Supplementary document

Estimating cholesterol variability using a linear mixed effects model

Given a dataset with N individuals and n_i cholesterol measurements from the i^{th} individual, $i = 1, \ldots, N$, let Y_{ij}, $j = 1, \ldots, n_i$,be the j^{th} measurement of individual i taken at measurement time t_{ij}.

Consider a standard linear mixed effects model

$$Y_{ij} = \beta^T X_{ij} + b_i^T Z_{ij} + \varepsilon_{ij}$$

where X_{ij} is a covariate vector for the fixed effects β and Z_{ij} is a covariate vector for the random effects b_i, assumed normally distributed $b_i \sim N(0, \Sigma_b)$. The residual errors ϵ_{ij} are assumed independent and normally distributed, $\epsilon_{ij} \sim N(0, \sigma^2)$. We can allow variability in the repeated measurements to differ between individuals by replacing the residual SD σ with an individual-specific residual SD σ_i and assuming that the σ_i are randomly distributed. We assume a log-normal distribution for the residual SD distribution, ensuring positivity of the SDs, $\sigma_i \sim \log N(\mu_\sigma, \tau_\sigma^2)$. The choice of log-normal distribution also allows a natural extension of the model to incorporate correlation between the usual level and the residual SD by assuming a multivariate normal distribution for the random effects and log residual SD

$$\varepsilon_{ij} \sim N(0, \sigma_i^2), \qquad \begin{pmatrix} b_i \\ \log \sigma_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ \mu_\sigma \end{pmatrix}, \begin{pmatrix} \Sigma_b & \Sigma_{b\sigma} \\ \Sigma_{b\sigma}^T & \tau_\sigma^2 \end{pmatrix}\right)$$

where $\Sigma_{b\sigma}$ is a vector of covariances between the random effects and the random residual errors. For this study, the model was

$$\begin{pmatrix} b_i \\ \log \sigma_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ \mu_\sigma \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \rho\tau_0\tau_\sigma \\ \rho\tau_0\tau_\sigma & \tau_\sigma^2 \end{pmatrix}\right)$$

For the Bayesian estimation, we used diffuse uniform prior distributions U[0, 100] for SDs, uniform U[−1, 1] prior distributions for correlation parameters, and diffuse normal prior distributions N(0, 100^2) for all other parameters. Priors were specified for the bivariate normal distribution by expressing it as two conditional univariate normal distributions. In the current study, we used a Markov Chain Monte Carlo (MCMC) of 2000 iterations (1000 burn-in) to update for the mixed effects models.

Supplementary table 1. Definition of the diseases

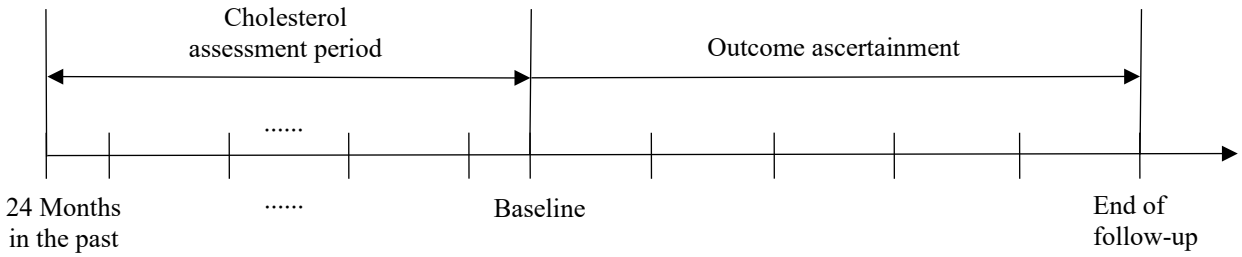
Event	ICPC-2 codes	ICD-9-CM codes
DM	T89, T90	NA
CHD	K74-K76	410-414
Heart Failure	K77	428
Stroke	K89-K91	430-438

ICPC-2 = the International Classification of Primary Care-2; ICD-9-CM = the International Classification of Diseases, Ninth Edition, Clinical Modification; DM = diabetes mellitus; CHD = coronary heart disease.

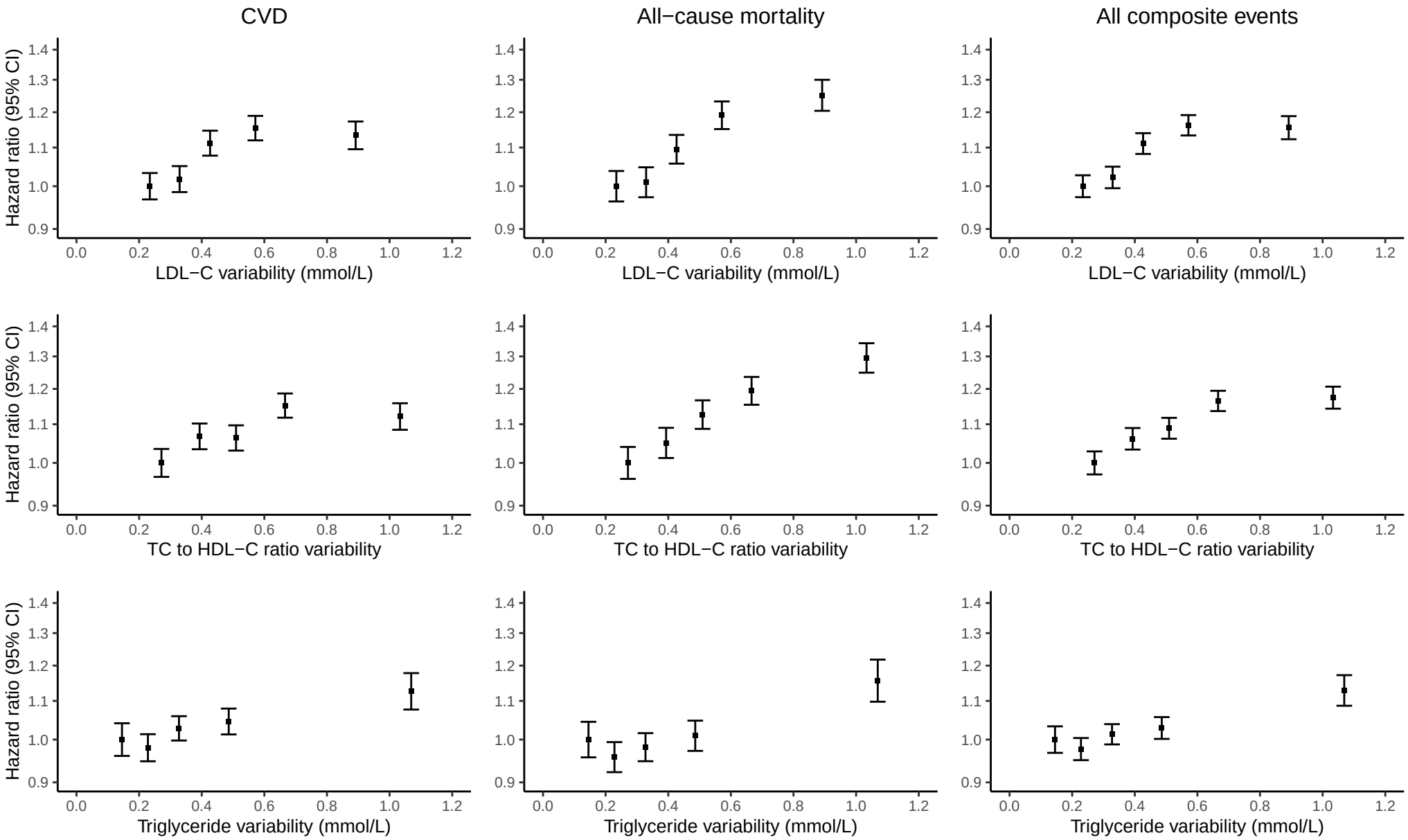
S.table 2. Data completion rate of the baseline characterstics in studied patients

