Appendix 1 (as supplied by the authors): Supplementary Methods

Cerebrovascular endpoints

For cerebrovascular disease validation of ICD codes has been described previously[1]. In brief, information on diagnoses of cerebrovascular disease (ICD8 431-438 and ICD10 I60-I69, G45), including ischemic cerebrovascular disease (transitory ischemic attacks, amaurosis fugax, and ischemic stroke) and hemorrhagic stroke, was collected by reviewing all hospital admissions and diagnoses entered in the national Danish Patient Registry and the national Danish Causes of Death Registry. Possible cerebrovascular events (hospitalized and nonhospitalized) were validated by trained physicians using the WHO definitions of cerebrovascular disease.

Supplementary cohort descriptions

The Copenhagen General Population Study
This prospective study of the Danish general population was initiated in 2003 with the first enrollment period from 2003 to 2015, and with follow-up examinations ongoing[2-5]. Data collection included a questionnaire, a physical examination, and blood sampling for biochemical and DNA analyses. We included 94,193 consecutive individuals in the current analyses; among these, 1,211 developed dementia and 6,106 developed cerebrovascular disease.

The Copenhagen City Heart Study
This prospective study of the Danish general population was initiated in 1976–78 with follow-up examinations in 1981–83, 1991–94, and 2001–03[2-5]. Participants were recruited and examined as in the CGPS. We included 10,344 individuals who gave blood for biochemical and DNA analyses at the 1991–94 or 2001–03 examinations; among these, 949 developed dementia and 1,414 developed cerebrovascular disease.

Biochemical and genetic analyses

Plasma total cholesterol, triglycerides and high-density lipoprotein (HDL) cholesterol were measured using colorimetric assays (Boehringer Mannheim GmbH, Mannheim, Germany; Konelab, ThermoFisher Scientific, Waltham, Massachusetts, USA). Low-density lipoprotein (LDL) cholesterol was calculated using the Friedewald equation[6] when plasma triglycerides were ≤4.0 mmol/L (≤352 mg/dL) and otherwise measured directly (Konelab). Apolipoprotein E was measured using nephelometry or turbidimetry (Dade Behring, Deerfield, Illinois, USA, or Dako, Glostrup, Denmark) as previously described[4].
An ABI PRISM 7900HT Sequence Detection System (Applied Biosystems Inc., Foster City, California, USA) and Taqman based assays were used to genotype for p.Cys130Arg (rs429358, legacy name Cys112Arg, c.388T>C) defining the ε4 allele and p.Arg176Cys (rs7412, legacy name Arg158Cys, c.526C>T) defining the ε2 allele. The genotyping call rate was 98%. Frequencies did not

deviate from the Hardy-Weinberg equilibrium (p=0.56 for p.Cys130Arg (ε4) and p=0.62 for p.Arg176Cys (ε2)).

Covariates

Body mass index was measured as weight in kilograms divided by measured height in meters squared.

Hypertension was defined as use of antihypertensive medication, systolic blood pressure of ≥140mmHg, and/or diastolic blood pressure of ≥90mmHg.

Diabetes mellitus was defined as self-reported disease, use of insulin or oral hypoglycemic agents, and/or nonfasting plasma glucose level of >11mmol/L (>198mg/dL).

Smoking was defined as current smoking.

Alcohol consumption was defined as >14/21U per week for women/men (1U=12g alcohol, equivalent to 1 glass of wine or spirit or 1 beer [33cl]).

Physical inactivity was defined as ≤4 hours per week of light physical activity in leisure time.

Women reported menopausal status and use of hormonal replacement therapy.

Lipid-lowering therapy was primarily statins (yes/no).

The cutoff for low education was 8 years.

Presence of atrial fibrillation was any diagnosis of atrial fibrillation ever.

Smoking, alcohol consumption, physical inactivity, menopausal status, hormonal replacement therapy, lipid-lowering therapy, and education were self-reported and reviewed together with an investigator on the day of attendance to clarify any uncertainties.

