Supplementary material

Associations of serum testosterone and sex hormone-binding globulin with incident cardiovascular events in men.

Authors: Bu B. Yeap, Ross J. Marriott, Leen Antonio, Suchitra Raj, Girish Dwivedi, Christopher M. Reid, Bradley D. Anawalt, Shalender Bhasin, Adrian S. Dobs, David J. Handelsman, Graeme J. Hankey, Robin Haring, Alvin M. Matsumoto, Paul E. Norman, Terence W. O'Neill, Claes Ohlsson, Eric S. Orwoll, Dirk Vanderschueren, Gary A. Wittert, Frederick C.W. Wu, Kevin Murray.

Methods

Exposures

Serum total testosterone was assayed using a competitive binding chemiluminescent immunoassay, analytical range 0.35-55.5 nmol/L (10-1,599 ng/dL, DXI 800, Beckman Coulter, UK) (1,2). Coefficients of variation were 8.3%, 3.7% and 4.2% for testosterone concentrations in low, medium and high ranges (1.0-2.2, 13.4-22.8 and 29.3-49.4 nmol/L, or 29-63, 386-657 and 844-1,424 ng/dL, respectively). Serum SHBG was assayed using a two-step sandwich chemiluminescent immunoassay, analytical range 0.33-242 nmol/L (DXI 800, Beckman Coulter, UK). Coefficients of variation were 5.7%, 5.3% and 5.2% for SHBG in low, medium and high ranges (15.0-27.7, 31.9-55.5 and 56.3-87.8 nmol/L).[1,2]

Covariates

Qualifications were categorised as below A-levels (high school), completed A-levels, completed college/university, or completed other professional qualification (not school/college/university). Alcohol consumption was categorised as abstainers <1, low 1-<7,

moderate 7-<14, medium 14-<21, high ≥21 units/week (1 unit=8 g pure alcohol).[3] Diet categories were high red meat (beef, lamb and pork >3 times/week), low red meat (≤3/week), poultry (no red meat), fish (no red meat or poultry), vegetarians (no red meat, poultry or fish), vegan (no red meat, poultry, fish, eggs or diary).[4] Vigorous and moderate physical activity undertaken per week was categorised to World Health Organization recommendations as sufficient: ≥75 minutes vigorous or ≥150 minutes moderate, or equivalent combination; insufficient: less than this; additional: ≥150 minutes vigorous- or ≥300 minutes moderate-level or equivalent combination (for "additional" health benefits).[5]

Diabetes was determined as blood glucose ≥11.1 mmol/L (200 mg/dL) unfasted or ≥7.0 mmol/L (126 mg/dL) fasting, or glycated haemoglobin ≥48 mmol/mol (6.5%), or previous hospital admission diagnosis, or insulin use, or self-report condition. Dementia was determined as use of anti-dementia medication(s), or previous hospital admission diagnosis, or self-report condition. Renal impairment was determined as creatinine >150 umol/L, or previous hospital admission diagnosis or self-report of renal/kidney failure, dialysis, nephropathy or nephritis. Hypertension was determined as systolic blood pressure ≥140 or diastolic ≥90 mmHg or taking blood pressure medication, prior hospital admission diagnosis, or self-report condition. Dyslipidemia was categorised by use of lipid medications, and number of medications was recoded into categories of 0, 1-2, 3-4, 5+ medications taken, consistent with recent NHS reporting.[6,7]

Statistical analysis

Median follow-up time was calculated using the reverse Kaplan-Meier method.[8] Cox proportional hazards models were fitted using the rms package in R version 4.0.2.[9,10] Site was modelled as a stratified factor to account for the relatedness of observations taken for

individuals recruited to different UK Biobank assessment centres at baseline. Continuous explanatory variables were modelled using pre-specified restricted cubic splines to account for the prospect of non-linearities in modelled associations. Restricted cubic splines were used to model non-linear associations with continuous covariates, with 3 internal knots at the 27.5th, 50th, 72.5th percentiles and linear constraints outside of the outer knots at the 5th and 95th percentiles of the marginal distribution of these covariates.

Complete-case analyses were done in the first instance. Analyses were repeated for: (i) multiply-imputed datasets; (ii) data excluding participants who recorded an event within the first two years or had less than two years of follow-up, for inferring the potential effects of reverse causation; and (iii) multivariable models without factor terms demonstrating a sparse distribution of events among the factor levels (i.e., by dropping the Dementia, Diet, Ethnicity, and HIV terms). For (i), Multiple Imputation using Chained Equations (MICE) was used to impute the missing values.[11,12] Five imputed datasets were initially generated from 20 MICE iterations for each analysis, including the following variables: log(Y), δ, hormone variables (testosterone, SHBG), participants' age, ethnicity, living with partner, qualifications, alcohol consumption, dietary patterns, total cholesterol, physical activity, smoking status, BMI, waist circumference, assessment centre (site), medication variables (anticonvulsants, lipid, glucocorticoids, opioids, total number of medications), and prevalent health condition variables (angina, atrial fibrillation, CVD, COPD, dementia, diabetes, HIV, hypertension, liver disease, renal impairment, thyroid disease). Region (i.e., instead of site: UK, Scotland, Wales) was included in multiple imputations for analyses of haemorrhagic stroke, ischaemic stroke, and heart failure instead of site because there were zero events for one of the levels of site (Wrexham). Diet was dropped from analyses of haemorrhagic stroke because of a failure of model converge due to zero events for fish eaters and vegans. The Y and δ variables are the

minimum of the event and censoring times, and the censoring variable, respectively. Logistic regression was used for δ and factors with two levels, polytomous regression used for factors with more than two levels, and predictive mean matching used for continuous covariates. The estimated minimum number of imputed datasets was determined using the formula of von Hippel applied to the log hazard ratio estimates from the first five imputed datasets.[13] The estimate of the minimum number of imputations exceeded five for multivariable analyses of the association of testosterone with myocardial infarction ($\widehat{m}$=14.6), heart failure ($\widehat{m}$=17.4), and MACE ($\widehat{m}$=12.7), and so an additional set of 10, 15, and 10 MICE imputations were conducted for those analyses. Multiply imputed estimates were pooled using Rubin's rules,[14] and, for tests of association, likelihood ratio statistics were pooled using the method of Meng and Rubin.[15]

Prior to fitting models, the number of events per variable (EPV[16]) was calculated for the most complex analysis model (Multivariable 2), as a guide to indicate the prospect for model overfitting, which could potentially lead to biased estimates (Supplementary Table S2). Since the EPV was much lower than the recommended value of at least 10 for the haemorrhagic stroke outcome, internal validation using 200 bootstrapped replicates of each fitted model was done to calculate statistics for reporting the extent of overfitting bias (optimism of the calibration slope) and discrimination (Somers' *D*).[9] The validity of the proportional hazards assumption was assessed using per-variable and global tests.[17] Plots of the Schoenfeld residuals, with estimated coefficients and 95% confidence intervals plotted against follow-up times, were inspected for statistically significant results. Test results were ignored when the detected temporal variation for proportional hazard estimates was shown to be negligible.[18]

Correlations of estimated coefficients with coefficients for the exposure variable were assessed and the maximum of the absolute values was 0.2 for all analyses.

Supplementary analyses were conducted to better understand the results of the multivariable Cox Proportional Hazards analyses investigating the associations of baseline total testosterone concentration with incident CVD events. Specifically, we explored to what extent the inclusion of BMI and age in multivariable modelling influenced findings on the association of testosterone with each of the CVD outcomes: MI, HS, IS, HF and MACE. We explored the addition of BMI to the Univariable model to show any effect of adding BMI on its own, and of dropping BMI from the Multivariable 2 model, to explore the relative importance of BMI, after controlling for all other model predictors. The statistical models for this analysis are presented in Supplementary Table S3. An additional set of analyses were done dropping both BMI and waist from Multivariable 2, as these were correlated variables, and except for those few cases mentioned, the results were substantively the same (data not shown). The procedure was repeated for analyses involving SHBG, cFT and FTZ.

References

1. UK Biobank. Biomarker assay quality procedures: approaches used to minimise systematic and random errors (and the wider epidemiological implications). Version 1.2, date 2 April 2019, 1–15. http://www.ukbiobank.ac.uk/ (accessed 31 January 2020).
2. Fry D, Almond R, Moffat S, Gordon M, Singh P. UK Biobank Biomarker Project. Companion document to accompany serum biomarker data. Version 1·0, date 11 March 2019, 1–16. http://www.ukbiobank.ac.uk/ (accessed 31 January 2020).

3. Zhao L, Matloff W, Ning K, Kim H, Dinov ID, Toga AW. Age-related differences in brain morphology and the modifiers in middle-aged and older adults. Cereb Cortex 2019; 29: 4169–4193.
4. Bradbury KE, Tong TYN, Key TJ. Dietary intake of high-protein foods and other major foods in meat-eaters, poultry-eaters, fish-eaters, vegetarians, and vegans in UK Biobank. Nutrients 2017; 9: E1317.
5. Global Recommendations on Physical Activity for Health: 18-64 years old. World Health Organisation, 2011. https://www.who.int/dietphysicalactivity/pa/en/ (accessed 18 February, 2020).
6. Zanetti D, Bergman H, Burgess S, Assimes TL, Bhalla V, Ingelsson E. Urinary albumin, sodium, and potassium and cardiovascular outcomes in the UK Biobank: Observation and Mendelian Randomization analyses. Hypertension 2020; 75: 714–722.
7. Moody A, Mindell J, Faulding S. Health Survey for England 2016. Prescribed medicines. NHS Digital. http://healthsurvey.hscic.gov.uk/media/63790/HSE2016-pres-med.pdf. December 2016 (accessed May 5, 2020).
8. Schemper, M, Smith, TL. A note on quantifying follow-up in studies of failure time. Control Clin Trials 1996; 17(4): 343–346.
9. Harrell FE. Jr. rms: Regression Modeling Strategies. R package version 6.0-1. 2020. https://CRAN.R-project.org/package=rms
10. R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2020. URL https://www.R-project.org/.
11. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res. 2011; 20(1): 40-9.
12. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. J Statistical Software. 2011; 45: 3.

13. von Hippel PT. How many imputations do you need? A two-stage calculation using a quadratic rule. Sociological Methods & Research. 2018; 49(3): 699-718.
14. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley, 1987.
15. Meng X, Rubin DB. Performing likelihood ratio tests with multiply-imputed data sets. Biometrika 1992; 79(1): 103-111.
16. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates. J Clin Epidemiol 1995; 48(12): 1503–1510.
17. Grambsch P, Therneau T. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 1994; 81: 515–526.
18. Therneau TM, Grambsch PM. Modeling survival data: extending the Cox model. Springer-Verlag: New York 2000.

Table S1. Hospital diagnosis and death registry codes used to identify incident health conditions of interest or prevalent conditions, for model covariates or exclusions.*

Health condition	**ICD-9**	**ICD-10**
Myocardial infarction**	ICD-10 only for incident outcomes	I21, I22
Haemorrhagic stroke**	ICD-10 only for incident outcomes	I60, I61
Ischaemic stroke**	ICD-10 only for incident outcomes	I63, I64
Heart failure**	ICD-10 only for incident outcomes	I50, I42.0
CVD death**	ICD-10 only for incident outcomes	I20, I21, I22, I23, I24, I25, I42.0, I46, I50, I60, I61, I63, I64, I70, I71
Major adverse cardiovascular event (MACE)**	ICD-10 only for incident outcomes	Myocardial infarction or Ischaemic stroke or CVD death
Androgenital / testicular disorders	255.2, 257	E25, E29
Pituitary disease	253.0-253.4, 253.7	E22.0, E22.1, E22.8, E22.9, E23.0, E22.1, E23.3, E24
Angina	413.0, 413.1, 413.9	I20.0, I20.1, I20.8, I20.9
Atrial Fibrillation	427.3	I48, and, where the 4th digit was provided***, I48.0, I48.1, I48.2, I48.9
COPD	490, 491, 492, 494, 496	J43.1, J43.2, J43.8, J43.9, J44.0, J44.1, J44.8, J44.9
CVD	410, 428, 430-437	I21, I22, I50, I60-I64, I69.0-I69.4
Diabetes	250.0 - 250.9	E10 - E14
Dementia	290.0, 290.1, 290.2, 290.3, 290.4, 290.8, 290.9, 294.1, 294.8, 331.0, 331.1, 331.2, 331.4, 331.82	F00, F01, F02.0, F02.2, F02.3, F02.8, F03, F05.1, G10, G30, G31.0, G31.1, G31.8, G31.9
HIV	042	B20-B24
Hypertension	401	I10
Liver disease	570-573	K70-K76

Renal Impairment	582, 583, 585, 586	N03, N04, N05, N08.1-NO8.3, NO8.5, N11.1, N11.8, N11.9, N14, N18.2, N18.3, N18.4, N18.5, N18.6, N18.9, N19
Thyroid disease	240-245	E00-E06

* = Additional data sources were also used for identifying prevalent conditions (e.g., from self-report medical conditions, self-report medication usage, physical examination and blood chemistry measurements).