	Total (N = 125,047)
Age	100.0% (125,047)
Gender	100.0% (125,047)
Duration of diabetic mellitus	95.9% (119,893)
Smoking status	99.4% (124,359)
Body mass index	94.8% (118,536)
Systolic blood pressure	99.4% (124,241)
Diastolic blood pressure	99.4% (124,242)
Haemoglobin A1c	99.8% (124,737)
Low-density lipoprotein-cholesterol	100.0% (125,047)
Total cholesterol to high-density lipoprotein-cholesterol ratio	100.0% (125,047)
Triglyceride	100.0% (125,047)
Estimated glomerular filtration rate	100.0% (125,035)
Charlson index	100.0% (125,047)
Use of anti-diabetic drugs	100.0% (125,047)
Use of anti-hypertensive drugs	100.0% (125,047)
Use of statins	100.0% (125,047)
Use of fibrates	100.0% (125,047)

S.figure 1. Study design for the investigation of the association between cholesterol variability and risk of cardiovascular diseases and all-cause mortality. The measurements of lipid between 2 years before baseline and baseline were used to calculate usual mean and variability of cholesterol. The median follow-up period was 77.5 months after baseline

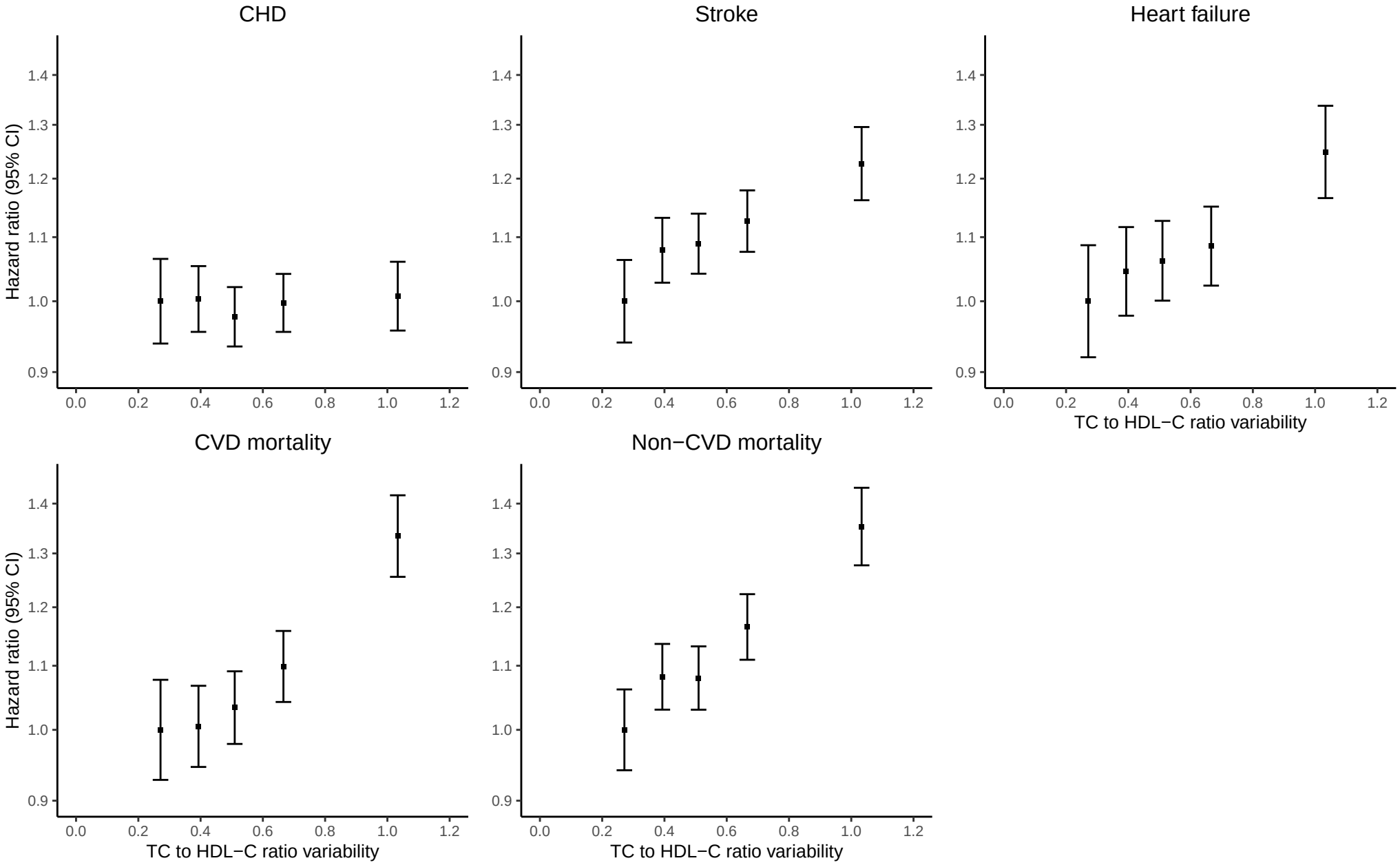


S.Figure 2a. Hazard ratios for the association of a unit increase in LDL−C, TC to HDL−C ratio and triglyceride variability, measured using the coefficient of variation, with CVD, all−cause mortality, and the composite outcome (CVD and mortality) from Cox regression models adjusted for baseline covariates



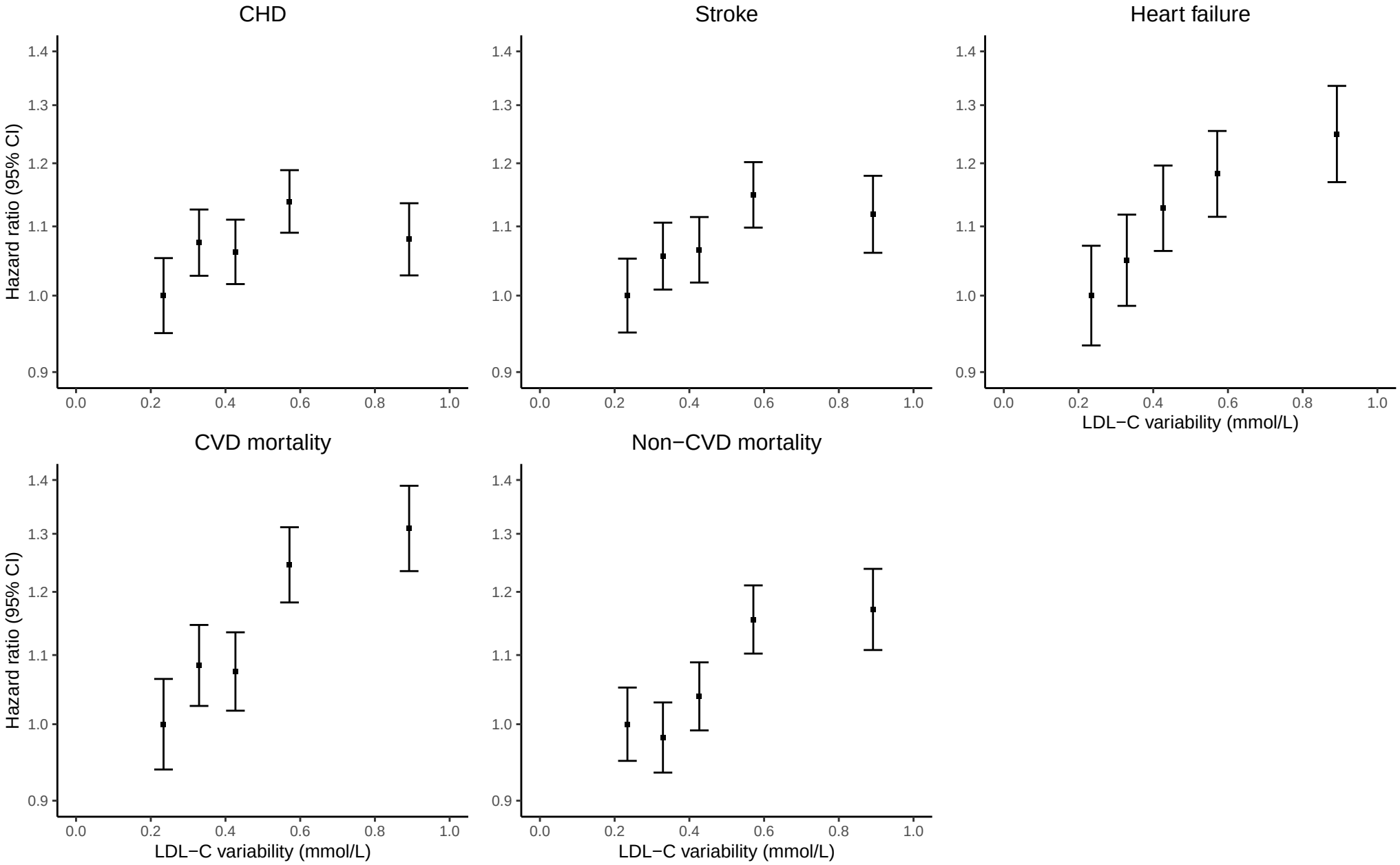
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti−diabetic drugs, anti−hypertensive drugs, statins and fibrates, Charlson's index and usual LDL−C, TC to HDL−C ratio or triglyceride (as appropriate). CIs are displayed as floating absolute risks.
LDL−C = Low−density lipoprotein−cholesterol; TC = Total cholesterol; HDL−C = High−density lipoproteincholesterol; CVD = Cardiovascular disease; CHD = Coronary heart disease.