Statistical analysis

We used Stata/S.E. version 13.1 (Stata Corp., College Station, Texas, USA). Absolute 10 year risk of Alzheimer disease and all dementia as a function of *APOE* genotype and stratified by age (in three age-groups) and by sex was estimated by competing risk models[7]. Data was presented as estimated incidence rates (number of events per 10 years) in percent. Cumulative incidence curves and respective trends across genotypes were generated from competing risk regression models, according to the method of Fine and Gray[7]. Cumulative incidence was further evaluated by Kaplan-Meier curves and log-rank trend tests. To test whether *APOE* genotypes were associated with increased risk of dementia and cerebrovascular disease, we used Cox regression models adjusted for known biologically relevant risk factors and markers of life style regardless of the sizes of their contributions and as given in Figure legends.

Imputing
Missing data on categorical and continuous covariates (<0.4%) were imputed from age, sex, and population using multiple imputation with 10 imputations:multinomial logistic regression was applied for categorical variables and linear regression for continuous variables, and was performed using the "mi impute mlogit" and "mi impute chained (regress)" commands in Stata.

Cox regression and interaction
Cox proportional hazards regression models estimated hazard ratios for dementia and cerebrovascular disease. For Cox regression models, proportionality of hazards over time was assessed by plotting –ln(–ln[survival]) versus ln(analysis time), and tested using Schoenfeld residuals. No major violations of the proportional hazards assumption were observed.
Interactions between *APOE* genotype and sex, and between *APOE* genotype and age, on risk of dementia were evaluated by the inclusion of two-factor interaction terms in the Cox regression model, using a likelihood ratio test between models excluding and including the interaction term (Appendix 3).

Cumulative incidence curves
The cumulative incidence from a competing risk model does not overestimate the incidence of disease as the traditional Kaplan-Meier method does if competing risk of death is present. The cumulative incidence reflects estimates from the time point of the first registered event and therefore takes death in that exact age group into account. This is part of the explanation why the curves for subtypes of dementia cannot exactly be added up to the all dementia curve. We used first registered event from each specific diagnosis group; e.g. follow up time for Alzheimer disease was from the first time this diagnose was registered. A total of 45 individuals were diagnosed with both Alzheimer disease and vascular dementia, whereas 360 individuals were diagnosed with both Alzheimer disease and unspecified dementia somewhere along the course of disease. A total of 101 individuals were diagnosed with both vascular dementia and unspecified dementia. It would not be correct to exclude patients with unspecified dementia diagnoses that later also are diagnosed with Alzheimer disease or vascular dementia, or to exclude mixed events of Alzheimer/vascular origin, since such exclusions would generate a follow up bias. As a consequence of these facts, the numbers and the cumulative incidence curves will never add exactly up.

Competing risk models
Unadjusted and multivariable-adjusted competing risk models estimated risk of all dementia after a cerebrovascular event as a function of *APOE* genotype. The chosen endpoint was all dementia, since the presence of substantial concomitant cerebrovascular disease excludes an Alzheimer diagnosis according to diagnostic criteria[8].

References to Supplementary Methods:

1. Nordestgaard LT, Tybjaerg-Hansen A, Nordestgaard BG, Frikke-Schmidt R. Loss-of-function mutation in ABCA1 and risk of Alzheimer's disease and cerebrovascular disease. *Alzheimers Dement* 2015;11:1430-1438.

2. Frikke-Schmidt R, Nordestgaard BG, Stene MC et al. Association of loss-of-function mutations in the ABCA1 gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. *JAMA* 2008;299:2524-2532.

3. Jorgensen AB, Frikke-Schmidt R, Nordestgaard BG, Tybjaerg-Hansen A. Loss-of-function mutations in APOC3 and risk of ischemic vascular disease. *N Engl J Med* 2014;371:32-41.

4. Rasmussen KL, Tybjaerg-Hansen A, Nordestgaard BG, Frikke-Schmidt R. Plasma levels of apolipoprotein E and risk of dementia in the general population. *Ann Neurol* 2015;**77**:301-311.

5. Rasmussen KL, Tybjaerg-Hansen A, Nordestgaard BG, Frikke-Schmidt R. Plasma apolipoprotein E levels and risk of dementia: A Mendelian randomization study of 106,562 individuals. *Alzheimers Dement* 2018;14:71-80.

6. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972;18:499-502.

7. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc* 1999;94:496-509.

8. McKhann GM, Knopman DS, Chertkow H et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011;**7**:263-269.

Supplemental Table 1: Age- and sex-adjusted and multivariable-adjusted estimates for risk of dementia and cerebrovascular disease as a function of *APOE* genotype.

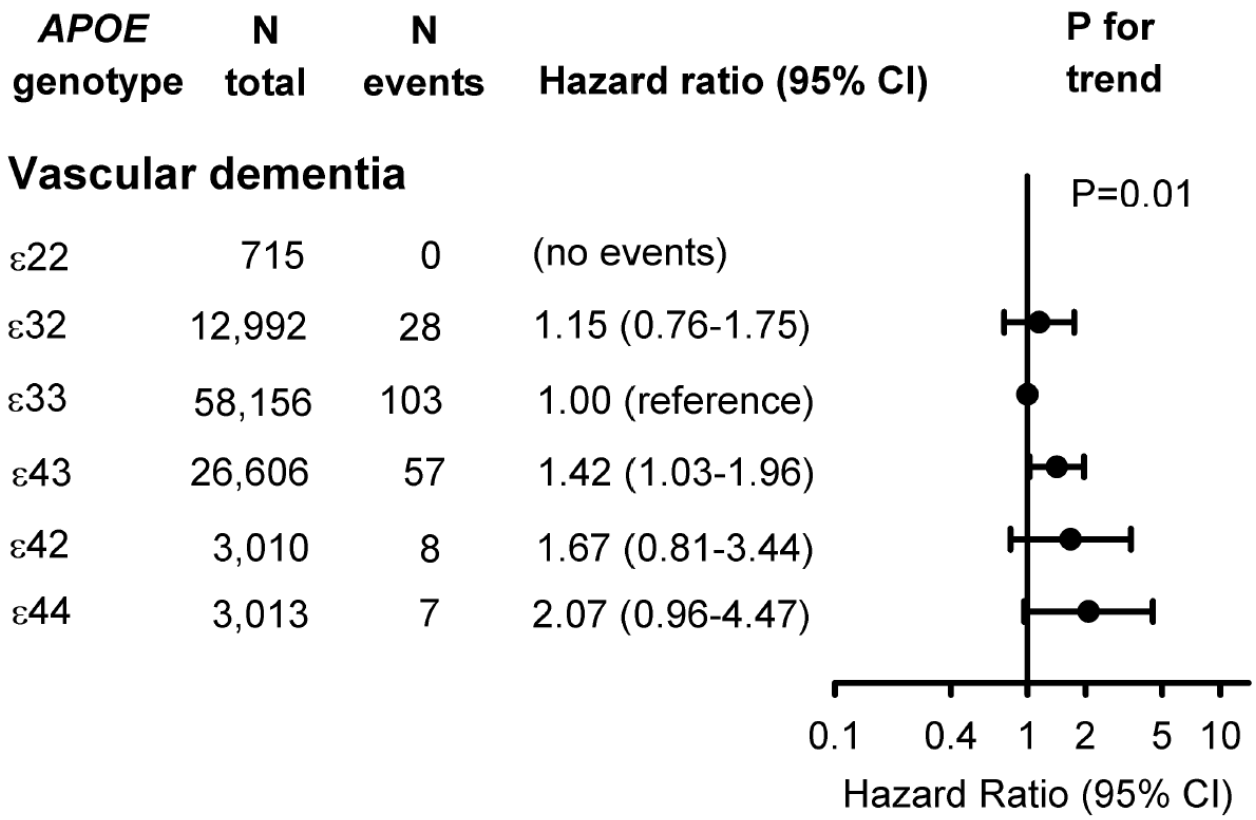
	Age- and sex adjusted hazard ratios (95% CI)	Multivariable-adjusted hazard ratios (95% CI)
Alzheimer disease	ε22: 1.34 (0.60-3.01)	ε22: 1.30 (0.58-2.91)
	ε32: 0.62 (0.48-0.82)	ε32: 0.62 (0.47-0.82)
	ε33: 1.00 (reference)	ε33: 1.00 (reference)
	ε42: 1.47 (1.00-2.16)	ε42: 1.48 (1.01-2.16)
	ε43: 2.48 (2.16-2.85)	ε43: 2.47 (2.15-2.84)
	ε44: 8.71 (7.06-10.75)	ε44: 8.74 (7.08-10.79)
Vascular dementia	ε22: (no events)	ε22: (no events)
	ε32: 1.07 (0.71-1.59)	ε32: 1.06 (0.71-1.58)
	ε33: 1.00 (reference)	ε33: 1.00 (reference)
	ε43: 1.65 (1.24-2.20)	ε43: 1.66 (1.24-2.21)
	ε42: 1.98 (1.07-3.67)	ε42: 1.99 (1.07-3.70)
	ε44: 2.82 (1.52-5.24)	ε44: 2.87 (1.54-5.33)
Unspecified dementia	ε22: 0.13 (0.02-0.93)	ε22: 0.13 (0.02-0.92)
	ε32: 0.83 (0.69-1.01)	ε32: 0.81 (0.67-0.98)
	ε33: 1.00 (reference)	ε33: 1.00 (reference)
	ε42: 1.34 (0.98-1.85)	ε42: 1.33 (0.97-1.83)
	ε43: 1.85 (1.64-2.09)	ε43: 1.85 (1.64-2.09)
	ε44: 4.50 (3.60-5.62)	ε44: 4.68 (3.74-5.85)
All dementia	ε22: 0.61 (0.29-1.28)	ε22: 0.60 (0.29-1.27)