** = Only the principal diagnosis (hospital admissions) or main/underlying cause of death (death registry data) was used to identify each incident event (secondary diagnoses were not used for incident events).

*** = In many cases only "I48" was provided but in others the full 4 digit code was provided.

Table S2. Events per variable and bootstrap validation statistics of Multivariable 2 for analyses of testosterone.

		Bootstrap validation statistics[§]			
		Calibration Slope		Somers' *D*	
Outcome	**EPV***	**Estimate**	**Optimism**	**Estimate**	**Optimism**
Complete case analyses					
Myocardial infarction	38.5	0.965	0.035	0.440	0.014
Haemorrhagic stroke	5.2	0.771	0.229	0.304	0.105
Ischaemic stroke	20.5	0.952	0.048	0.494	0.022
Heart failure	10.1	0.963	0.037	0.702	0.016
MACE	66.7	0.977	0.023	0.454	0.011
Multiple imputation analyses					
Myocardial infarction	58.9	0.976	0.024	0.458	0.012
Haemorrhagic stroke	7.3	0.826	0.174	0.329	0.073
Ischaemic stroke	30.9	0.968	0.032	0.505	0.014
Heart failure	17.3	0.977	0.023	0.707	0.015
MACE	102.1	0.982	0.018	0.474	0.007

* 'EPV' = events per variable, which was used as an initial guide to likely model overfitting, with an EPV of at least 10 generally considered acceptable.[17] Diet was dropped from the model fitted in the analysis of haemorrhagic strokes because of zero events for fish eaters and vegans, which led to convergence failure.

'Calibration Slope' measures the agreement between model predictions and test set observations, with a value of 1 indicating no average bias;[20] Somers' *D* measures overall discrimination of fit;[21] 'Estimate' is the optimism-corrected estimate; 'Optimism' measures the extent of overfitting in the estimate, which is the difference between that obtained from the fit of the model to 200 bootstrapped 'training' datasets and then applied to describe the original (test) data set, as a method of "internal model validation".[20] Accordingly, the Optimism for Calibration Slope measures the extent of bias in model estimates due to overfitting.

§ = Validation statistics for the fit to multiply-imputed data, obtained from the fit to the first imputed dataset.

Table S3. Statistical models to investigate the importance of including BMI as a term in analyses of the association of testosterone with incident CVD[a]

Model	Model terms[b]		
	rcs(T)[c]	rcs(BMI)	All other predictors[d]
Univariable	✓		
Univar+BMI	✓	✓	
Multivariable 2	✓	✓	✓
Multivar2-BMI	✓		✓

[a] = Time to incident myocardial infarction, haemorrhagic stroke, ischaemic stroke, heart failure, or MACE, with right censoring applied to participants with no event observed during follow-up.

[b] = To be consistent with the main analyses, all Cox Proportional Hazards models include site (or region) as a stratification variable (refer to Methods in main text).

[c] = 'rcs(·)' denotes a restricted cubic spline with 3 internal knots at the 27.5th, 50th, 72.5th percentiles and linear constraints outside of the outer knots at the 5th and 95th percentiles of the marginal distribution of this covariate; 'T' = total testosterone (nmol/L).

[d] = Age, ethnicity, living with partner, qualifications, alcohol consumption, smoking status, diet, physical activity, waist circumference, cholesterol, prevalent CVD, diabetes, hypertension, angina, atrial fibrillation, COPD, dementia, renal impairment, liver disease, thyroid disease, HIV, use of lipid medications (a proxy for hyperlipidemia), glucocorticoids, opioids, anticonvulsants, the number of medications (proxy for overall comorbidity status), SHBG.

The analyses were also repeated substituting: (i) age for BMI and (ii) SHBG, cFT and FTZ for T.

Table S4. Hazard ratios of different types of incident cardiovascular disease events by quintiles of testosterone (T): Pooled estimates from analysis of multiply-imputed data.

Model	Q5 (highest T) (n = 41,675) [§]	Q4 (n = 41,673)	Q3 (n = 41,686)	Q2 (n = 41,686)	Q1 (lowest T) (n = 41,699)	Overall trend *P* Value
Myocardial infarction: 3,771 events						
	646 events	*722 events*	*771 events*	*784 events*	*848 events*	
Univariable	ref.	1.08 (1.02-1.14)	1.18 (1.06-1.31)	1.20 (1.09-1.31)	1.29 (1.19-1.41)	< 0.001*
Multivariable 1[#]	ref.	1.01 (0.95-1.07)	1.04 (0.93-1.16)	1.00 (0.91-1.10)	0.96 (0.88-1.06)	0.544
Multivariable 2	ref.	1.00 (0.94-1.06)	1.01 (0.90-1.13)	0.94 (0.85-1.05)	0.88 (0.79-0.99)	0.078
Haemorrhagic stroke: 466 events						
	95 events	*94 events*	*95 events*	*90 events*	*92 events*	
Univariable	ref.	0.93 (0.79-1.09)	0.86 (0.64-1.15)	0.95 (0.74-1.21)	0.95 (0.74-1.21)	0.845
Multivariable 1	ref.	0.89 (0.76-1.05)	0.79 (0.58-1.06)	0.84 (0.65-1.08)	0.73 (0.56-0.95)	0.079
Multivariable 2	ref.	0.91 (0.76-1.08)	0.82 (0.60-1.13)	0.91 (0.68-1.22)	0.84 (0.61-1.16)	0.470
Ischaemic stroke: 1,975 events						
	373 events	*367 events*	*395 events*	*420 events*	*420 events*	
Univariable	ref.	1.02 (0.95-1.10)	1.09 (0.94-1.26)	1.11 (0.98-1.25)	1.19 (1.06-1.34)	0.018
Multivariable 1	ref.	0.98 (0.90-1.06)	0.98 (0.85-1.14)	0.95 (0.83-1.07)	0.87 (0.76-0.99)	0.056
Multivariable 2	ref.	1.00 (0.92-1.09)	1.03 (0.88-1.21)	1.02 (0.88-1.19)	0.98 (0.83-1.14)	0.759
Heart failure: 1,110 events						
	188 events	*162 events*	*188 events*	*203 events*	*369 events*	
Univariable	ref.	1.00 (0.89-1.12)	1.02 (0.83-1.26)	1.19 (1.00-1.42)	1.79 (1.53-2.10)	< 0.001*
Multivariable 1	ref.	0.90 (0.80-1.00)	0.79 (0.64-0.97)	0.77 (0.64-0.93)	0.86 (0.72-1.02)	0.017
Multivariable 2	ref.	0.91 (0.80-1.03)	0.81 (0.65-1.02)	0.82 (0.67-1.02)	0.94 (0.76-1.17)	0.015
MACE: 6,536 events						
	1,194 events	*1,227 events*	*1,310 events*	*1,348 events*	*1,457 events*	
Univariable	ref.	1.04 (0.99-1.08)	1.09 (1.01-1.19)	1.11 (1.04-1.19)	1.23 (1.15-1.32)	< 0.001*
Multivariable 1	ref.	0.98 (0.94-1.03)	0.97 (0.89-1.05)	0.93 (0.87-1.00)	0.89 (0.83-0.96)	0.013
Multivariable 2	ref.	0.99 (0.94-1.04)	0.98 (0.90-1.07)	0.94 (0.87-1.03)	0.91 (0.83-0.99)	0.135

[§] = Quintile boundaries Q1/2 8.9 nmol/L (256 ng/dL), Q2/3 10.8 nmol/L (311 ng/dL), Q3/4 12.5 nmol/L (360 ng/dL) and Q4/5 14.8 nmol/L (427 ng/dL). 2.5th percentile = 5.9 nmol/L (170 ng/dL), 97.5th = 20.1 nmol/L (579 ng/dL); Presented numbers are for imputed data, after exclusions.

[#] = Multivariable 1 models included age, ethnicity, living with partner, qualifications, alcohol consumption, smoking status, diet, physical activity, BMI, waist circumference, cholesterol, prevalent CVD, diabetes, hypertension, angina, atrial fibrillation, COPD, dementia, renal impairment, liver disease, thyroid disease, HIV, and use of lipid medications (a proxy for hyperlipidemia), glucocorticoids, opioids, and anticonvulsants, with the number of medications included as a proxy for overall comorbidity status and assessment centre as a stratification factor. Multivariable 2 models for testosterone included SHBG as an additional covariate.
* = Result interpreted as significant.

Table S5. Hazard ratios of different types of incident cardiovascular disease events by quintiles of testosterone (T): Analysis of data excluding participants who recorded an event within the first two years or had less than two years of follow-up.

Model	Q5 (highest T)[§]	Q4	Q3	Q2	Q1 (lowest T)	Overall trend *P* Value
Myocardial infarction: 1,904 events						
	n = 30,199	**n = 30,271**	**n = 29,981**	**n = 29,515**	**n = 28,142**	
	333 events	***382 events***	***396 events***	***402 events***	***391 events***	
Univariable	ref.	1.08 (1.00-1.18)	1.17 (1.01-1.36)	1.18 (1.04-1.34)	1.26 (1.11-1.42)	0.006*
Multivariable 1[#]	ref.	1.02 (0.93-1.11)	1.03 (0.88-1.20)	0.98 (0.86-1.11)	0.93 (0.82-1.06)	0.621
Multivariable 2	ref.	0.99 (0.90-1.08)	0.97 (0.82-1.14)	0.88 (0.76-1.02)	0.81 (0.69-0.96)	0.076
Haemorrhagic stroke: 237 events						
	n = 30,272	**n = 30,371**	**n = 30,070**	**n = 29,602**	**n = 28,237**	
	49 events	***49 events***	***45 events***	***48 events***	***46 events***	
Univariable	ref.	0.84 (0.67-1.07)	0.79 (0.52-1.20)	1.00 (0.71-1.41)	0.96 (0.68-1.35)	0.575
Multivariable 1	ref.	0.82 (0.64-1.03)	0.73 (0.48-1.12)	0.92 (0.64-1.31)	0.78 (0.53-1.13)	0.204
Multivariable 2	ref.	0.83 (0.65-1.07)	0.78 (0.50-1.23)	1.06 (0.70-1.59)	1.00 (0.64-1.57)	0.369
Ischaemic stroke: 1,050 events						
	n = 30,241	**n = 30,335**	**n = 30,048**	**n = 28,556**	**n = 28,193**	
	209 events	***207 events***	***202 events***	***226 events***	***206 events***	
Univariable	ref.	0.98 (0.88-1.09)	1.00 (0.82-1.22)	1.06 (0.90-1.26)	1.15 (0.97-1.35)	0.353
Multivariable 1	ref.	0.93 (0.83-1.04)	0.88 (0.72-1.08)	0.90 (0.75-1.07)	0.84 (0.70-1.00)	0.252
Multivariable 2	ref.	0.94 (0.83-1.05)	0.89 (0.72-1.11)	0.93 (0.76-1.13)	0.90 (0.73-1.12)	0.798
Heart failure: 518 events						
	n = 30,265	**n = 30,362**	**n = 30,059**	**n = 29,594**	**n = 28,213**	
	91 events	***83 events***	***81 events***	***96 events***	***167 events***	
Univariable	ref.	0.95 (0.80-1.12)	0.92 (0.67-1.25)	1.14 (0.89-1.47)	1.78 (1.41-2.24)	< 0.001*
Multivariable 1	ref.	0.83 (0.70-0.99)	0.67 (0.49-0.92)	0.74 (0.57-0.96)	0.86 (0.67-1.11)	0.055
Multivariable 2	ref.	0.81 (0.67-0.98)	0.65 (0.47-0.91)	0.71 (0.53-0.96)	0.86 (0.63-1.16)	0.022
MACE: 3,338 events						
	n = 30,158	**n = 30,230**	**n = 29,946**	**n = 29,460**	**n = 28,091**	

	619 events	*668 events*	*669 events*	*701 events*	*681 events*	
Univariable	ref.	1.03 (0.97-1.10)	1.08 (0.96-1.20)	1.11 (1.01-1.22)	1.20 (1.09-1.32)	0.001*
Multivariable 1	ref.	0.97 (0.91-1.03)	0.94 (0.84-1.06)	0.92 (0.84-1.02)	0.88 (0.79-0.97)	0.091
Multivariable 2	ref.	0.96 (0.90-1.03)	0.92 (0.82-1.04)	0.90 (0.80-1.00)	0.85 (0.76-0.96)	0.136

§ = Quintile boundaries Q1/2 8.9 nmol/L (256 ng/dL), Q2/3 10.8 nmol/L (311 ng/dL), Q3/4 12.5 nmol/L (360 ng/dL) and Q4/5 14.8 nmol/L (427 ng/dL). 2.5th percentile = 5.9 nmol/L (170 ng/dL), 97.5th = 20.1 nmol/L (579 ng/dL); Presented numbers are for complete cases, after exclusions, and after excluding participants recording an event within the first two years or with less than two years of follow-up.