S.Figure 3b. Hazard ratios for the association of a unit increase in TC to HDL−C ratio variability with CHD, stroke, heart failure, CVD mortality and non−CVD mortality from Cox regression models adjusted for baseline covariates



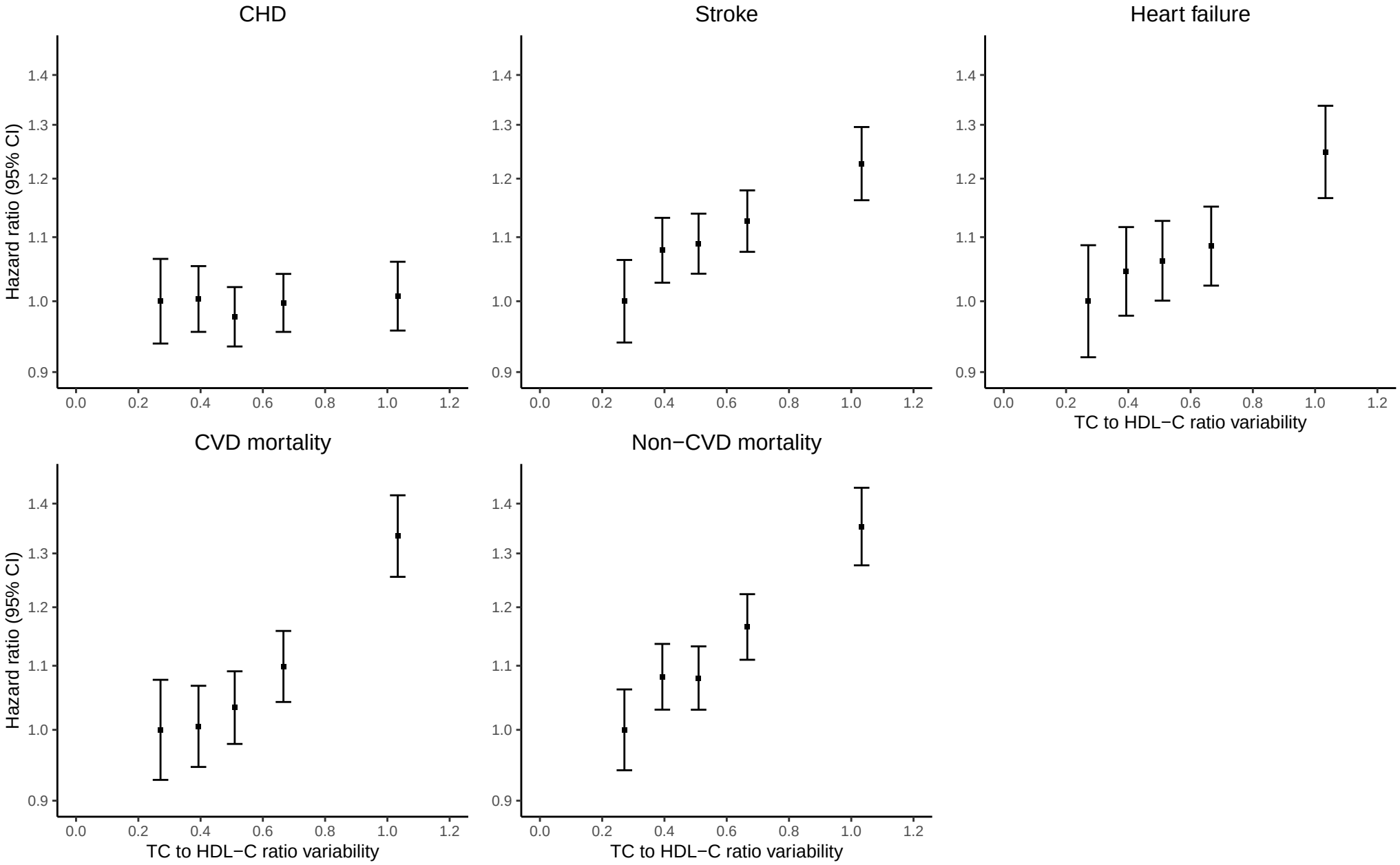
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti−diabetic drugs, anti−hypertensive drugs, statins and fibrates, Charlson's index and usual TC to HDL−C ratio. CIs are displayed as floating absolute risks. TC = Total cholesterol; HDL−C = High−density lipoprotein−cholesterol; CHD = Coronary heart disease.

S.Figure 3a. Hazard ratios for the association of a unit increase in LDL−C variability with CHD, stroke, heart failure, CVD mortality and non−CVD mortality from Cox regression models adjusted for baseline covariates



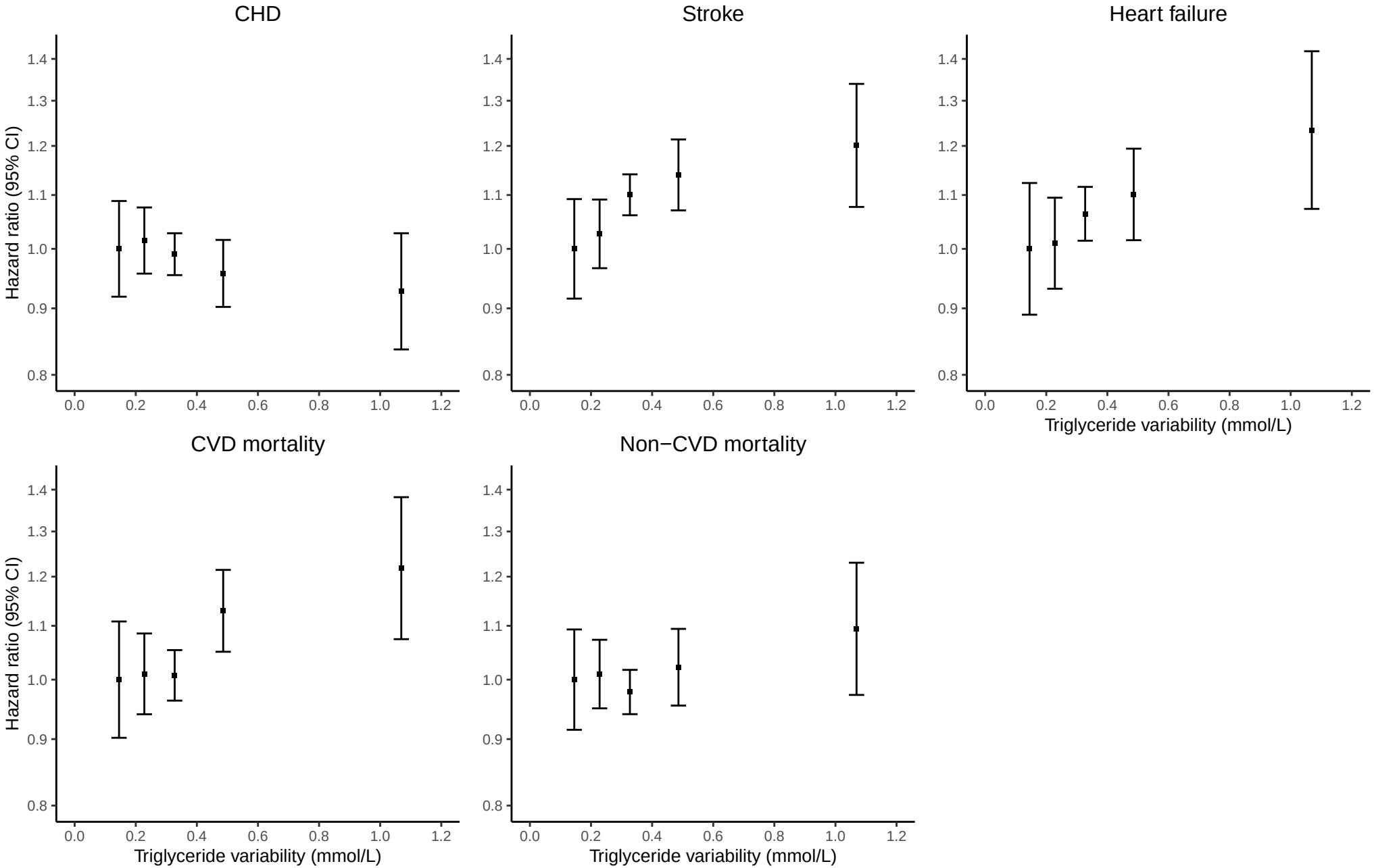
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti−diabetic drugs, anti−hypertensive drugs, statins and fibrates, Charlson's index and usual LDL−C. CIs are displayed as floating absolute risks. LDL−C = Low−density lipoprotein−cholesterol; CHD = Coronary heart disease.