	ε32: 0.81 (0.70-0.95)	ε32: 0.80 (0.69-0.94)
	ε33: 1.00 (reference)	ε33: 1.00 (reference)
	ε42: 1.35 (1.05-1.75)	ε42: 1.34 (1.04-1.74)
	ε43: 2.03 (1.84-2.23)	ε43: 2.03 (1.85-2.23)
	ε44: 5.62 (4.76-6.63)	ε44: 5.77 (4.89-6.81)
Ischemic stroke	ε22: 0.89 (0.58-1.37)	ε22: 0.94 (0.61-1.44)
	ε32: 0.92 (0.83-1.03)	ε32: 0.96 (0.87-1.07)
	ε33: 1.00 (reference)	ε33: 1.00 (reference)
	ε43: 1.03 (0.95-1.12)	ε43: 1.01 (0.94-1.10)
	ε42: 1.19 (0.99-1.44)	ε42: 1.22 (1.01-1.48)
	ε44: 1.33 (1.10-1.62)	ε44: 1.31 (1.08-1.58)
Ischemic cerebrovascular disease	ε22: 0.87 (0.63-1.21)	ε22: 0.92 (0.66-1.27)
	ε32: 0.96 (0.89-1.04)	ε32: 1.01 (0.93-1.09)
	ε33: 1.00 (reference)	ε33: 1.00 (reference)
	ε43: 1.03 (0.97-1.09)	ε43: 1.01 (0.95-1.07)
	ε42: 1.16 (1.01-1.34)	ε42: 1.19 (1.03-1.37)
	ε44:1.35 (1.17-1.56)	ε44: 1.29 (1.12-1.48)
Hemorrhagic stroke	ε22: 0.73 (0.27-1.95)	ε22: 0.75 (0.28-2.00)
	ε32: 1.07 (0.88-1.32)	ε32: 1.10 (0.89-1.35)
	ε33: 1.00 (reference)	ε33: 1.00 (reference)
	ε43: 0.99 (0.84-1.17)	ε43: 0.99 (0.84-1.17)
	ε42: 1.14 (0.77-1.69)	ε42: 1.15 (0.78-1.71)
	ε44: 1.83 (1.31-2.56)	ε44: 1.81 (1.29-2.53)
All cerebrovascular disease	ε22: 0.85 (0.63-1.15)	ε22: 0.89 (0.66-1.21)
	ε32: 0.96 (0.89-1.03)	ε32: 1.00 (0.93-1.08)
	ε33: 1.00 (reference)	ε33: 1.00 (reference)
	ε43: 1.03 (0.97-1.09)	ε43: 1.01 (0.95-1.06)
	ε42: 1.15 (1.01-1.31)	ε42: 1.17 (1.03-1.34)
	ε44: 1.36-1.20-1.55)	ε44: 1.30 (1.15-1.48)

Supplemental Table 2: Tests for interaction between *APOE* genotype and sex and between *APOE* genotype and age.