= Multivariable 1 models included age, ethnicity, living with partner, qualifications, alcohol consumption, smoking status, diet, physical activity, BMI, waist circumference, cholesterol, prevalent CVD, diabetes, hypertension, angina, atrial fibrillation, COPD, dementia, renal impairment, liver disease, thyroid disease, HIV, and use of lipid medications (a proxy for hyperlipidemia), glucocorticoids, opioids, and anticonvulsants, with the number of medications included as a proxy for overall comorbidity status and assessment centre as a stratification factor. Multivariable 2 models for testosterone included SHBG as an additional covariate.

* = Result interpreted as significant.

Table S6. Hazard ratios of different types of incident cardiovascular disease events by quintiles of testosterone (T): Analyses repeated for multivariable models dropping terms with sparse events among factor levels (diet, ethnicity, dementia, HIV).

Model	Q5 (highest T) (n = 30,467) [§]	Q4 (n = 30,515)	Q3 (n = 30,224)	Q2 (n = 29,787)	Q1 (lowest T) (n = 28,441)	Overall trend *P* Value
Myocardial infarction: 2,467 events						
	430 events	***502 events***	***512 events***	***513 events***	***510 events***	
Multivariable 1[#]	ref.	1.01 (0.94-1.09)	1.02 (0.89-1.16)	0.98 (0.87-1.10)	0.94 (0.83-1.05)	0.593
Multivariable 2	ref.	1.00 (0.92-1.08)	0.98 (0.85-1.13)	0.91 (0.80-1.04)	0.84 (0.73-0.97)	0.072
Haemorrhagic stroke: 307 events						
	66 events	***60 events***	***63 events***	***60 events***	***58 events***	
Multivariable 1	ref.	0.86 (0.70-1.05)	0.81 (0.56-1.17)	0.88 (0.64-1.20)	0.78 (0.56-1.08)	0.309
Multivariable 2	ref.	0.90 (0.72-1.11)	0.90 (0.61-1.33)	1.05 (0.73-1.51)	1.04 (0.70-1.54)	0.693
Ischaemic stroke: 1,313 events						
	257 events	***254 events***	***244 events***	***291 events***	***267 events***	
Multivariable 1	ref.	0.91 (0.82-1.01)	0.87 (0.72-1.04)	0.91 (0.78-1.07)	0.86 (0.74-1.01)	0.148
Multivariable 2	ref.	0.93 (0.83-1.04)	0.90 (0.74-1.09)	0.97 (0.81-1.16)	0.96 (0.79-1.16)	0.656
Heart failure: 649 events						
	113 events	***99 events***	***109 events***	***120 events***	***208 events***	
Multivariable 1	ref.	0.87 (0.74-1.01)	0.74 (0.56-0.98)	0.79 (0.63-1.00)	0.90 (0.72-1.13)	0.109
Multivariable 2	ref.	0.87 (0.73-1.02)	0.73 (0.55-0.99)	0.78 (0.60-1.02)	0.91 (0.69-1.19)	0.077
MACE: 4,266 events						
	787 events	***848 events***	***845 events***	***902 events***	***884 events***	
Multivariable 1	ref.	0.96 (0.91-1.01)	0.93 (0.84-1.03)	0.93 (0.85-1.02)	0.89 (0.81-0.97)	0.055
Multivariable 2	ref.	0.96 (0.91-1.02)	0.94 (0.84-1.04)	0.93 (0.84-1.03)	0.89 (0.80-0.99)	0.179

[§] = Quintile boundaries Q1/2 8.9 nmol/L (256 ng/dL), Q2/3 10.8 nmol/L (311 ng/dL), Q3/4 12.5 nmol/L (360 ng/dL) and Q4/5 14.8 nmol/L (427 ng/dL). 2.5th percentile = 5.9 nmol/L (170 ng/dL), 97.5th = 20.1 nmol/L (579 ng/dL); Presented numbers are for complete cases, after exclusions.

[#] = Multivariable 1 models included age, living with partner, qualifications, alcohol consumption, smoking status, physical activity, BMI, waist circumference, cholesterol, prevalent CVD, diabetes, hypertension, angina, atrial fibrillation, COPD, renal impairment, liver disease, thyroid disease, and use of lipid medications (a proxy for hyperlipidemia), glucocorticoids, opioids, and anticonvulsants, with the number of medications included as a proxy for overall comorbidity status and assessment centre as a stratification factor. Multivariable 2 models for testosterone included SHBG as an additional covariate.

* = Result interpreted as significant.

Table S7. P values from exploratory tests investigating the importance of including BMI as a term in analyses of the association of testosterone with incident cardiovascular disease (CVD). Results are arranged in columns, in increasing order of model complexity (from left to right)[a].

Analysis model: Test of term: CVD outcome variable	Univariable rcs(T)	Univar+BMI rcs(BMI)	Univar+BMI rcs(T)	Multivar2-BMI rcs(T)	Multivariable 2 rcs(BMI)	Multivariable 2 rcs(T)
Myocardial infarction	0.002*	< 0.001*	0.803	0.069	0.077	0.078
Haemorrhagic stroke	0.755	0.003*	0.609	0.706	0.085	0.689
Ischaemic stroke	0.038	< 0.001*	0.628	0.665	0.137	0.647
Heart Failure	< 0.001*	< 0.001*	< 0.001*	0.077	0.393	0.085
MACE	< 0.001*	< 0.001*	0.725	0.163	0.512	0.168

[a] = Refer to Supplementary Table S3 for model definitions.

* = Result interpreted as significant.

Table S8. P values from exploratory tests investigating the importance of including age as a term in analyses of the association of testosterone with incident cardiovascular disease (CVD). Results are arranged in columns, in increasing order of model complexity (from left to right)[a].

Analysis model: Test of term: CVD outcome variable	Univariable rcs(T)	Univar+Age rcs(age)	Univar+Age rcs(T)	Multivar2-Age rcs(T)	Multivariable 2 rcs(age)	Multivariable 2 rcs(T)
Myocardial infarction	0.002*	< 0.001*	0.147	0.723	< 0.001*	0.078
Haemorrhagic stroke	0.755	< 0.001*	0.600	0.563	0.004*	0.689
Ischaemic stroke	0.038	< 0.001*	0.255	0.035	< 0.001*	0.647
Heart Failure	< 0.001*	< 0.001*	< 0.001*	0.003*	< 0.001*	0.085
MACE	< 0.001*	< 0.001*	0.057	0.078	< 0.001*	0.168

[a] = Refer to Supplementary Table S3 for model definitions, substituting rcs(age) for rcs(BMI).

* = Result interpreted as significant.

Table S9. Hazard ratios of different types of incident cardiovascular disease events by quintiles of SHBG: Pooled estimates from analysis of multiply-imputed data.

Model	Q5 (highest) (n = 38,543) [§]	Q4 (n = 38,592)	Q3 (n = 38,531)	Q2 (n = 38,608)	Q1 (lowest) (n = 38,597)	Overall trend *P* Value
Myocardial infarction: 3,468 events						
	738 events	*664 events*	*714 events*	*671 events*	*681 events*	
Univariable	ref.	0.96 (0.90-1.02)	0.93 (0.84-1.04)	0.94 (0.86-1.03)	0.92 (0.84-1.01)	0.407
Multivariable 1[#]	ref.	0.98 (0.92-1.05)	0.98 (0.88-1.10)	1.03 (0.94-1.14)	1.07 (0.96-1.18)	0.531
Multivariable 2	ref.	1.00 (0.94-1.08)	1.02 (0.91-1.15)	1.10 (0.98-1.23)	1.17 (1.03-1.33)	0.086
Haemorrhagic stroke: 433 events						
	108 events	*98 events*	*86 events*	*77 events*	*64 events*	
Univariable	ref.	0.92 (0.77-1.09)	0.85 (0.63-1.14)	0.74 (0.57-0.96)	0.56 (0.43-0.75)	< 0.001*
Multivariable 1	ref.	0.96 (0.81-1.15)	0.90 (0.67-1.22)	0.83 (0.63-1.09)	0.66 (0.49-0.90)	0.089
Multivariable 2	ref.	0.99 (0.82-1.20)	0.95 (0.69-1.32)	0.88 (0.64-1.21)	0.73 (0.50-1.06)	0.443
Ischaemic stroke: 1,829 events						
	465 events	*405 events*	*369 events*	*324 events*	*266 events*	
Univariable	ref.	0.89 (0.82-0.97)	0.82 (0.71-0.95)	0.73 (0.64-0.83)	0.58 (0.50-0.66)	< 0.001*
Multivariable 1	ref.	0.97 (0.89-1.06)	0.96 (0.83-1.11)	0.91 (0.80-1.05)	0.78 (0.68-0.91)	0.006*
Multivariable 2	ref.	0.98 (0.89-1.07)	0.96 (0.82-1.13)	0.92 (0.79-1.08)	0.81 (0.68-0.97)	0.101
Heart failure: 1,019 events						
	226 events	*203 events*	*207 events*	*191 events*	*192 events*	
Univariable	ref.	0.97 (0.86-1.09)	0.97 (0.79-1.18)	0.87 (0.73-1.03)	0.87 (0.73-1.03)	0.073
Multivariable 1	ref.	0.97 (0.86-1.09)	0.92 (0.75-1.13)	0.84 (0.70-1.01)	0.83 (0.69-1.00)	0.237
Multivariable 2	ref.	1.00 (0.88-1.13)	0.97 (0.77-1.20)	0.86 (0.70-1.06)	0.79 (0.63-0.99)	0.232
MACE: 6,020 events						
	1,406 events	*1,221 events*	*1,222 events*	*1,121 events*	*1,050 events*	
Univariable	ref.	0.92 (0.87-0.96)	0.87 (0.80-0.94)	0.82 (0.77-0.88)	0.75 (0.70-0.81)	< 0.001*
Multivariable 1	ref.	0.96 (0.92-1.01)	0.94 (0.87-1.02)	0.94 (0.88-1.02)	0.92 (0.85-0.99)	0.192
Multivariable 2	ref.	0.98 (0.93-1.03)	0.97 (0.89-1.06)	0.99 (0.91-1.08)	0.99 (0.90-1.09)	0.928

§ = Quintile boundaries Q1/2 26.0 nmol/L, Q2/3 33.3 nmol/L, Q3/4 40.9 nmol/L and Q4/5 51.4 nmol/L. 2.5th percentile = 15.2 nmol/L, 97.5th = 79.1 nmol/L; Presented numbers are for imputed data, after exclusions.
# = Multivariable 1 models included age, ethnicity, living with partner, qualifications, alcohol consumption, smoking status, diet, physical activity, BMI, waist circumference, cholesterol, prevalent CVD, diabetes, hypertension, angina, atrial fibrillation, COPD, dementia, renal impairment, liver disease, thyroid disease, HIV, and use of lipid medications (a proxy for hyperlipidemia), glucocorticoids, opioids, and anticonvulsants, with the total number of medications used included as a proxy for overall comorbidity status and assessment centre as a stratification factor. Multivariable 2 models for SHBG included testosterone as an additional covariate.
* = Result interpreted as significant.

Table S10. Hazard ratios of different types of incident cardiovascular disease events by quintiles of SHBG: Analysis of data excluding participants who recorded an event within the first two years or had less than two years of follow-up.