S.Figure 3b. Hazard ratios for the association of a unit increase in TC to HDL−C ratio variability with CHD, stroke, heart failure, CVD mortality and non−CVD mortality from Cox regression models adjusted for baseline covariates



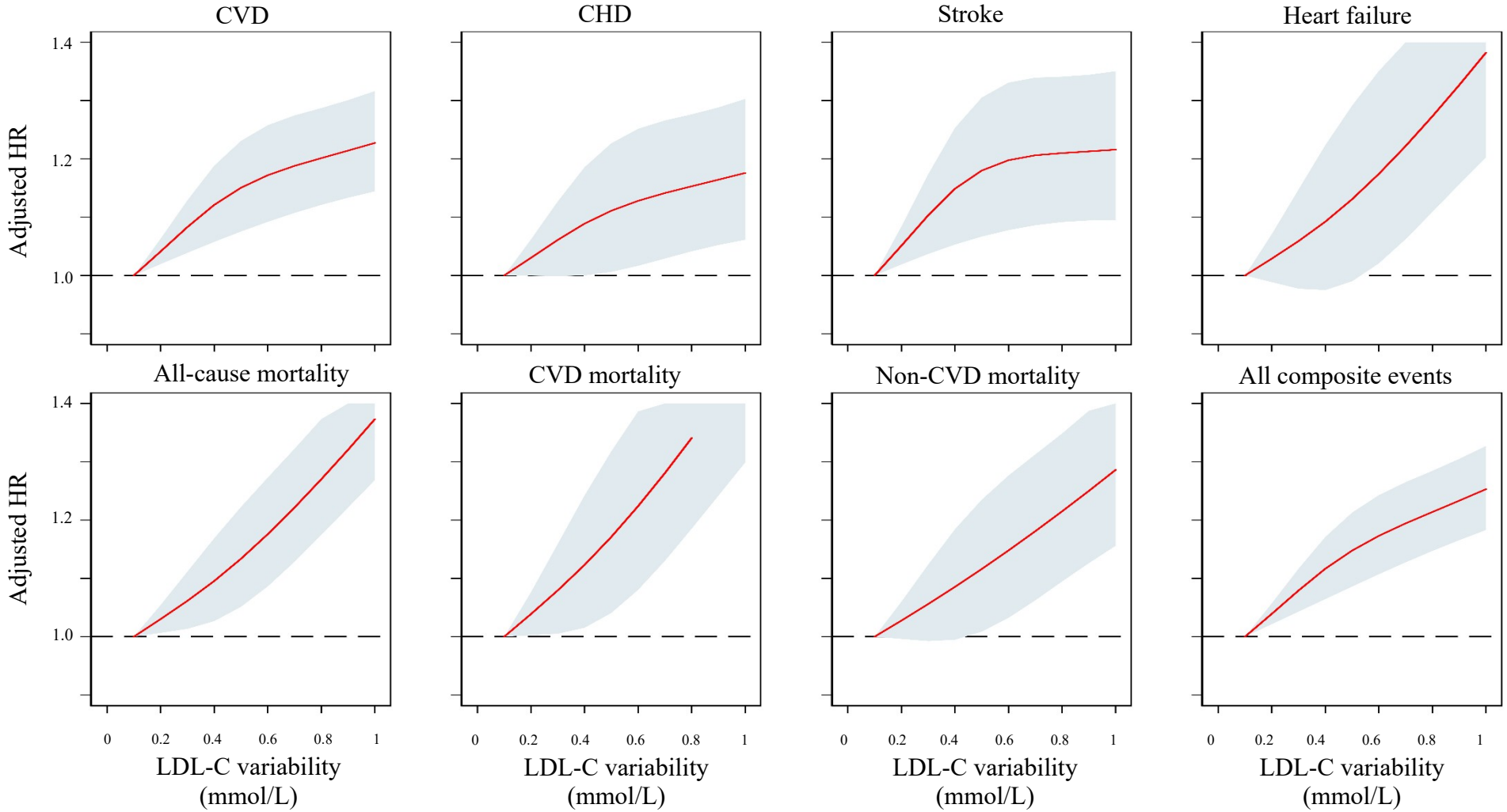
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti−diabetic drugs, anti−hypertensive drugs, statins and fibrates, Charlson's index and usual TC to HDL−C ratio. CIs are displayed as floating absolute risks. TC = Total cholesterol; HDL−C = High−density lipoprotein−cholesterol; CHD = Coronary heart disease.

S.Figure 3c. Hazard ratios for the association of a unit increase in triglyceride variability with CHD, stroke, heart failure, CVD mortality and non−CVD mortality from Cox regression models adjusted for baseline covariates