	P for interaction between *APOE* and sex	P for interaction between *APOE* and age
Alzheimer disease	0.03	0.31
Vascular dementia	0.77	0.95
Unspecified dementia	0.12	0.07
All dementia	0.05	0.08

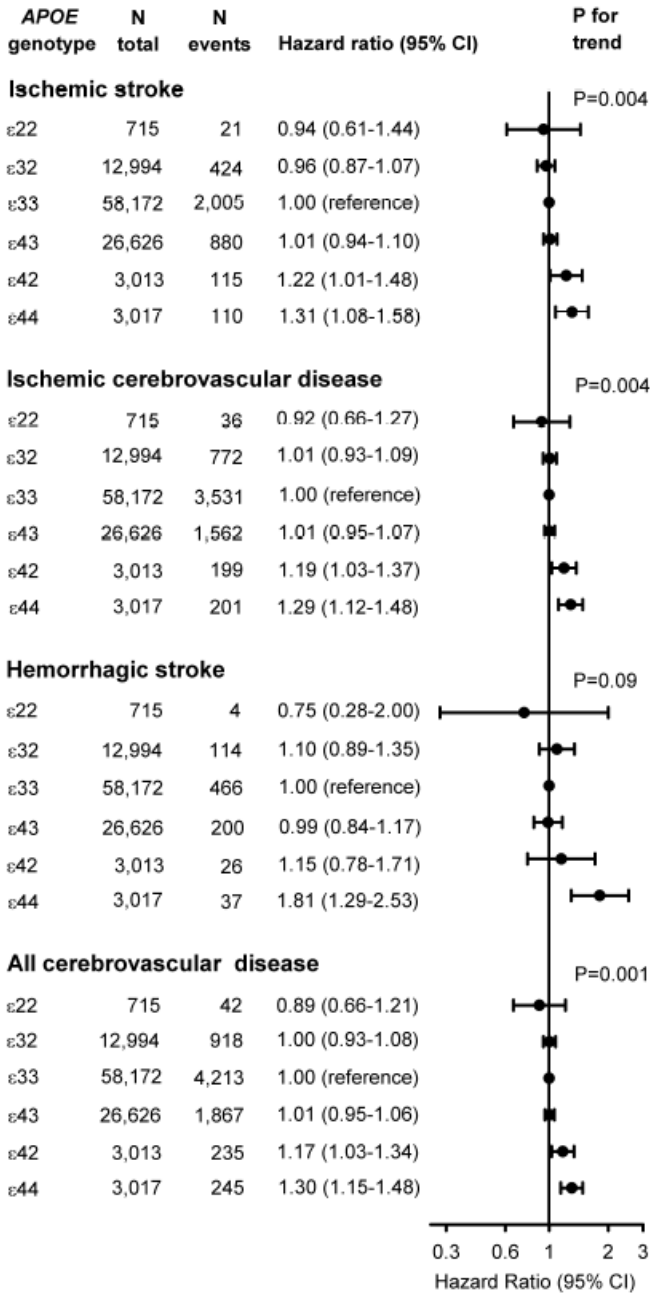
The models were adjusted for sex, body mass index, hypertension, diabetes mellitus, smoking, alcohol consumption, physical inactivity, menopausal status and hormonal replacement therapy (only women), lipid-lowering therapy, and education, using age as time scale when testing for interaction between *APOE* and sex, and using time since blood sampling and adjustment for age in four age groups (<60, 60-69, 70-79, 80+ years) when testing for interaction between *APOE* and age.

Supplemental Figure 1: Risk of vascular dementia as a function of *APOE* genotype after exclusion of individuals with both vascular dementia and Alzheimer disease.