Model	Q5 (highest)[§]	Q4	Q3	Q2	Q1 (lowest)	Overall trend *P* Value
Myocardial infarction: 1,904 events						
	n = 29,838	**n = 30,108**	**n = 29,880**	**n = 29,606**	**n = 28,676**	
	403 events	***357 events***	***397 events***	***382 events***	***365 events***	
Univariable	ref.	0.95 (0.87-1.04)	0.94 (0.81-1.09)	1.01 (0.89-1.14)	0.96 (0.84-1.08)	0.539
Multivariable 1[#]	ref.	0.98 (0.90-1.07)	0.99 (0.85-1.15)	1.10 (0.96-1.25)	1.10 (0.95-1.26)	0.296
Multivariable 2	ref.	1.01 (0.92-1.11)	1.05 (0.89-1.23)	1.20 (1.03-1.40)	1.26 (1.06-1.49)	0.030
Haemorrhagic stroke: 237 events						
	n = 29,936	**n = 30,209**	**n = 29,961**	**n = 29,691**	**n = 28,755**	
	63 events	***59 events***	***49 events***	***38 events***	***28 events***	
Univariable	ref.	0.91 (0.72-1.15)	0.83 (0.56-1.22)	0.67 (0.47-0.95)	0.42 (0.27-0.64)	< 0.001*
Multivariable 1	ref.	0.99 (0.78-1.25)	0.94 (0.63-1.41)	0.82 (0.56-1.19)	0.56 (0.35-0.88)	0.081
Multivariable 2	ref.	1.00 (0.77-1.29)	0.94 (0.61-1.46)	0.80 (0.52-1.23)	0.55 (0.32-0.94)	0.154
Ischaemic stroke: 1,050 events						
	n = 29,888	**n = 30,171**	**n = 29,929**	**n = 29,658**	**n = 28,727**	
	256 events	***245 events***	***216 events***	***191 events***	***142 events***	
Univariable	ref.	0.94 (0.84-1.05)	0.92 (0.76-1.11)	0.76 (0.64-0.90)	0.58 (0.49-0.70)	< 0.001*
Multivariable 1	ref.	1.03 (0.92-1.16)	1.08 (0.89-1.31)	0.97 (0.81-1.16)	0.81 (0.67-0.99)	0.024
Multivariable 2	ref.	1.05 (0.93-1.19)	1.12 (0.91-1.38)	1.01 (0.82-1.24)	0.86 (0.68-1.10)	0.110
Heart failure: 518 events						
	n = 29,930	**n = 30,190**	**n = 29,954**	**n = 29,678**	**n = 28,741**	
	111 events	***103 events***	***103 events***	***111 events***	***90 events***	
Univariable	ref.	0.95 (0.81-1.12)	0.98 (0.74-1.30)	1.04 (0.82-1.32)	0.89 (0.69-1.14)	0.392
Multivariable 1	ref.	0.97 (0.82-1.15)	0.99 (0.74-1.33)	1.08 (0.84-1.39)	0.95 (0.72-1.25)	0.703
Multivariable 2	ref.	1.03 (0.85-1.23)	1.08 (0.79-1.47)	1.15 (0.85-1.54)	0.94 (0.68-1.31)	0.358
MACE: 3,338 events						
	n = 29,774	**n = 30,064**	**n = 29,840**	**n = 29,566**	**n = 28,641**	

	764 events	*693 events*	*682 events*	*648 events*	*551 events*	
Univariable	ref.	0.93 (0.87-0.99)	0.89 (0.80-1.00)	0.88 (0.80-0.96)	0.76 (0.69-0.83)	< 0.001*
Multivariable 1	ref.	0.98 (0.92-1.05)	0.98 (0.87-1.09)	1.01 (0.91-1.12)	0.93 (0.84-1.04)	0.342
Multivariable 2	ref.	1.01 (0.94-1.08)	1.03 (0.91-1.16)	1.08 (0.97-1.22)	1.03 (0.91-1.18)	0.479

§ = Quintile boundaries Q1/2 26.0 nmol/L, Q2/3 33.3 nmol/L, Q3/4 40.9 nmol/L and Q4/5 51.4 nmol/L. 2.5th percentile = 15.2 nmol/L, 97.5th = 79.1 nmol/L; Presented numbers are for complete cases, after exclusions and after excluding participants recording an event within the first two years or with less than two years of follow-up.

= Multivariable 1 models included age, ethnicity, living with partner, qualifications, alcohol consumption, smoking status, diet, physical activity, BMI, waist circumference, cholesterol, prevalent CVD, diabetes, hypertension, angina, atrial fibrillation, COPD, dementia, renal impairment, liver disease, thyroid disease, HIV, and use of lipid medications (a proxy for hyperlipidemia), glucocorticoids, opioids, and anticonvulsants, with the total number of medications used included as a proxy for overall comorbidity status and assessment centre as a stratification factor. Multivariable 2 models for SHBG included testosterone as an additional covariate.

* = Result interpreted as significant.

Table S11. Hazard ratios of different types of incident cardiovascular disease events by quintiles of SHBG: Analyses repeated for multivariable models dropping terms with sparse events among factor levels (diet, ethnicity, dementia, HIV).

Model	Q5 (highest) (n = 30,207) [§]	Q4 (n = 30,380)	Q3 (n = 30,143)	Q2 (n = 29,826)	Q1 (lowest) (n = 28,878)	Overall trend *P* Value
Myocardial infarction: 2,467 events						
	535 events	***479 events***	***502 events***	***492 events***	***459 events***	
Multivariable 1[#]	ref.	0.96 (0.89-1.04)	0.96 (0.84-1.09)	1.05 (0.93-1.17)	1.08 (0.95-1.22)	0.316
Multivariable 2	ref.	0.98 (0.90-1.07)	1.00 (0.87-1.15)	1.13 (0.99-1.29)	1.21 (1.04-1.41)	0.032
Haemorrhagic stroke: 307 events						
	86 events	***69 events***	***63 events***	***49 events***	***40 events***	
Multivariable 1	ref.	0.91 (0.74-1.12)	0.83 (0.58-1.18)	0.76 (0.55-1.06)	0.55 (0.37-0.81)	0.029
Multivariable 2	ref.	0.92 (0.73-1.15)	0.83 (0.57-1.21)	0.75 (0.51-1.09)	0.53 (0.34-0.85)	0.083
Ischaemic stroke: 1,313 events						
	336 events	***294 events***	***265 events***	***237 events***	***181 events***	
Multivariable 1	ref.	0.99 (0.89-1.09)	1.01 (0.84-1.20)	0.95 (0.81-1.12)	0.79 (0.66-0.94)	0.010*
Multivariable 2	ref.	1.00 (0.90-1.12)	1.02 (0.85-1.23)	0.96 (0.80-1.16)	0.80 (0.65-1.00)	0.061
Heart failure: 649 events						
	142 events	***131 events***	***122 events***	***136 events***	***118 events***	
Multivariable 1	ref.	0.94 (0.81-1.10)	0.94 (0.72-1.22)	1.04 (0.83-1.30)	0.97 (0.76-1.23)	0.796
Multivariable 2	ref.	0.98 (0.83-1.15)	0.99 (0.75-1.31)	1.08 (0.83-1.40)	0.95 (0.71-1.27)	0.640
MACE: 4,266 events						
	1,005 events	***886 events***	***851 events***	***821 events***	***703 events***	
Multivariable 1	ref.	0.95 (0.90-1.01)	0.94 (0.85-1.03)	0.97 (0.89-1.06)	0.93 (0.84-1.02)	0.288
Multivariable 2	ref.	0.97 (0.91-1.03)	0.97 (0.87-1.08)	1.02 (0.92-1.13)	1.00 (0.89-1.12)	0.733

[§] = Quintile boundaries Q1/2 26.0 nmol/L, Q2/3 33.3 nmol/L, Q3/4 40.9 nmol/L and Q4/5 51.4 nmol/L. 2.5th percentile = 15.2 nmol/L, 97.5th = 79.1 nmol/L; Presented numbers are for complete cases, after exclusions.

[#] = Multivariable 1 models included age, living with partner, qualifications, alcohol consumption, smoking status, physical activity, BMI, waist circumference, cholesterol, prevalent CVD, diabetes, hypertension, angina, atrial fibrillation, COPD, renal impairment, liver disease, thyroid disease, and use of lipid medications (a proxy for hyperlipidemia), glucocorticoids, opioids, and anticonvulsants, with the total number of medications used included as a proxy for overall comorbidity status and assessment centre as a stratification factor. Multivariable 2 models for SHBG included testosterone as an additional covariate.

* = Result interpreted as significant.

Table S12. P values from exploratory tests investigating the importance of including BMI as a term in analyses of the association of SHBG with incident cardiovascular disease (CVD). Results are arranged in columns, in increasing order of model complexity (from left to right)[a].

Analysis model:	Univariable	Univar+BMI	Univar+BMI	Multivar2-BMI	Multivariable 2	Multivariable 2
Test of term:	rcs(SHBG)	rcs(BMI)	rcs(SHBG)	rcs(SHBG)	rcs(BMI)	rcs(SHBG)
CVD outcome variable						
Myocardial infarction	0.294	< 0.001*	< 0.001*	0.033	0.077	0.055
Haemorrhagic stroke	< 0.001*	< 0.001*	< 0.001*	0.063	0.085	0.079
Ischaemic stroke	< 0.001*	< 0.001*	< 0.001*	0.056	0.137	0.070
Heart Failure	0.267	< 0.001*	< 0.001*	0.617	0.393	0.614
MACE	< 0.001*	< 0.001*	< 0.001*	0.715	0.512	0.722

[a] = Refer to Supplementary Table S3 for model definitions, substituting rcs(SHBG) for rcs(T).

* = Result interpreted as significant.

Table S13. P values from exploratory tests investigating the importance of including age as a term in analyses of the association of SHBG with incident cardiovascular disease (CVD). Results are arranged in columns, in increasing order of model complexity (from left to right)[a].

Analysis model:	Univariable	Univar+Age	Univar+Age	Multivar2-Age	Multivariable 2	Multivariable 2
Test of term:	rcs(SHBG)	rcs(age)	rcs(SHBG)	rcs(SHBG)	rcs(age)	rcs(SHBG)
CVD outcome variable						
Myocardial infarction	0.294	< 0.001*	< 0.001*	0.004*	< 0.001*	0.055
Haemorrhagic stroke	< 0.001*	< 0.001*	0.088	< 0.001*	0.004*	0.079
Ischaemic stroke	< 0.001*	< 0.001*	0.084	< 0.001*	< 0.001*	0.070
Heart Failure	0.267	< 0.001*	< 0.001*	< 0.001*	< 0.001*	0.614
MACE	< 0.001*	< 0.001*	0.005*	< 0.001*	< 0.001*	0.722

[a] = Refer to Supplementary Table S3 for model definitions, substituting rcs(age) for rcs(BMI) and rcs(SHBG) for rcs(T).

* = Result interpreted as significant.

Table S14. P values from exploratory tests investigating the importance of including BMI as a term in analyses of the association of cFT with incident cardiovascular disease (CVD). Results are arranged in columns, in increasing order of model complexity (from left to right).[a]

Analysis model: Test of term: CVD outcome variable	Univariable rcs(cFT)	Univar+BMI rcs(BMI)	Univar+BMI rcs(cFT)	Multivar-BMI rcs(cFT)	Multivariable rcs(BMI)	Multivariable rcs(cFT)
Myocardial infarction	< 0.001*	< 0.001*	< 0.001*	0.065	0.067	0.081
Haemorrhagic stroke	0.007*	0.011	0.018	0.984	0.072	0.989
Ischaemic stroke	< 0.001*	< 0.001*	< 0.001*	0.744	0.101	0.729
Heart Failure	< 0.001*	< 0.001*	< 0.001*	0.379	0.402	0.428
MACE	< 0.001*	< 0.001*	< 0.001*	0.245	0.598	0.257

[a] = Refer to Supplementary Table S3 for model definitions, substituting cFT for T.

* = Result interpreted as significant.

Table S15. P values from exploratory tests investigating the importance of including age as a term in analyses of the association of cFT with incident CVD. Results are arranged in columns, in increasing order of model complexity (from left to right).[a]

Analysis model:	Univariable	Univar+Age	Univar+Age	Multivar-Age	Multivariable	Multivariable
Test of term:	rcs(cFT)	rcs(age)	rcs(cFT)	rcs(cFT)	rcs(age)	rcs(cFT)
CVD outcome variable						
Myocardial infarction	< 0.001*	< 0.001*	0.282	0.271	< 0.001*	0.081
Haemorrhagic stroke	0.007*	< 0.001*	0.746	0.480	< 0.001*	0.989
Ischaemic stroke	< 0.001*	< 0.001*	0.112	< 0.001*	< 0.001*	0.729
Heart Failure	< 0.001*	< 0.001*	< 0.001*	0.003*	< 0.001*	0.428
MACE	< 0.001*	< 0.001*	0.213	< 0.001*	< 0.001*	0.257

[a] = Refer to Supplementary Table S3 for model definitions, substituting cFT for T and rcs(age) for rcs(BMI).

* = Result interpreted as significant.

Table S16. P values from exploratory tests investigating the importance of including BMI as a term in analyses of the association of FTZ with incident cardiovascular disease (CVD). Results are arranged in columns, in increasing order of model complexity (from left to right).[a]

Analysis model: Test of term: CVD outcome variable	Univariable rcs(FTZ)	Univar+BMI rcs(BMI)	Univar+BMI rcs(FTZ)	Multivar-BMI rcs(FTZ)	Multivariable rcs(BMI)	Multivariable rcs(FTZ)
Myocardial infarction	< 0.001*	< 0.001*	< 0.001*	0.112	0.049	0.103
Haemorrhagic stroke	0.365	0.009*	0.613	0.847	0.069	0.843
Ischaemic stroke	< 0.001*	< 0.001*	< 0.001*	0.481	0.100	0.464
Heart Failure	< 0.001*	< 0.001*	< 0.001*	0.040	0.386	0.044
MACE	< 0.001*	< 0.001*	< 0.001*	0.115	0.550	0.108

[a] = Refer to Supplementary Table S3 for model definitions, substituting FTZ for T.