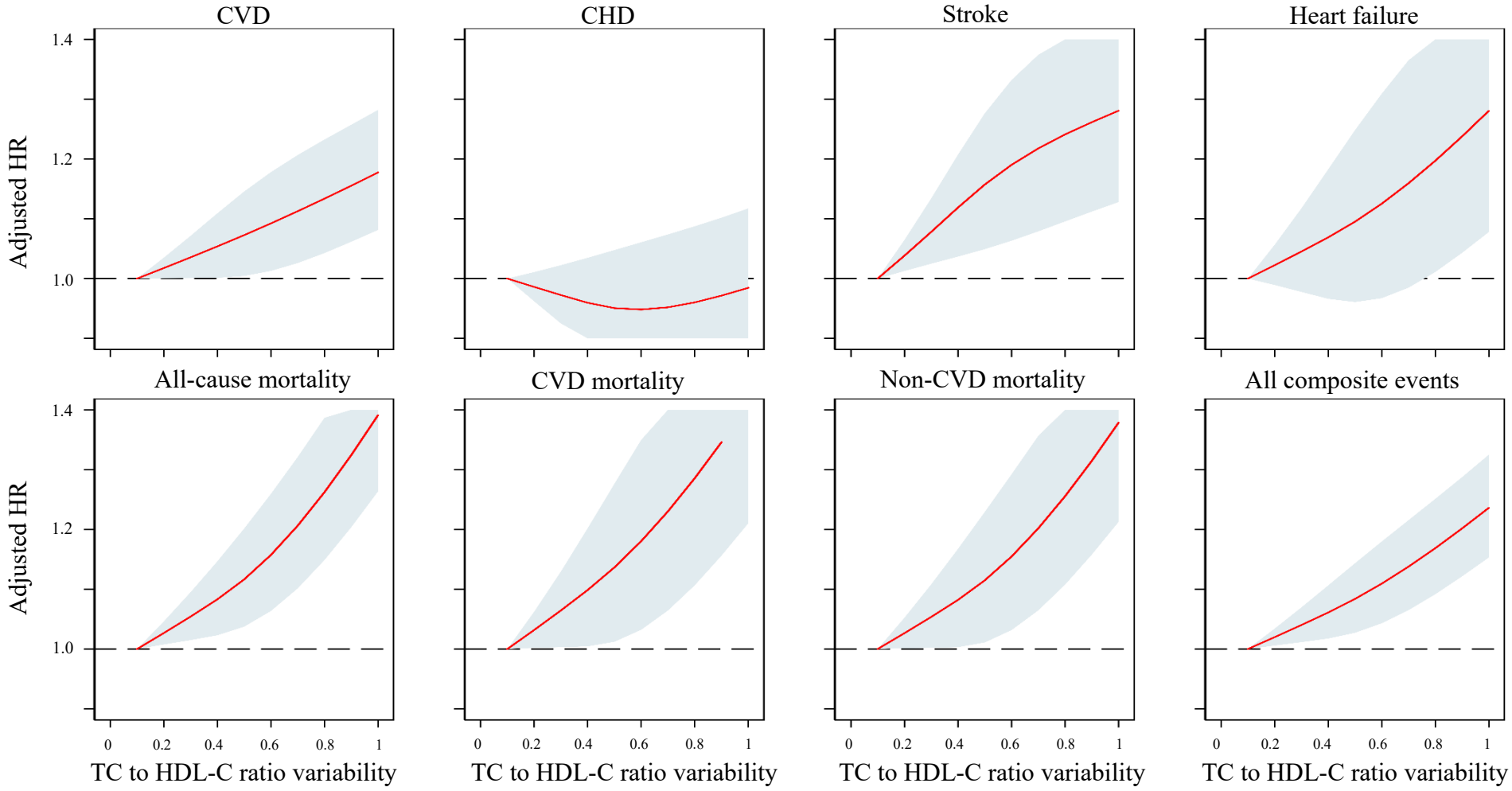
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti−diabetic drugs, anti−hypertensive drugs, statins and fibrates, Charlson's index and usual triglyceride. CIs are displayed as floating absolute risks. CHD = Coronary heart disease.

S.Figure 4a. Time-varying hazard ratios for the association of a unit increase in LDL-C variability with CVD, CHD, stroke, heart failure, all-cause mortality, CVD mortality, non-CVD mortality and the composite outcome (CVD and mortality), estimated using restricted cubic splines in Cox regression models adjusted for baseline covariates



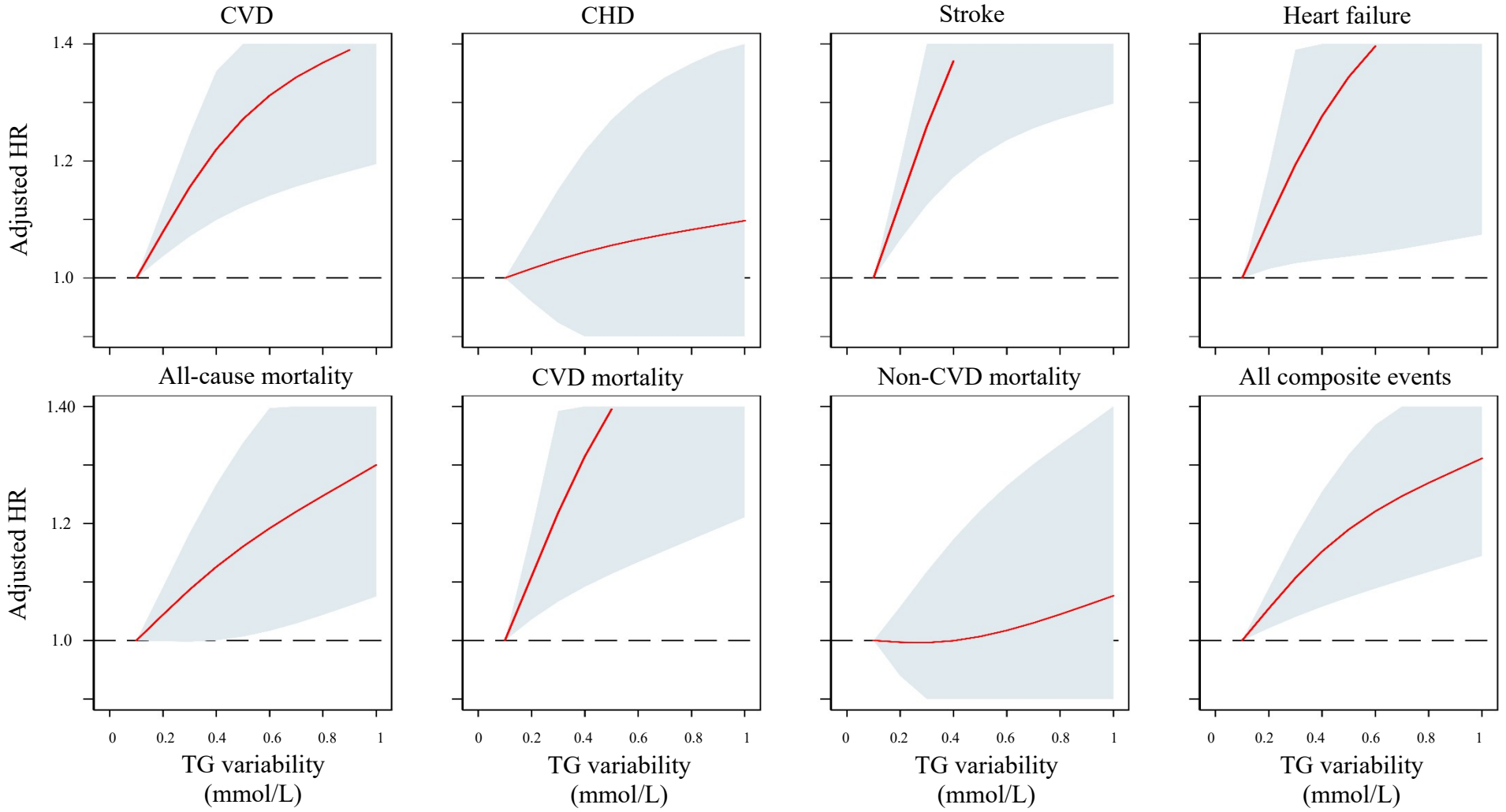
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti-diabetic drugs, anti-hypertensive drugs, statins and fibrates, Charlson's index and usual LDL-C. Shaded region represents 95% confidence intervals. CVD = Cardiovascular disease; CHD = Coronary heart disease; LDL-C = Low-density lipoprotein-cholesterol; HR = Hazard ratio.

S.Figure 4b. Time-varying hazard ratios for the association of a unit increase in TC to HDL-C ratio variability with CVD, CHD, stroke, heart failure, all-cause mortality, CVD mortality, non-CVD mortality and the composite outcome (CVD and mortality), estimated using restricted cubic splines in Cox regression models adjusted for baseline covariates