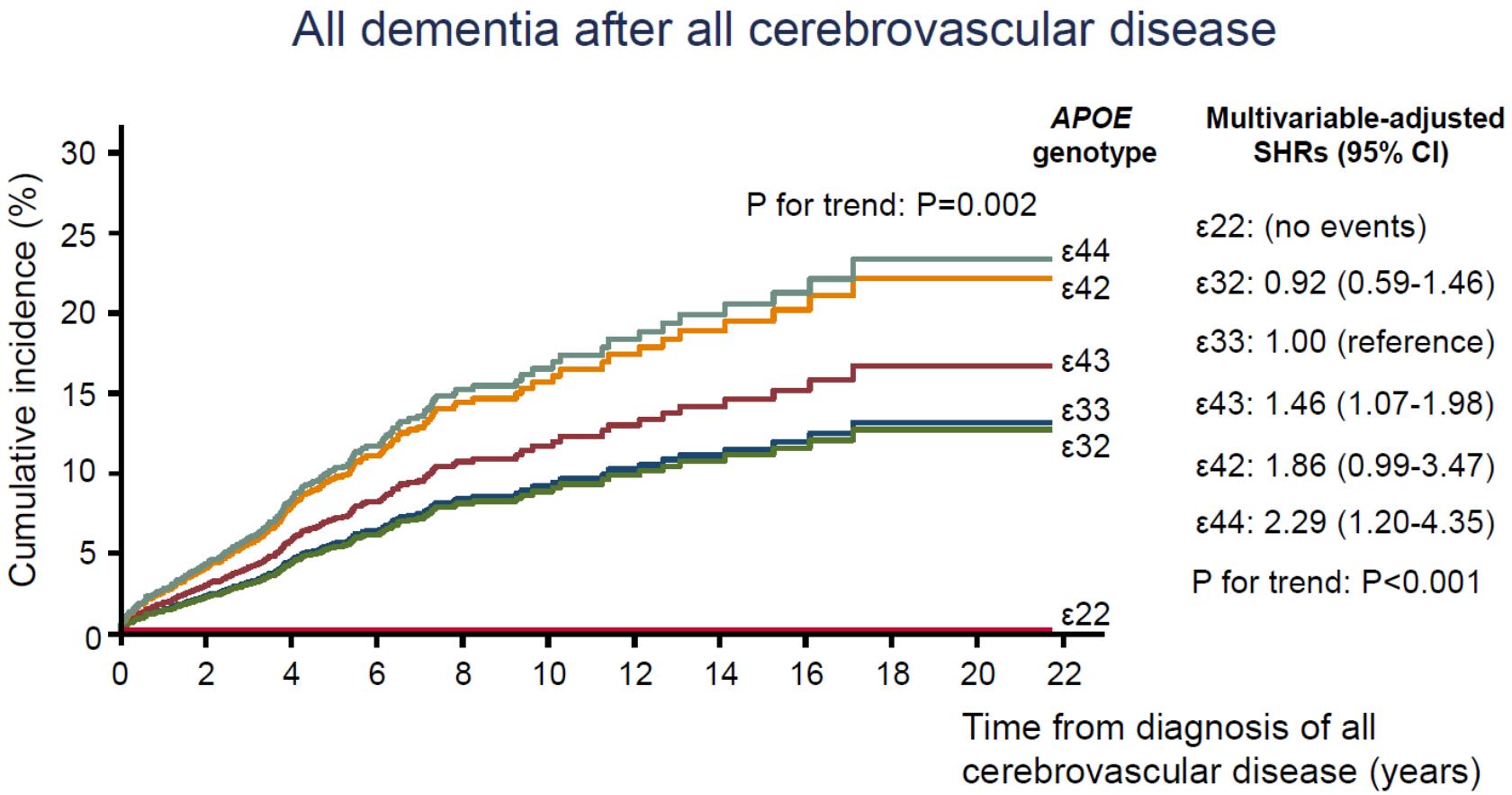
45 individuals with both vascular dementia and Alzheimer disease were excluded in these analyses. Hazard ratios were adjusted for age (time scale), sex, body mass index, hypertension, diabetes mellitus, smoking, alcohol consumption, physical inactivity, menopausal status and hormonal replacement therapy (only women), lipid-lowering therapy, and education. The p-value is the trend across genotypes from ε22 to ε32 to ε33 to ε43 to ε42 to ε44. CI=confidence interval.