* = Result interpreted as significant.

Table S17. P values from exploratory tests investigating the importance of including age as a term in analyses of the association of FTZ with incident CVD. Results are arranged in columns, in increasing order of model complexity (from left to right).[a]

Analysis model:	Univariable	Univar+Age	Univar+Age	Multivar-Age	Multivariable	Multivariable
Test of term:	rcs(FTZ)	rcs(age)	rcs(FTZ)	rcs(FTZ)	rcs(age)	rcs(FTZ)
CVD outcome variable						
Myocardial infarction	< 0.001*	< 0.001*	0.500	0.202	< 0.001*	0.103
Haemorrhagic stroke	0.365	< 0.001*	0.822	0.954	< 0.001*	0.843
Ischaemic stroke	< 0.001*	< 0.001*	0.230	0.365	< 0.001*	0.464
Heart Failure	< 0.001*	< 0.001*	< 0.001*	0.009*	< 0.001*	0.044
MACE	< 0.001*	< 0.001*	0.265	0.459	< 0.001*	0.108

[a] = Refer to Supplementary Table S17 for model definitions, substituting FTZ for T and rcs(age) for rcs(BMI).

* = Result interpreted as significant.

Figure S1. Derivation of the study cohort.

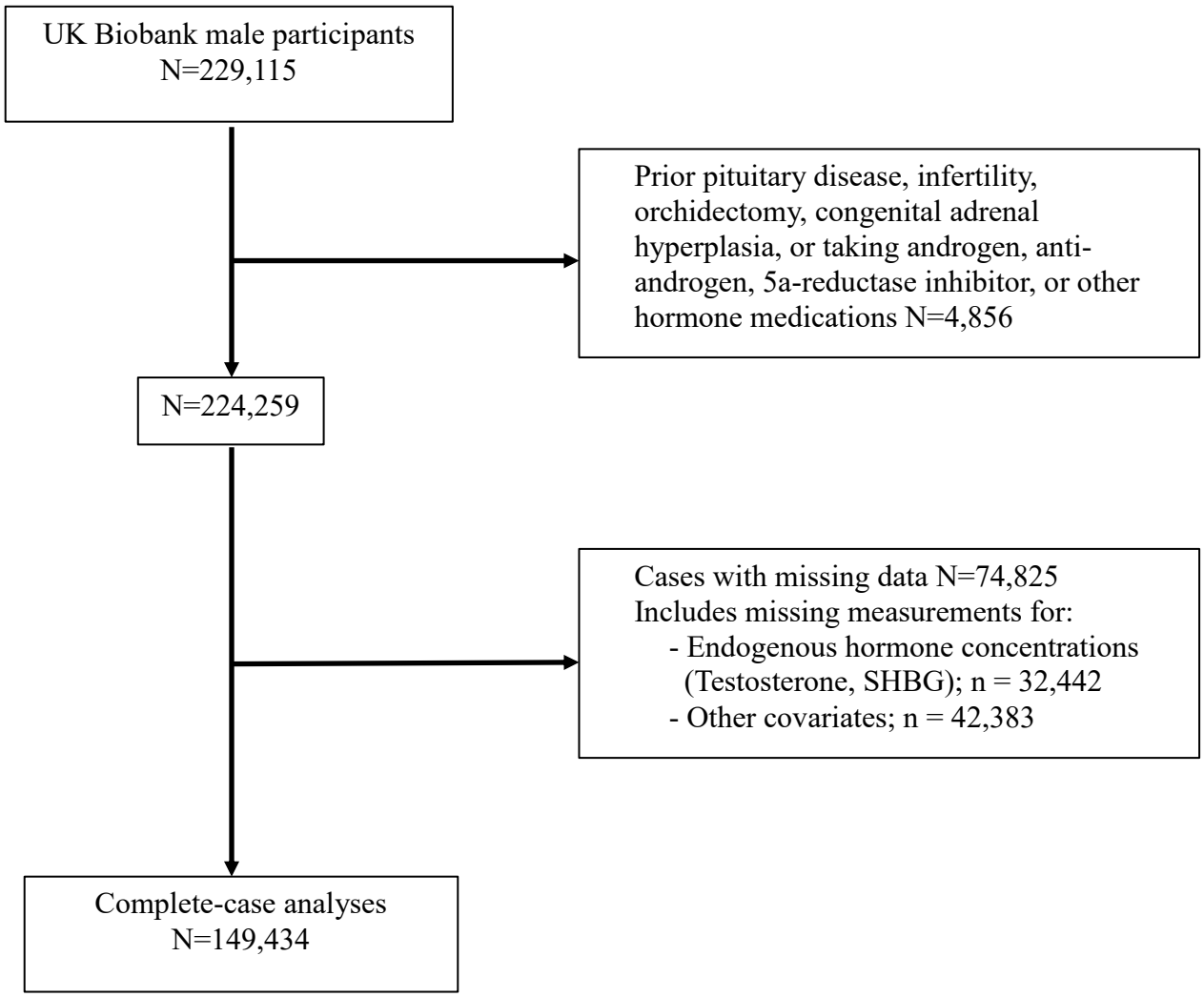

Figure S2. Kaplan-Meier survival plot according to quintiles of testosterone, showing average times to event for different types of cardiovascular disease.

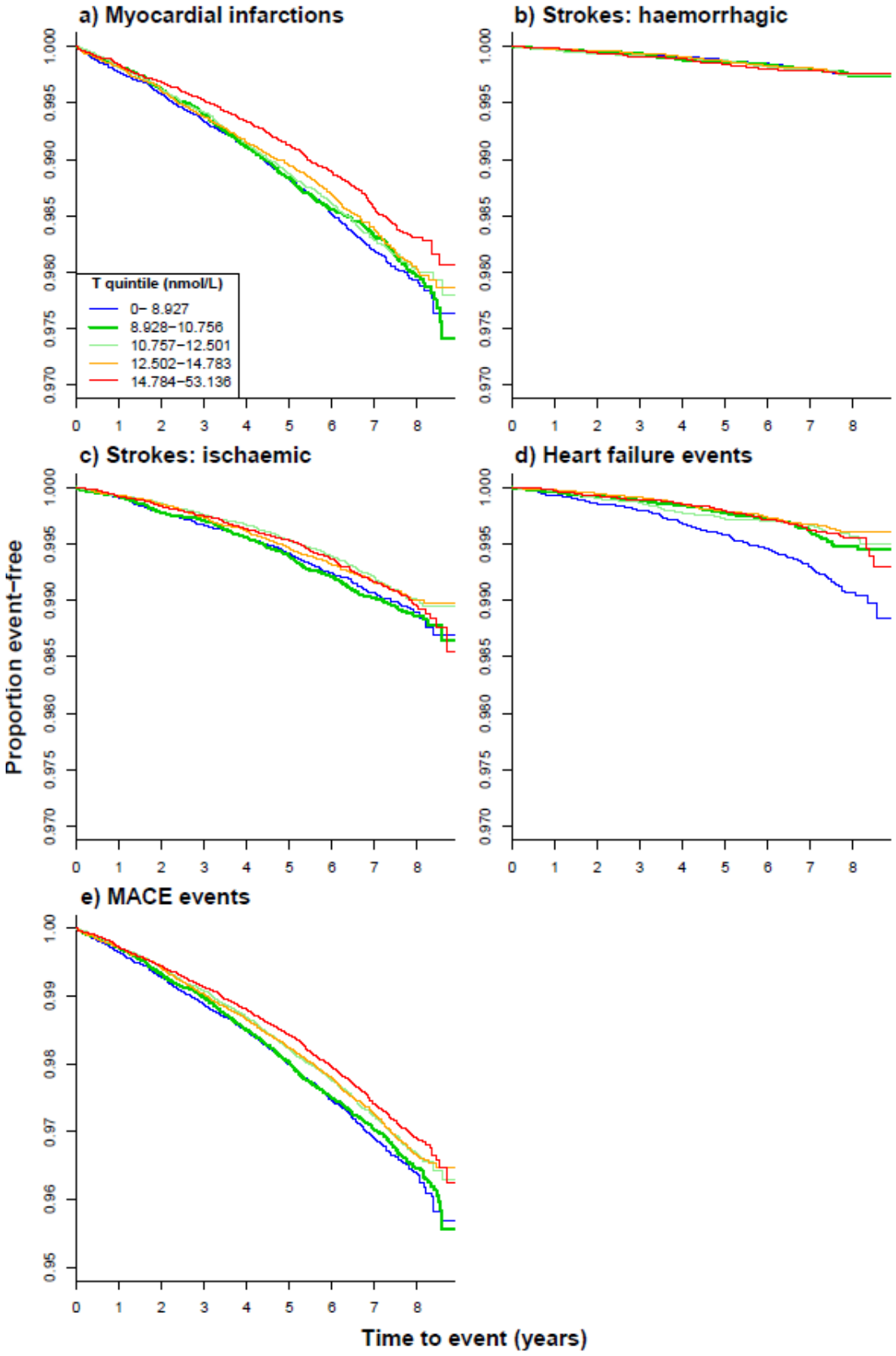

Figure S3. Univariable Cox proportional hazards regression model showing effect of baseline serum testosterone on risk of different types of cardiovascular disease. Shaded areas are the 95% confidence intervals. The horizontal dashed line is at the reference hazard (median of the fifth quintile). The vertical dashed lines are at medians for quintiles of testosterone.

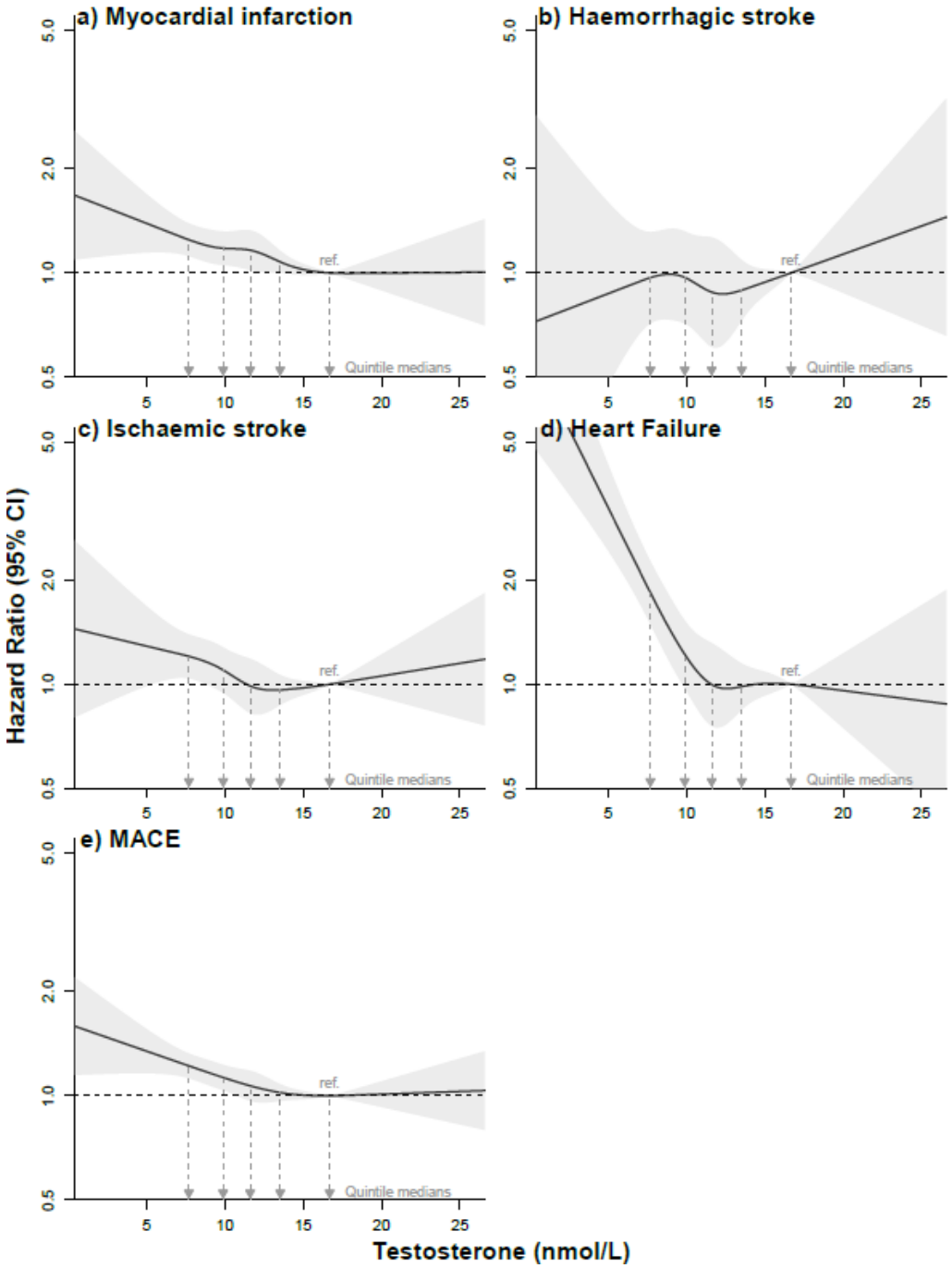

Figure S4. Multivariable Cox proportional hazards regression model showing effect of baseline serum testosterone on risk of different types of cardiovascular disease, adjusted for risk factors and potential confounders (not including SHBG). Shaded areas are the 95% confidence intervals. The horizontal dashed line is at the reference hazard (median of the fifth quintile). The vertical dashed lines are at medians for quintiles of testosterone.