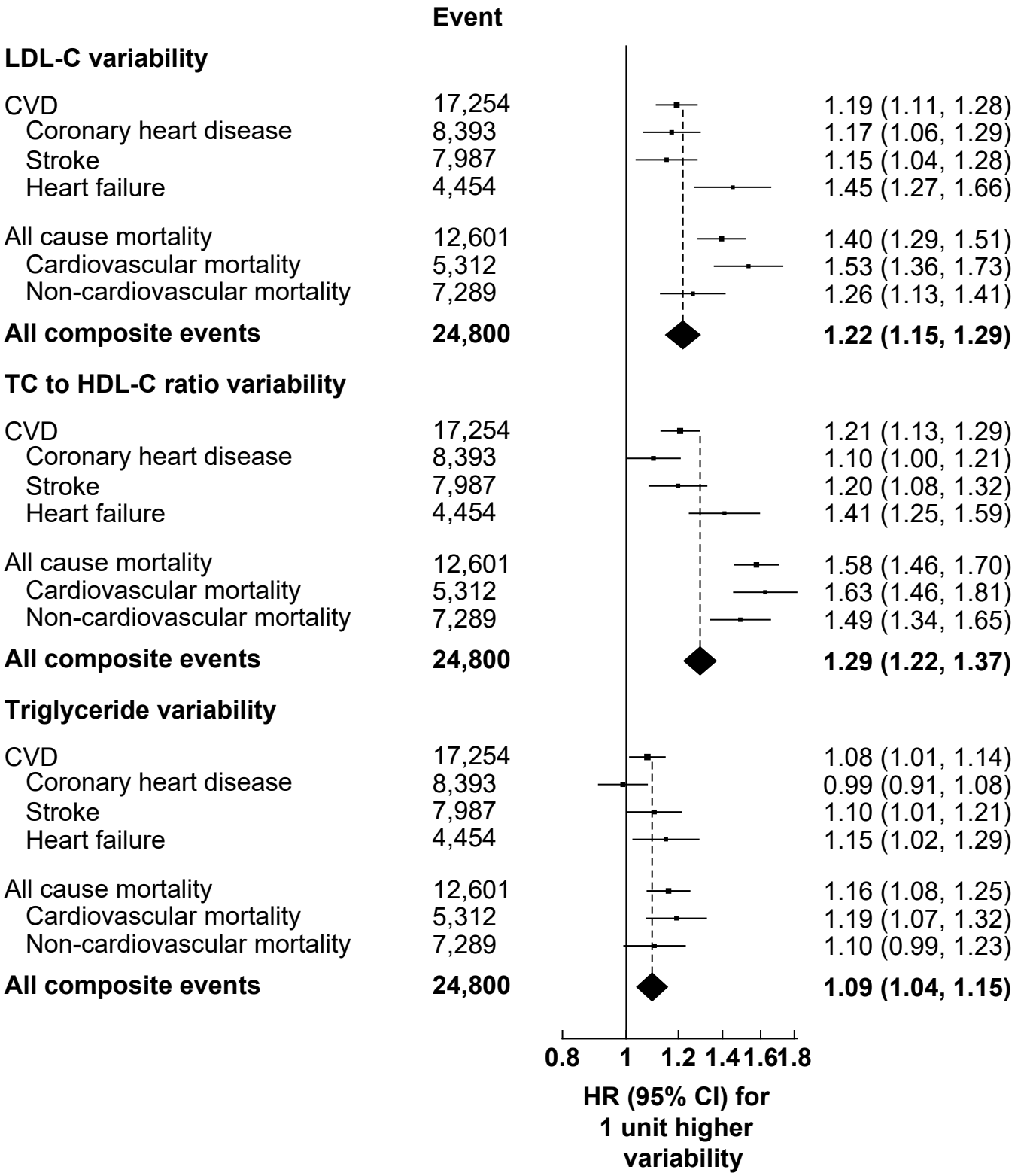
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti-diabetic drugs, anti-hypertensive drugs, statins and fibrates, Charlson's index and usual TC to HDL-C ratio. Shaded region represents 95% confidence intervals. CVD = Cardiovascular disease; CHD = Coronary heart disease; HDL-C = High-density lipoprotein-cholesterol; TC = Total cholesterol; HR = Hazard ratio.

S.Figure 4c. Time-varying hazard ratios for the association of a unit increase in triglyceride variability with CVD, CHD, stroke, heart failure, all-cause mortality, CVD mortality, non-CVD mortality and the composite outcome (CVD and mortality), estimated using restricted cubic splines in Cox regression models adjusted for baseline covariates



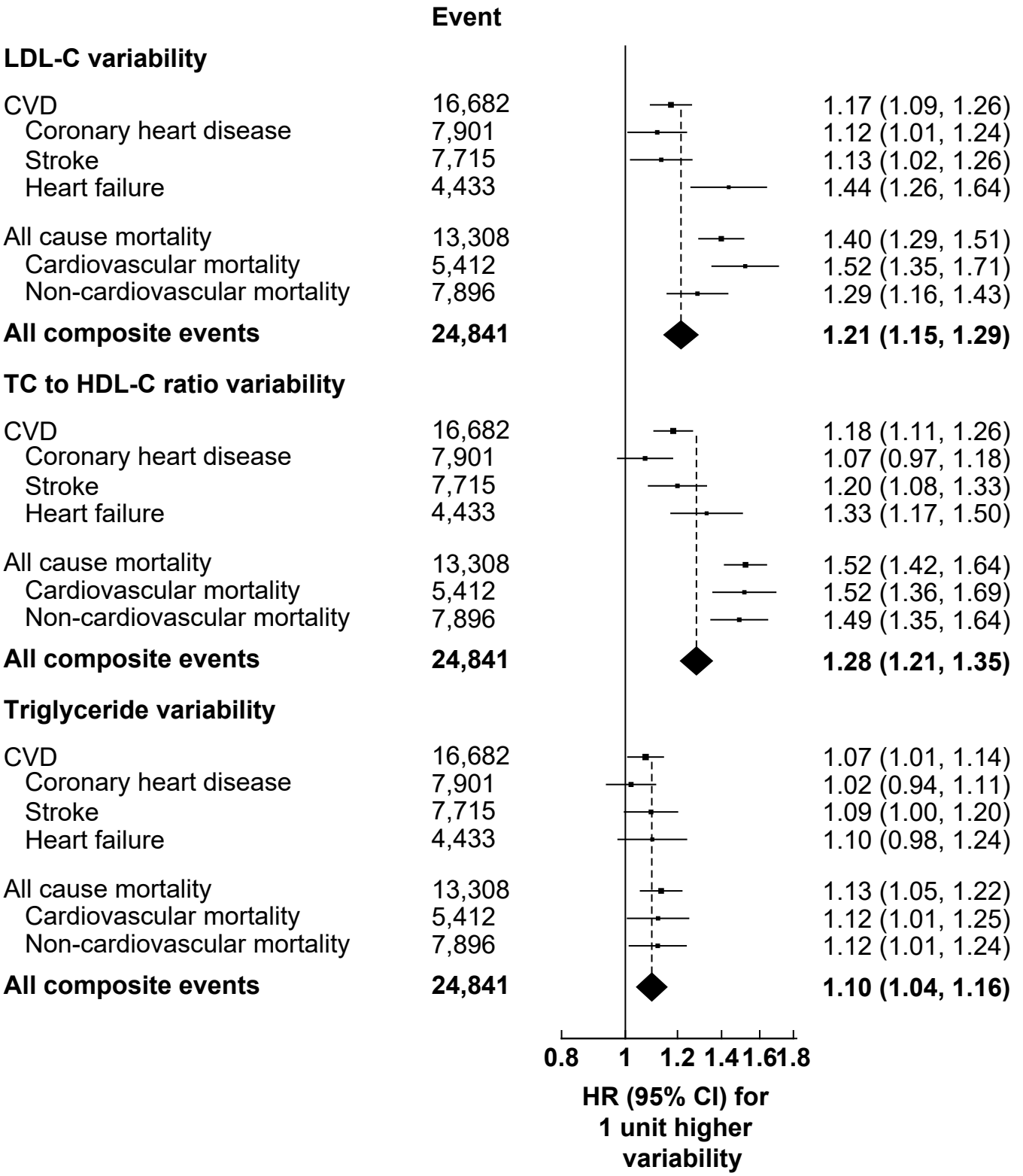
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti-diabetic drugs, anti-hypertensive drugs, statins and fibrates, Charlson's index and usual TG. Shaded region represents 95% confidence intervals. CVD = Cardiovascular disease; CHD = Coronary heart disease; TG = Triglyceride; HR = Hazard ratio.