Supplemental Figure 2: Risk of cerebrovascular disease as a function of *APOE* genotype.



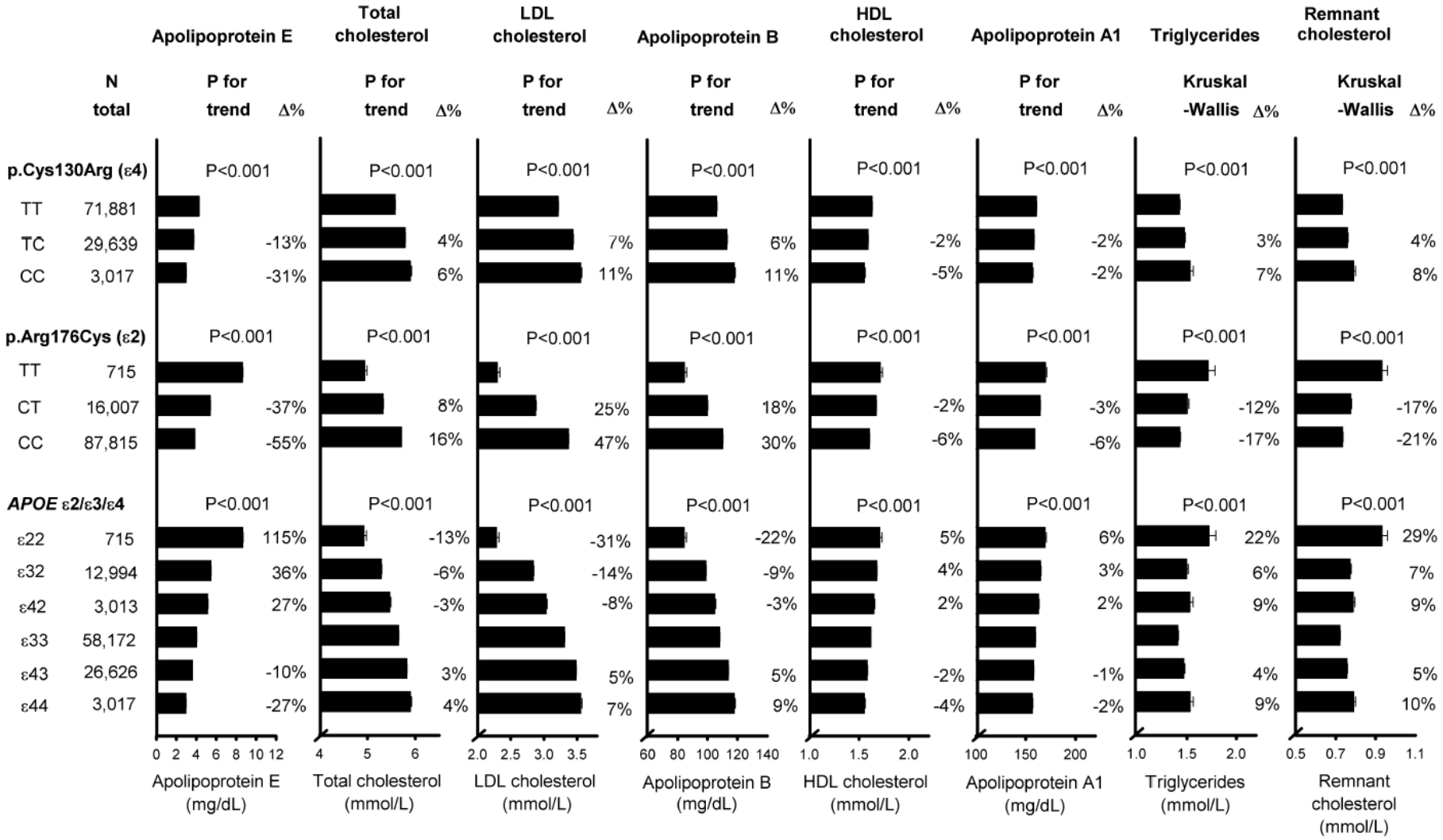
Hazard ratios were adjusted for age (time scale), sex, body mass index, hypertension, diabetes mellitus, smoking, alcohol consumption, physical inactivity, menopausal status and hormonal replacement therapy (only women), lipid-lowering therapy, education and atrial fibrillation. P-values are trends across genotypes from ε22 to ε32 to ε33 to ε43 to ε42 to ε44. CI=confidence interval.