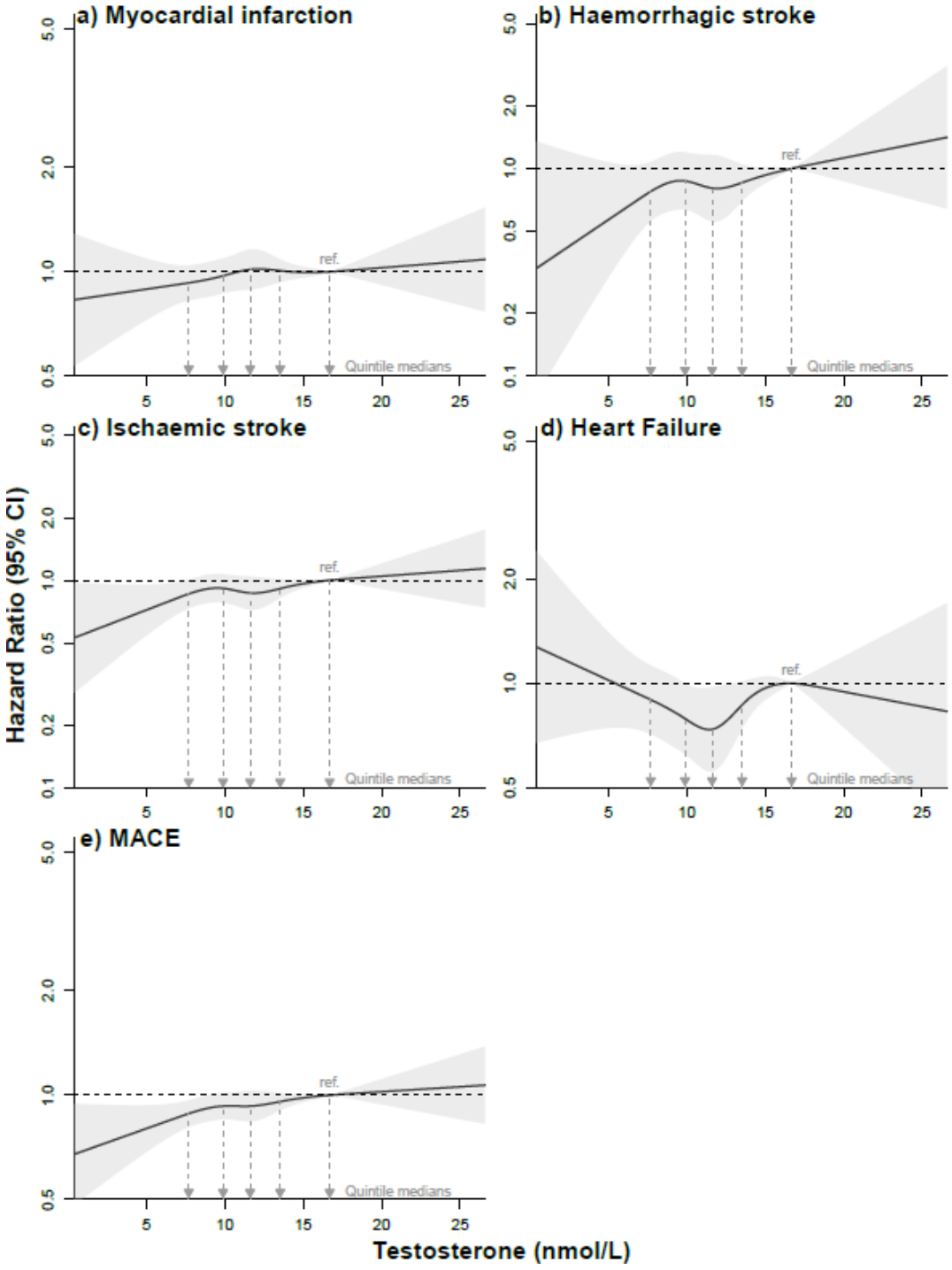

Figure S5. Kaplan-Meier survival plot according to quintiles of SHBG, showing average times to event for different types of cardiovascular disease.

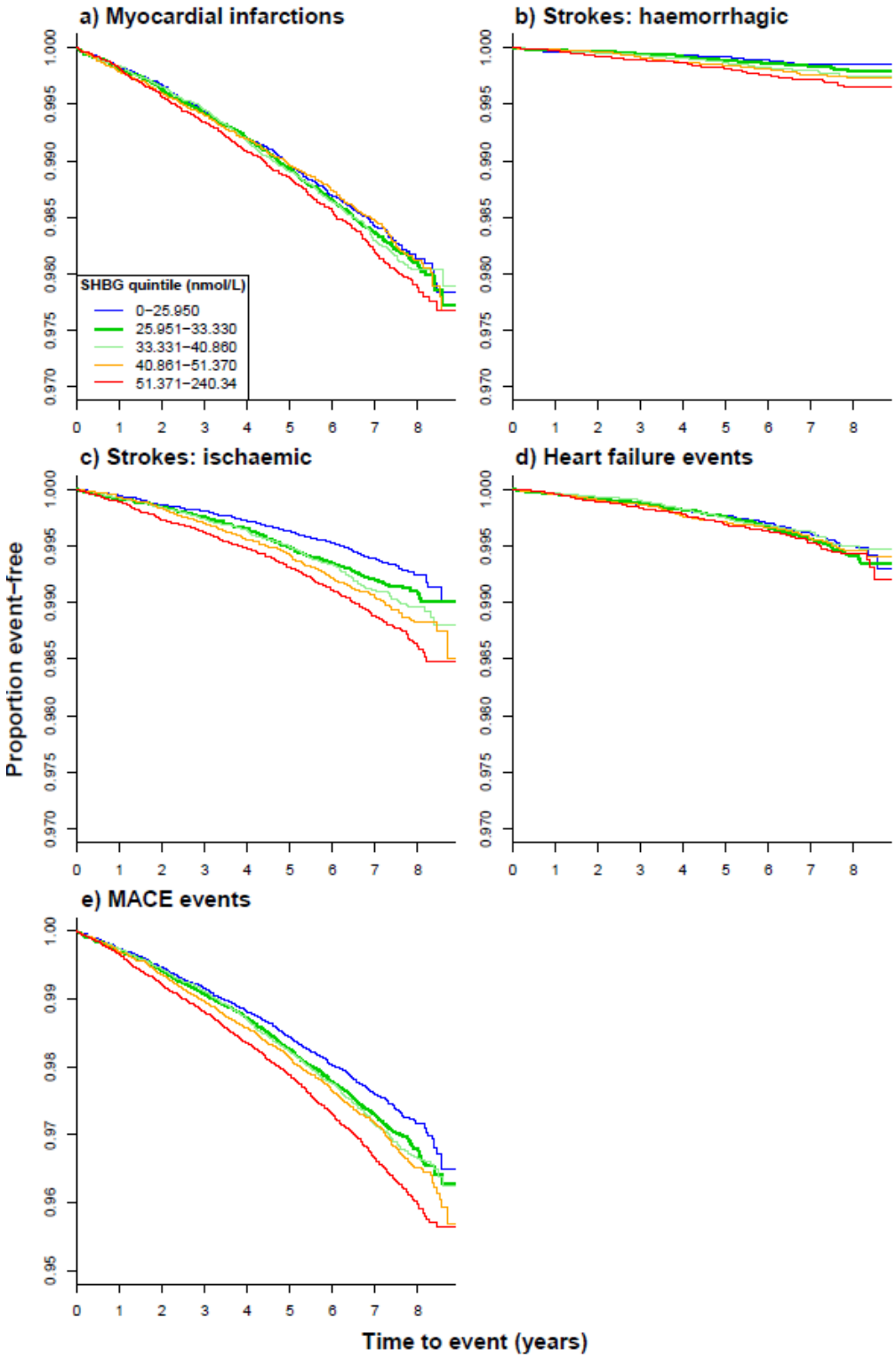

Figure S6. Univariable Cox proportional hazards regression model showing effect of baseline serum SHBG on risk of different types of cardiovascular disease. Shaded areas are the 95% confidence intervals. The horizontal dashed line is at the reference hazard (median of the fifth quintile). The vertical dashed lines are at medians for quintiles of SHBG.

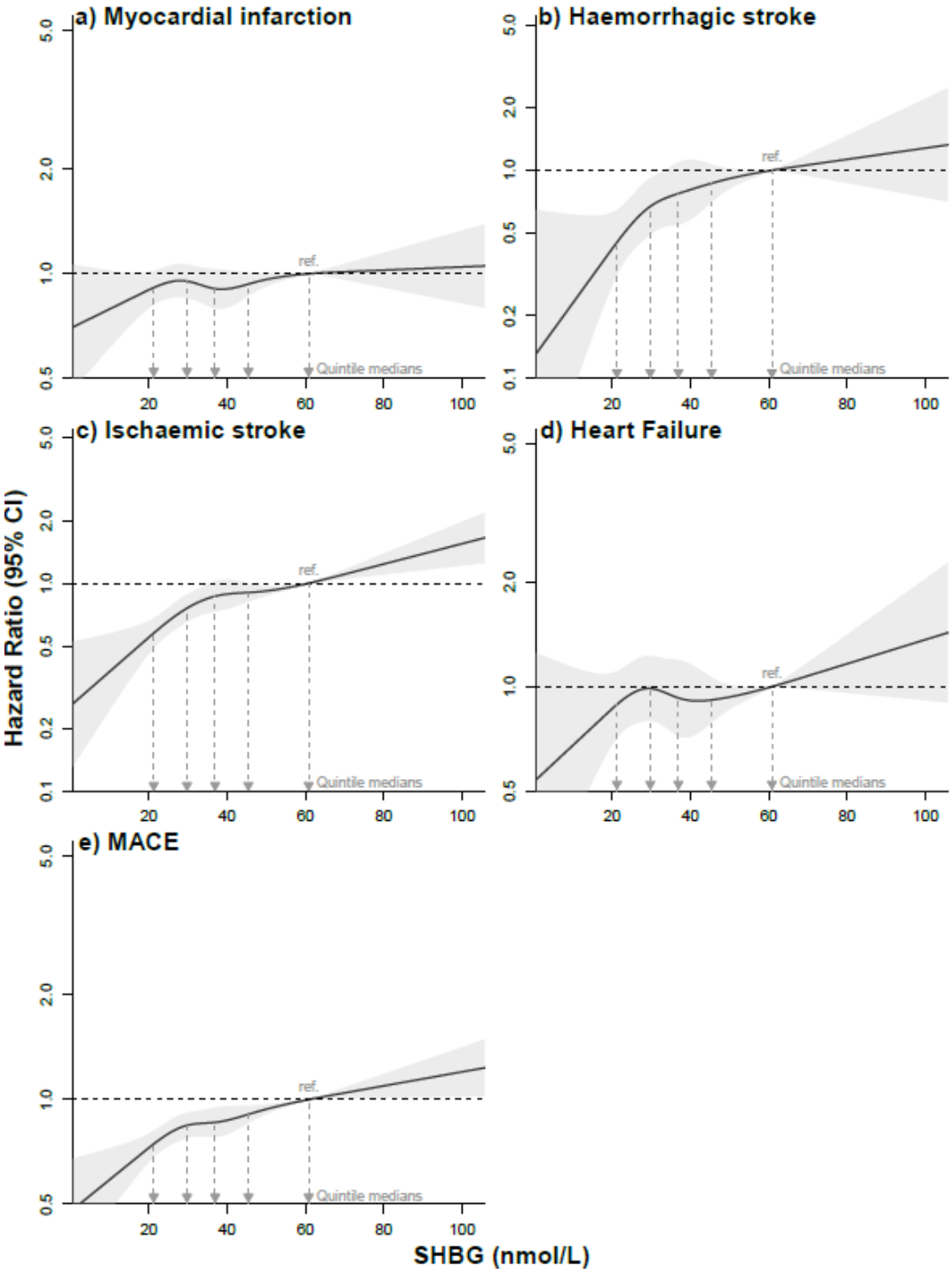

Figure S7. Multivariable Cox proportional hazards regression model showing effect of baseline serum SHBG on risk of different types of cardiovascular disease, adjusted for risk factors and potential confounders (not including testosterone). Shaded areas are the 95% confidence intervals. The horizontal dashed line is at the reference hazard (median of the fifth quintile). The vertical dashed lines are at medians for quintiles of SHBG.

