause-specific mortality for a 1mg/dL change in observational and causal, genetically determined plasma apolipoprotein E level for the simple allele score.



The hazard ratio for a 1 mg/dL change in observational plasma apoE was calculated using Cox regression with adjustment for age (time scale), body mass index, smoking, hypertension, diabetes, lipid-lowering therapy, alcohol consumption, physical inactivity, education, postmenopausal status, hormonal replacement therapy, LDL cholesterol, HDL cholesterol, plasma triglycerides and *APOE* genotype, whereas the corresponding risk ratios for the genetic change in plasma apoE for the simple allele score was derived from univariable and multivariable instrumental variable analyses, including LDL cholesterol and triglycerides in the multivariable analyses. For cancer mortality, simultaneous inclusion in the model of LDL cholesterol and triglycerides was not possible; the estimate with inclusion of LDL cholesterol is shown in above and the estimate with inclusion of triglycerides is 1.00 (0.92-1.07), with p=0.90. 74,560 individuals are included in these analyses. P value=significance of hazard ratios or risk ratios. F=strength of the genetic instrument (>10 indicates sufficient statistical strength). R²=percent contribution of genetic instrument to the variation in plasma apoE. apoE=apolipoprotein E. CI=confidence interval.