Supplemental Figure 3: Cumulative incidence of all dementia as a function of time since diagnosis of all cerebrovascular disease and *APOE* genotype.



Analyses were performed for participants with an event of all cerebrovascular disease after the date of inclusion in the study (N=3,965), using competing risk models by the method of Fine and Gray to estimate the risk for future development of dementia. Cumulative incidence estimates and trend across genotypes from ε22 to ε32 to ε33 to ε43 to ε42 to ε44 for the unadjusted competing risk regression are shown (left panel). Subhazard ratios were adjusted for age, sex, body mass index, hypertension, diabetes mellitus, smoking, alcohol consumption, physical inactivity, menopausal status and hormonal replacement therapy (only women), lipid-lowering therapy, and education and trend across genotypes from ε22 to ε32 to ε33 to ε43 to ε42 to ε44 for the multivariable-adjusted competing risk regression is given (right panel). CI=confidence interval. SHRs=subhazard ratios. *APOE*=apolipoprotein E gene.

Supplemental Figure 4: Plasma levels of lipids, lipoproteins, and apolipoproteins as a function of *APOE* genotype.



Geometric mean±standard errors of the mean are given for apolipoprotein E and triglycerides; arithmetic mean±standard errors of the mean are given for total cholesterol, LDL cholesterol, apolipoprotein B, HDL cholesterol, apolipoprotein A1, and remnant cholesterol. Both the genotypes for the separate SNPs as well as the six common *APOE* ε2/ε3/ε4 genotypes are ordered from ε22 to ε32 to ε42 to ε33 to ε43 to ε44

for decreasing plasma apoE level according to the hypothesized hierarchy of associations for lipoprotein and apolipoprotein trends. Differences in plasma levels of lipids, lipoproteins and apolipoproteins are given in percent (Δ%); the genotype with the highest plasma apoE level serves as the reference for the two individual SNPs, whereas ε33 serves as the reference for the six common *APOE* genotypes. P for trend or Kruskal-Wallis analysis of variance are given for each SNP and for the six common *APOE* genotypes (from ε22 to ε32 to ε42 to ε33 to ε43 to ε44). Remnant cholesterol was calculated as total cholesterol minus HDL cholesterol minus LDL cholesterol. *APOE*=apolipoprotein E gene. Arg=arginine. Cys=cysteine. HDL=high-density lipoprotein. LDL=low-density lipoprotein.

Supplemental Figure 5: *APOE* ε2/ε3/ε4 genotype combinations and frequencies.

Arg176Cys (ε2) **Cys130Arg (ε4)**	**CC**	**CT**	**TT**
TT	**TTCC** ε33 58,172 (56)	**TTCT** ε32 12,994 (12)	**TTTT** ε22 715 (1)
TC	**TCCC** ε43 26,626 (25)	**TCCT** ε42 3,013 (3)	**TCTT** ε422 11 (0)
CC	**CCCC** ε44 3,017 (3)	**CCCT** ε442 15 (0)	**CCTT** ε4422 1 (0)

Values are number (%). Three homozygous phenotypes (apoE3/3, apoE4/4, apoE2/2) and three heterozygous phenotypes (apoE4/3, apoE3/2, apoE4/2) are encoded by the six common genotypes. The combination of genotypes for rs429358 (p.Cys130Arg, legacy name Cys112Arg, c.388T>C, ε4) and rs7412 (p.Arg176Cys, legacy name Arg158Cys, c.526C>T, ε2), yields nine possible genotype combinations. Arg=arginine. Cys=cysteine.