S.Figure 5a: Hazard ratios for the risk of CVD, coronary heart disease, stroke, heart failure, all-cause mortality, CVD mortality, non-CVD mortality and their composite with each 1 unit increase in LDL-C (mmol/L), TC to HDL-C ratio or triglyceride (mmol/L) variability using Cox regressions adjusted for baseline covariates in sensitivity analysis 1 (complete case analysis, i.e. exclude records with any missing value)



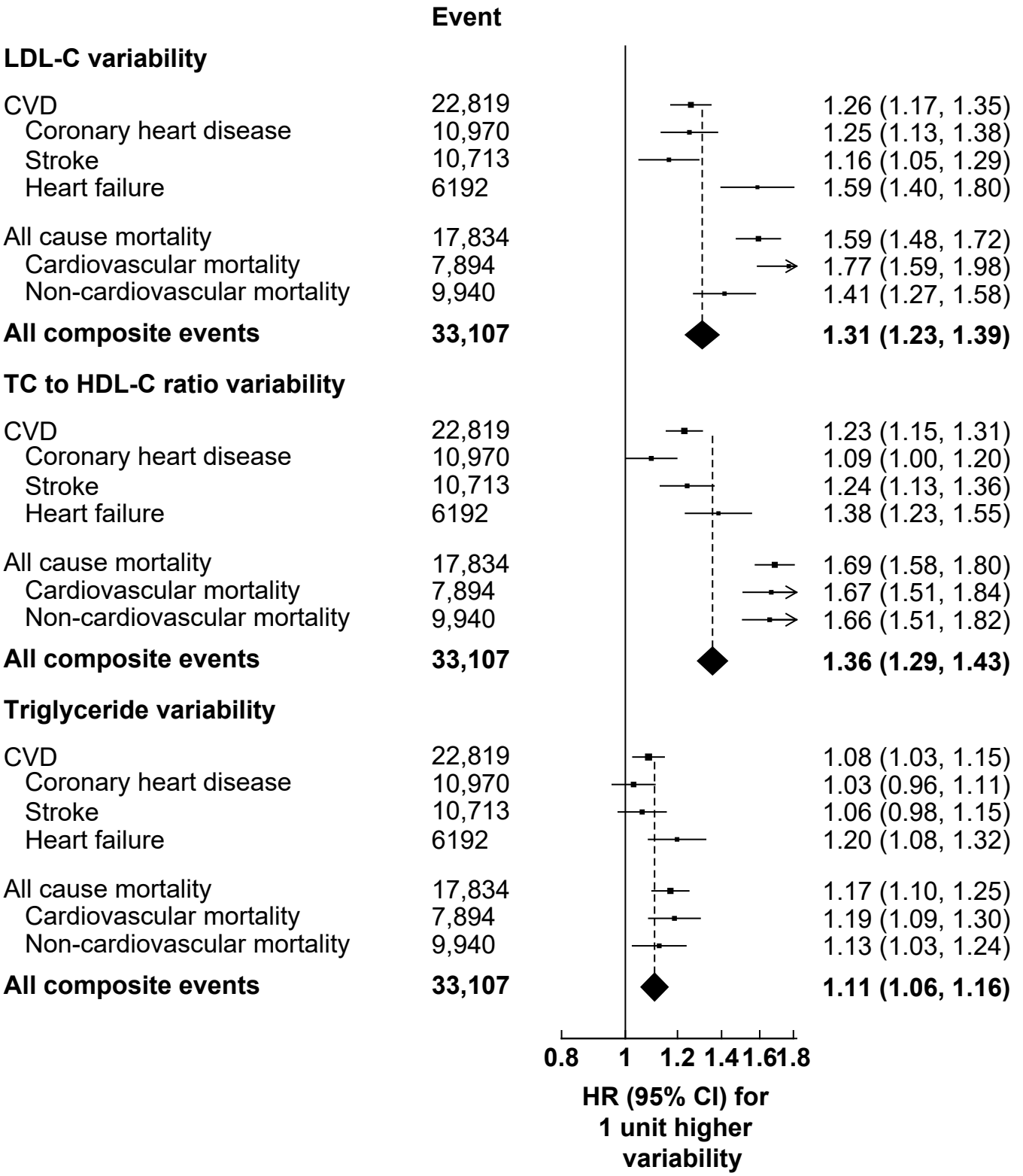
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti-diabetic drugs, anti-hypertensive drugs, statins and fibrates, Charlson's index and usual LDL-C, TC to HDL-C ratio or triglyceride (as appropriate). LDL-C = Low-density lipoprotein-cholesterol; TC = Total cholesterol; HDL-C = High-density lipoprotein-cholesterol; CVD = Cardiovascular disease; HR = Hazard ratio; CI = Confidence interval.

S.Figure 5b: Hazard ratios for the risk of CVD, coronary heart disease, stroke, heart failure, all-cause mortality, CVD mortality, non-CVD mortality and their composite with each 1 unit increase in LDL-C (mmol/L), TC to HDL-C ratio or triglyceride (mmol/L) variability using Cox regressions adjusted for baseline covariates in sensitivity analysis 2 (excluding patients with less than 12 months follow-up)



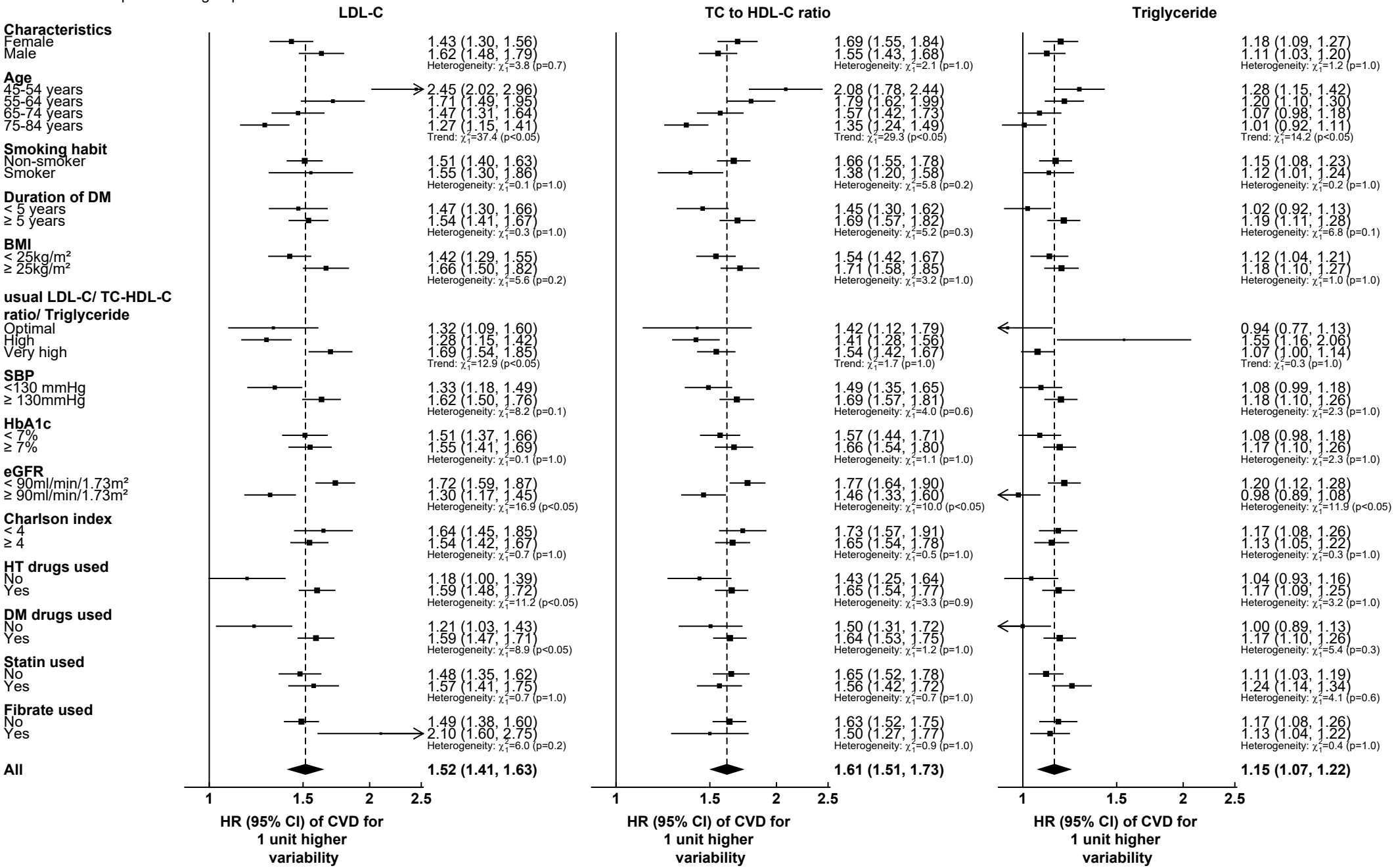
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti-diabetic drugs, anti-hypertensive drugs, statins and fibrates, Charlson's index and usual LDL-C, TC to HDL-C ratio or triglyceride (as appropriate). LDL-C = Low-density lipoprotein-cholesterol; TC = Total cholesterol; HDL-C = High-density lipoprotein-cholesterol; CVD = Cardiovascular disease; HR = Hazard ratio; CI = Confidence interval.