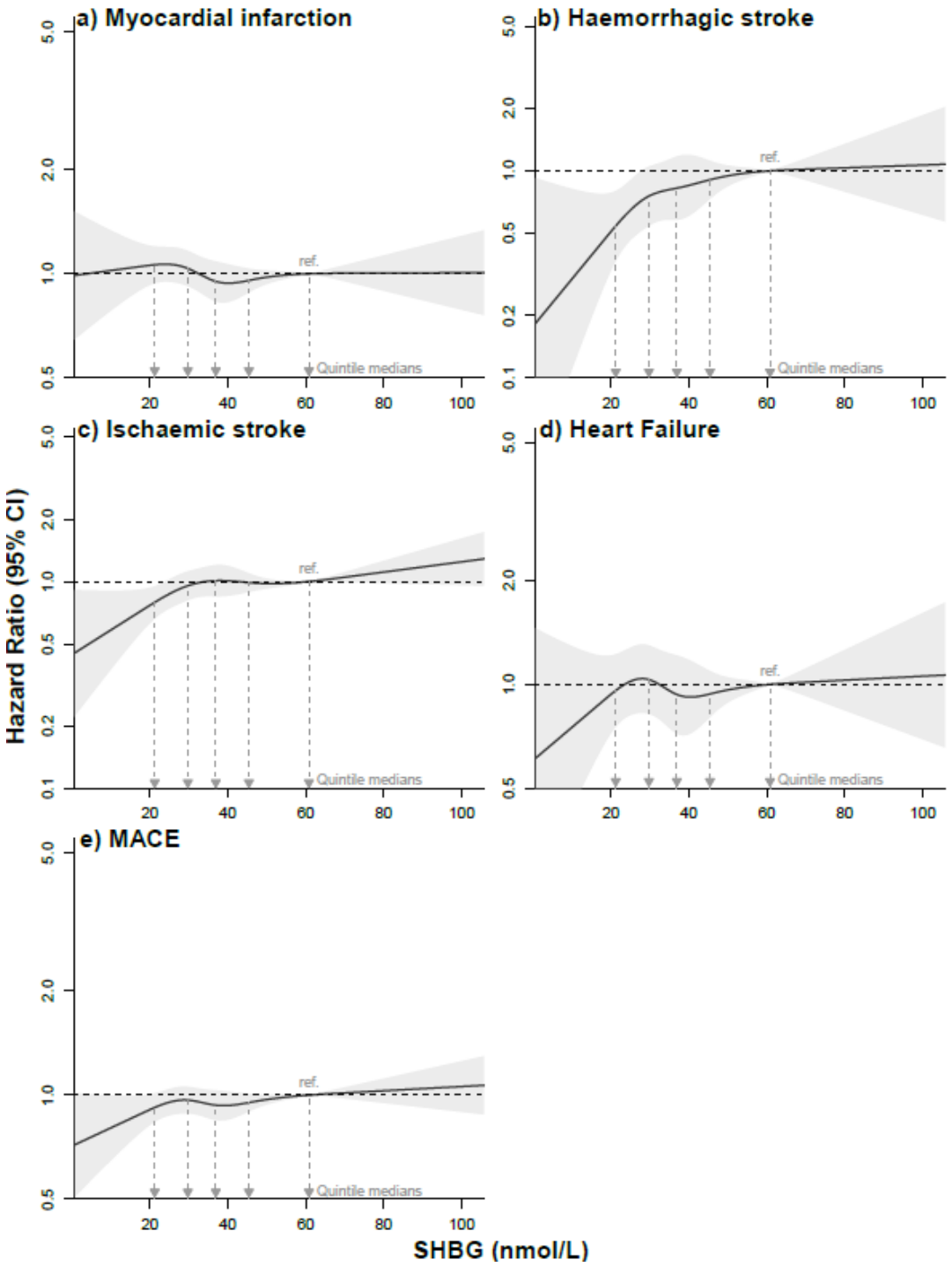

Figure S8. Kaplan-Meier survival plot according to quintiles of calculated free testosterone (cFT), showing average times to event for different types of cardiovascular disease.

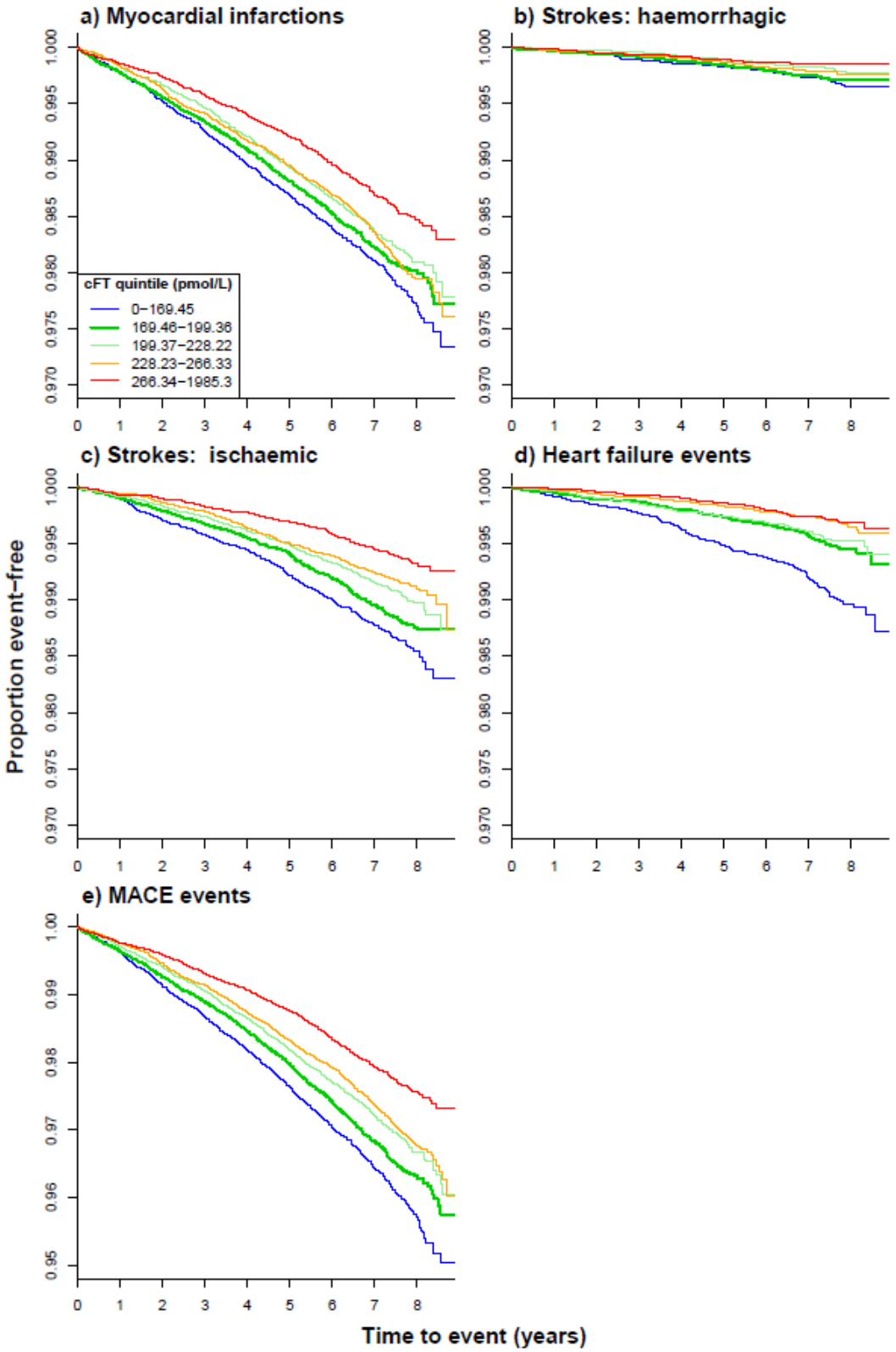

Figure S9. Univariable Cox proportional hazards regression model showing effect of baseline calculated free testosterone (cFT) value on risk of different types of cardiovascular disease. Shaded areas are the 95% confidence intervals. The horizontal dashed line is at the reference hazard (median of the fifth quintile). The vertical dashed lines are at medians for quintiles of cFT.

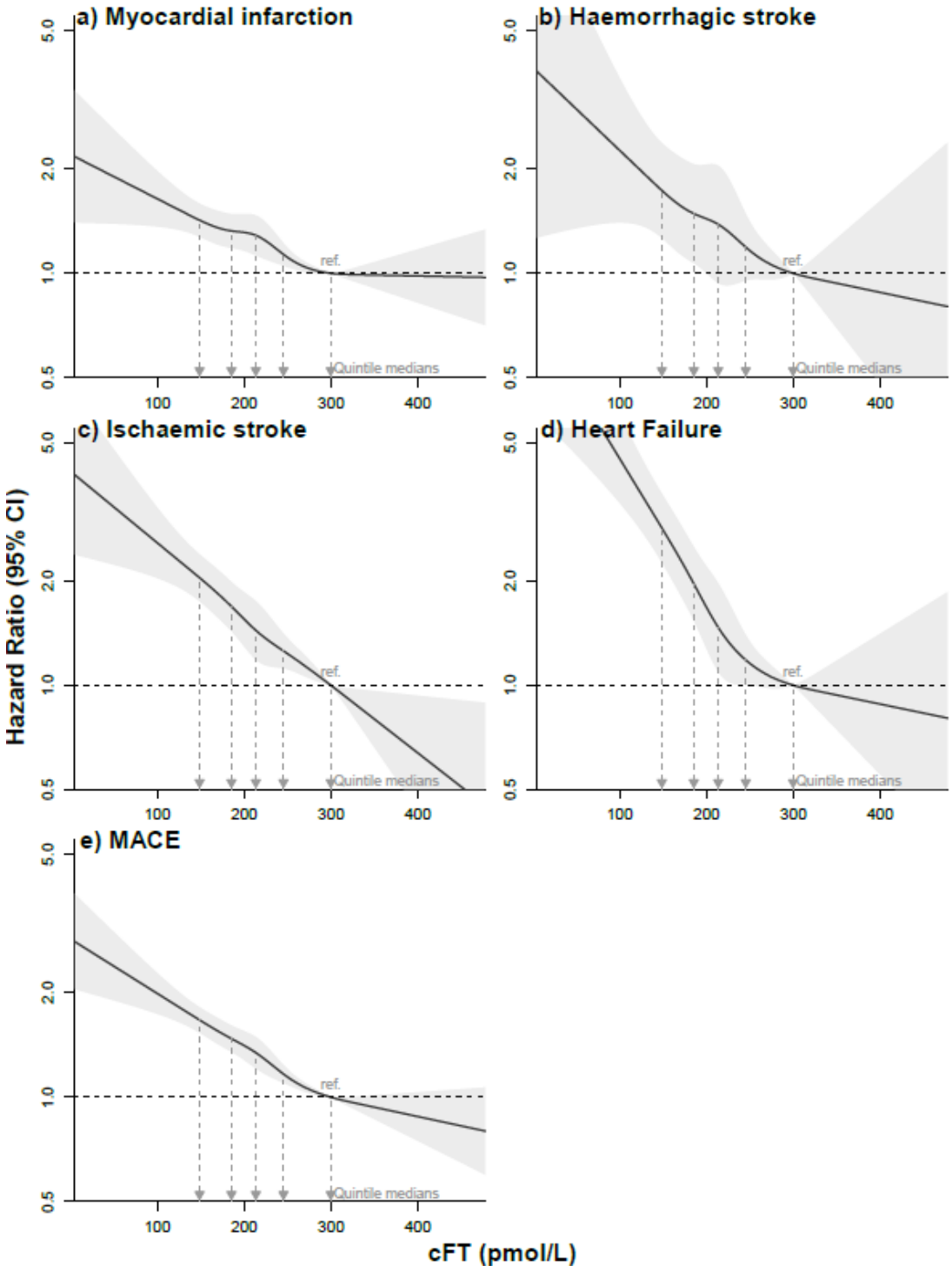

Figure S10. Multivariable model showing effect of baseline calculated free testosterone (cFT) value on risk of different types of cardiovascular disease. Shaded areas are the 95% confidence intervals. The horizontal dashed line is at the reference hazard (median of the fifth quintile). The vertical dashed lines are at medians for quintiles of cFT.