S.Figure 5c: Hazard ratios for the risk of CVD, coronary heart disease, stroke, heart failure, all-cause mortality, CVD mortality, non-CVD mortality and their composite with each 1 unit increase in LDL-C (mmol/L), TC to HDL-C ratio or triglyceride (mmol/L) variability using Cox regressions adjusted for baseline covariates in sensitivity analysis 3 (using 36 months collection period for cholesterol)



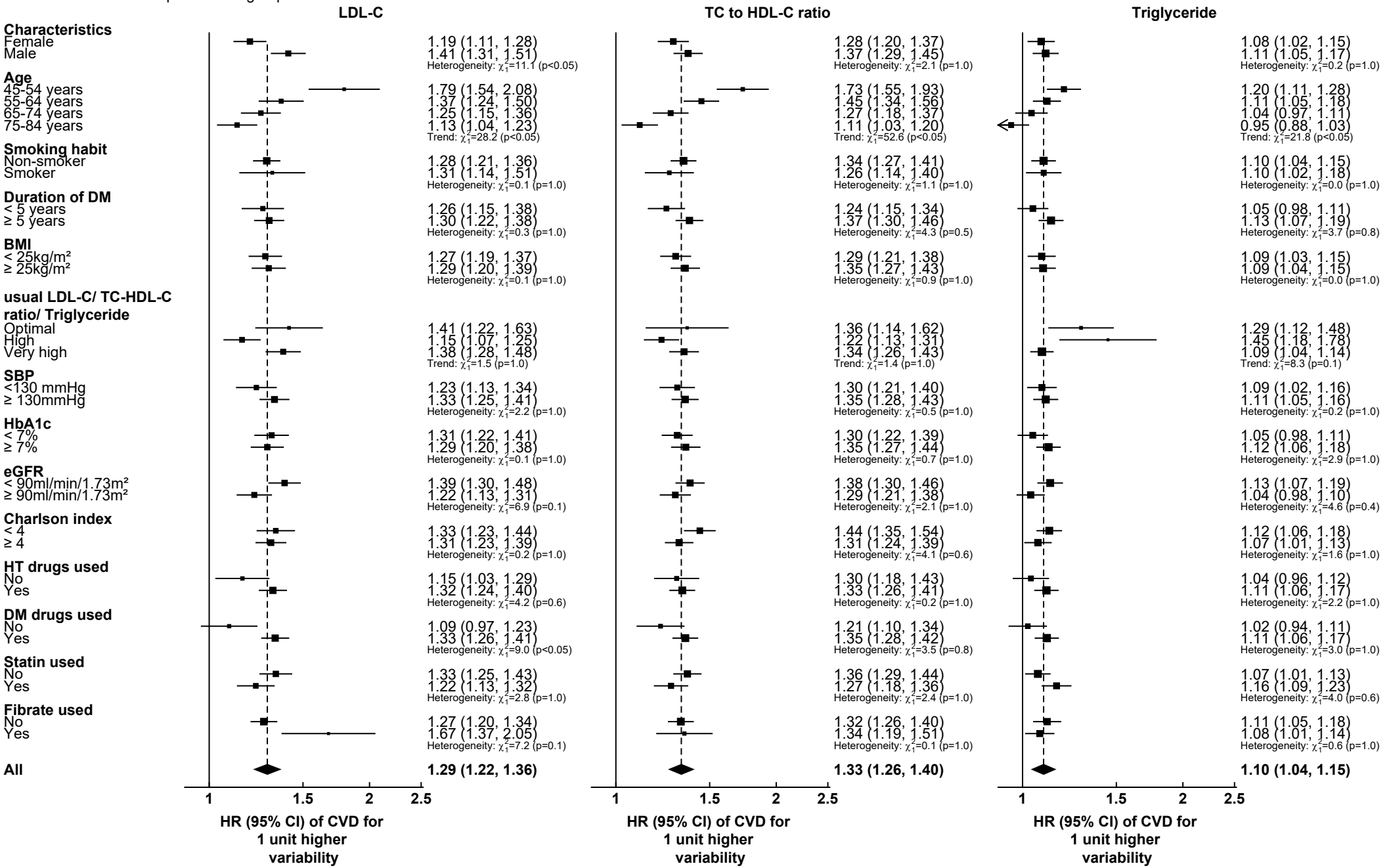
Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti-diabetic drugs, anti-hypertensive drugs, statins and fibrates, Charlson's index and usual LDL-C, TC to HDL-C ratio or triglyceride (as appropriate). LDL-C = Low-density lipoprotein-cholesterol; TC = Total cholesterol; HDL-C=High-density lipoprotein-cholesterol; CVD = Cardiovascular disease; HR = Hazard ratio; CI = Confidence interval.

S.Figure 6a: Hazard ratios for the association of a unit increase in LDL-C, TC to HDL-C ratio and triglyceride variability with CVD from Cox regression models adjusted for baseline covariates. A separate model is fitted for each specified subgroup



Hazard ratio was adjusted by age, gender, duration of diabetic mellitus, smoking status, body mass index, systolic blood pressure, diastolic blood pressure, haemoglobin A1c, estimated glomerular filtration rate, the usages of anti-diabetic drugs, anti-hypertensive drugs, statins and fibrates, Charlson's index and usual LDL-C, TC to HDL-C ratio or triglyceride (as appropriate). Optimal, high and very high level of LDL-C were defined as patients with <2.6, 2.6-3.4 and >3.4mmol/L, respectively; those for TC to HDL-C ratio were <3.5, 3.5-4.9 and ≥5, respectively; those for triglyceride were <1.8, 1.8-2.2 and ≥2.3mmol/L, respectively. CVD = Cardiovascular disease; BMI = Body Mass Index; LDL-C = Low-density lipoprotein-cholesterol; TC = Total cholesterol; HDL-C = High-density lipoprotein-cholesterol; HbA1c = Haemoglobin A1c; HT = Hypertension; DM = Diabetes mellitus; HR = Hazard ratio; eGFR = Estimated glomerular filtration rate; CI = Confidence interval.

S.Figure 6b: Hazard ratios for the association of a unit increase in LDL-C, TC to HDL-C ratio and triglyceride variability with mortality from Cox regression models adjusted for baseline covariates. A separate model is fitted for each specified subgroup



Hazard ratio was adjusted by age, gender, duration of DM, smoking status, BMI, SBP, DBP, HbA1c, eGFR, the usages of anti-diabetic drugs, anti-hypertensive drugs, statins and fibrates, Charlson's index and usual LDL-C, TC to HDL-C ratio or triglyceride (as appropriate). Optimal, high and very high level of LDL-C were defined as patients with <2.6, 2.6-3.4 and >3.4mmol/L, respectively; those for TC to HDL-C ratio were <3.5, 3.5-4.9 and ≥5, respectively; those for triglyceride were <1.8, 1.8-2.2 and ≥2.3mmol/L, respectively. BMI = Body Mass Index; SBP = Systolic blood pressure; DBP = Diastolic blood pressure; LDL-C = Low-density lipoprotein-cholesterol; TC = Total cholesterol; HDL-C = High-density lipoprotein-cholesterol; HbA1c = Haemoglobin A1c; HT = Hypertension; DM = Diabetes mellitus; HR = Hazard ratio; eGFR = Estimated glomerular filtration rate; CI = Confidence interval.