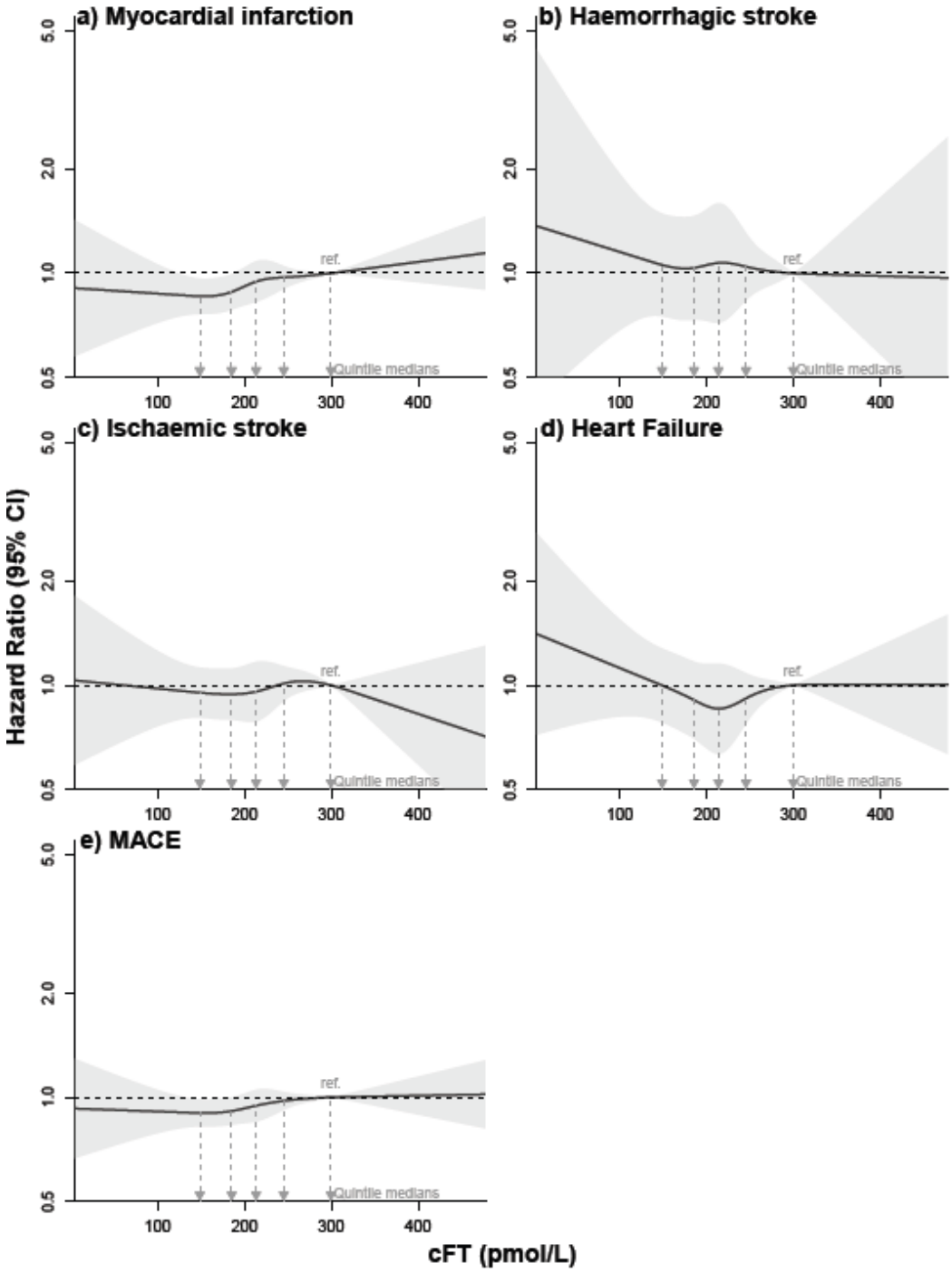

Figure S11.

Kaplan-Meier survival plot according to quintiles of calculated free testosterone (FTZ), showing average times to event for different types of cardiovascular disease.

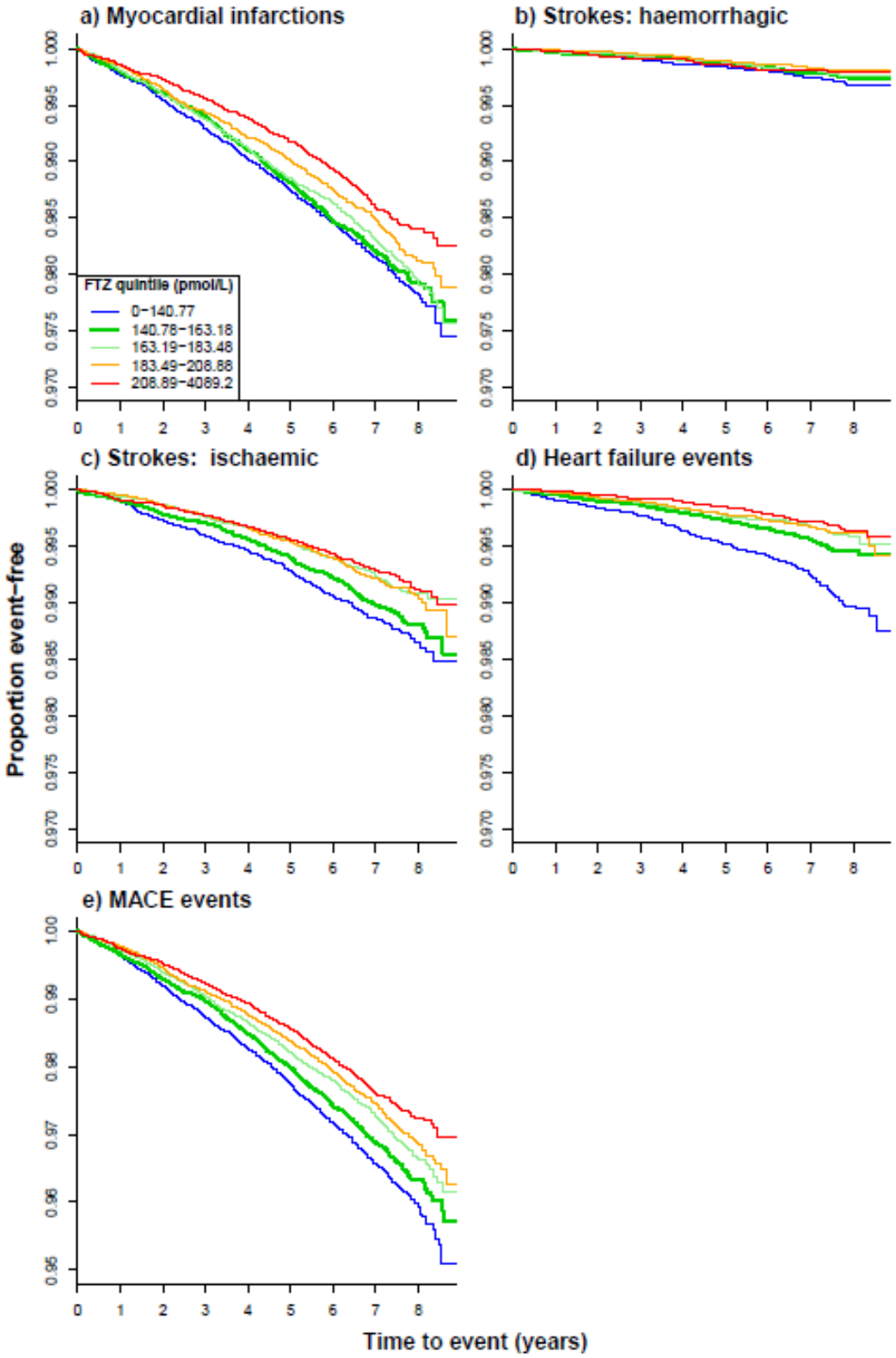

Figure S12. Univariable Cox proportional hazards regression model showing effect of baseline calculated free testosterone (FTZ) value on risk of different types of cardiovascular disease. Shaded areas are the 95% confidence intervals. The horizontal dashed line is at the reference hazard (median of the fifth quintile). The vertical dashed lines are at medians for quintiles of FTZ.

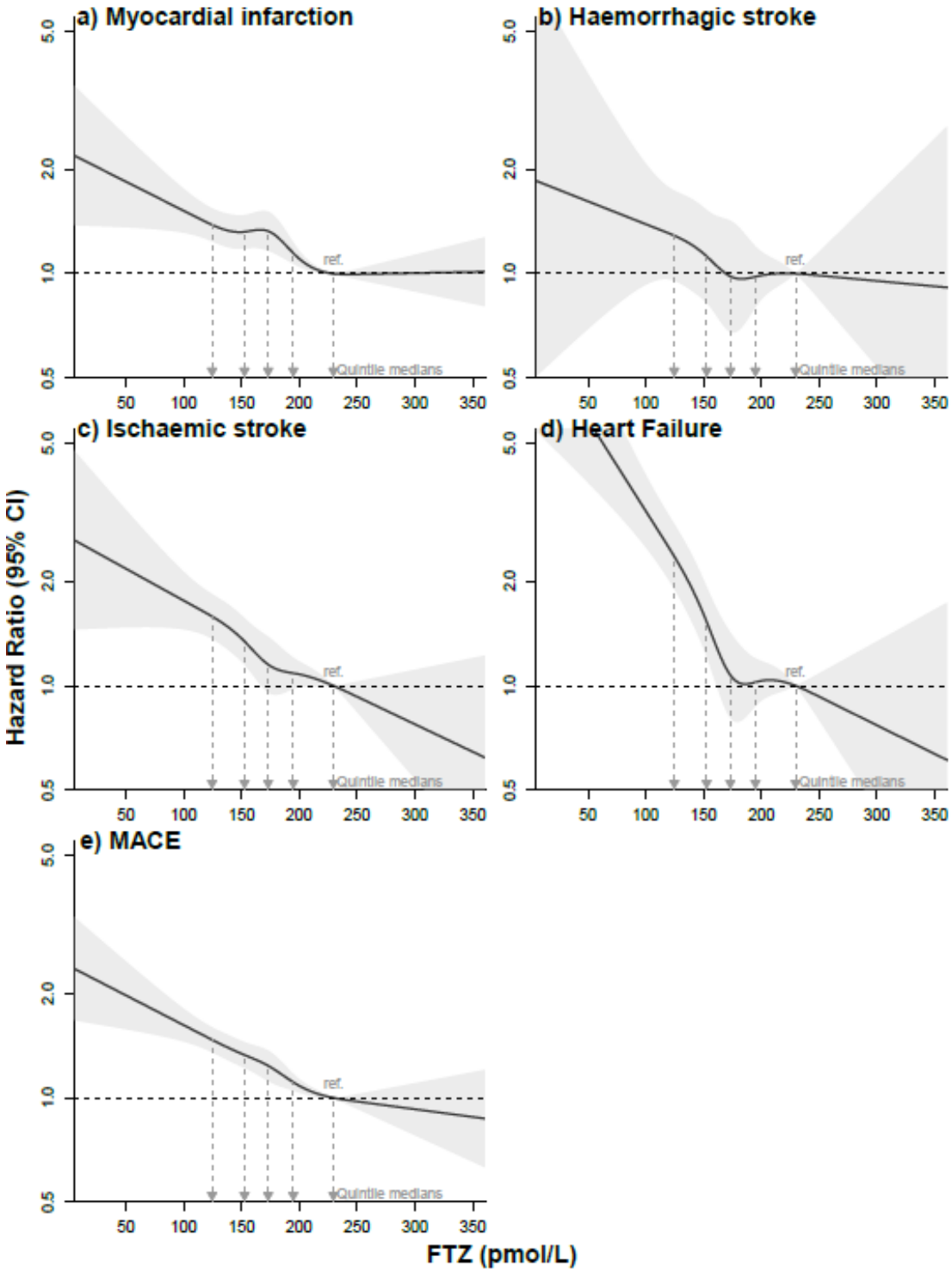

Figure S13. Multivariable model showing effect of baseline calculated free testosterone (FTZ) value on risk of different types of cardiovascular disease. Shaded areas are the 95% confidence intervals. The horizontal dashed line is at the reference hazard (median of the fifth quintile). The vertical dashed lines are at medians for quintiles of cFTZ.