Supplementary Materials

Figures

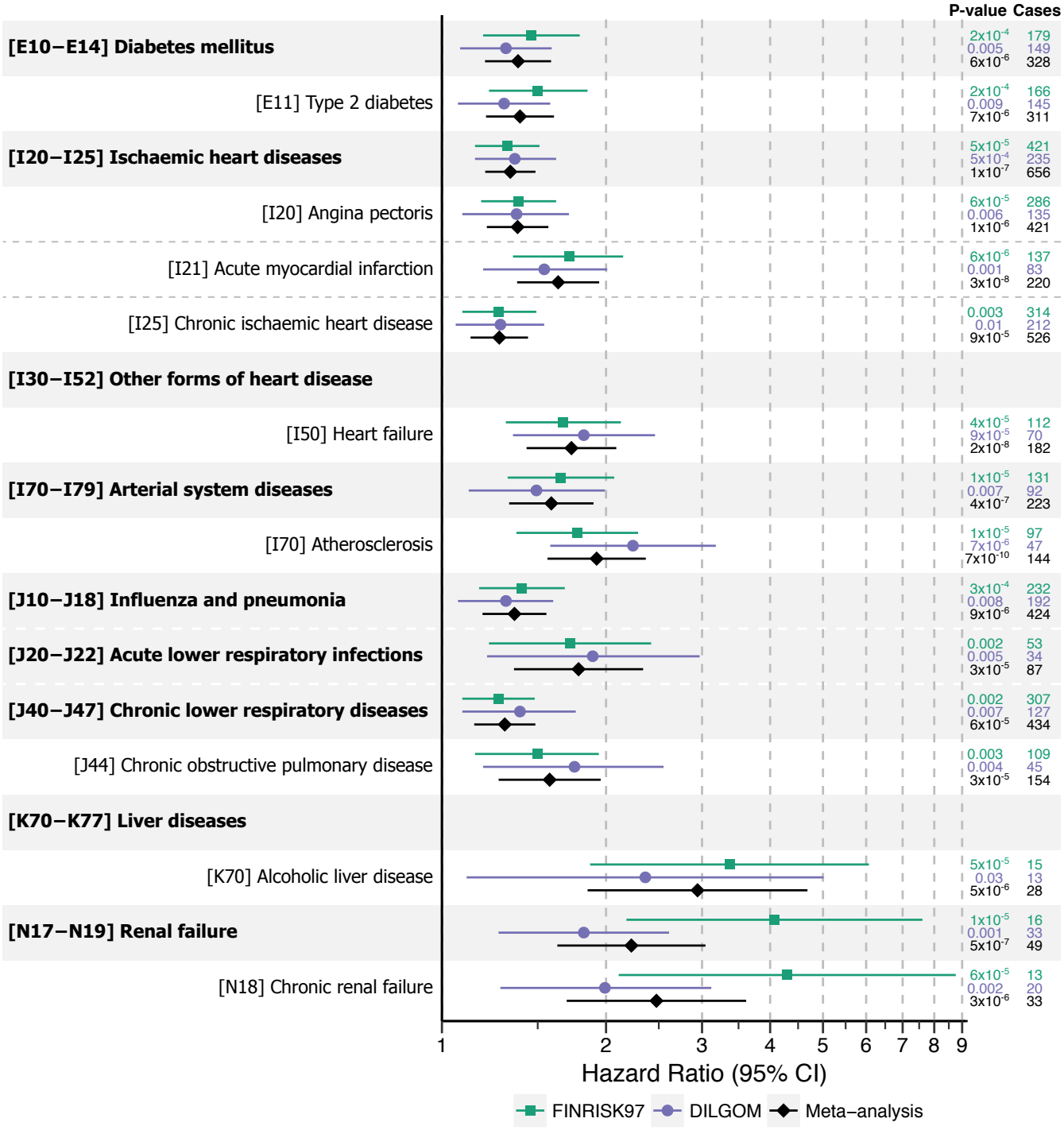

Figure S1. Cohort specific associations between GlycA and 8-year disease risk. Hazard ratios (diamonds) and 95% confidence intervals for the first diagnosis occurrence (hospitalisation or mortality) conferred per standard deviation increase of GlycA in FINRISK97 (green), DILGOM (blue), and in meta-analysis (black). The figure shows all ICD categories and codes where the association GlycA was

nominally significant ($P < 0.05$) and the association in meta-analysis was significant after multiple testing correction ($P < 1.1 \times 10^{-4}$; adjusting for the 468 tested outcomes). HRs and 95% CIs are shown on a natural logarithm scale. Diagnosis data were analysed for 468 outcomes with > 10 events in both cohorts over a matched 8-year follow-up period. Alphanumeric codes in the square brackets indicate the ICD10 codes/categories for each outcome. Cox models were fit using age as the time scale and adjusting for sex, smoking status, BMI, blood pressure, alcohol consumption, and biomarkers for all-cause mortality identified alongside GlycA: citrate, albumin, and VLDL particle size. Prevalent cases were excluded for each association test (**Methods**).

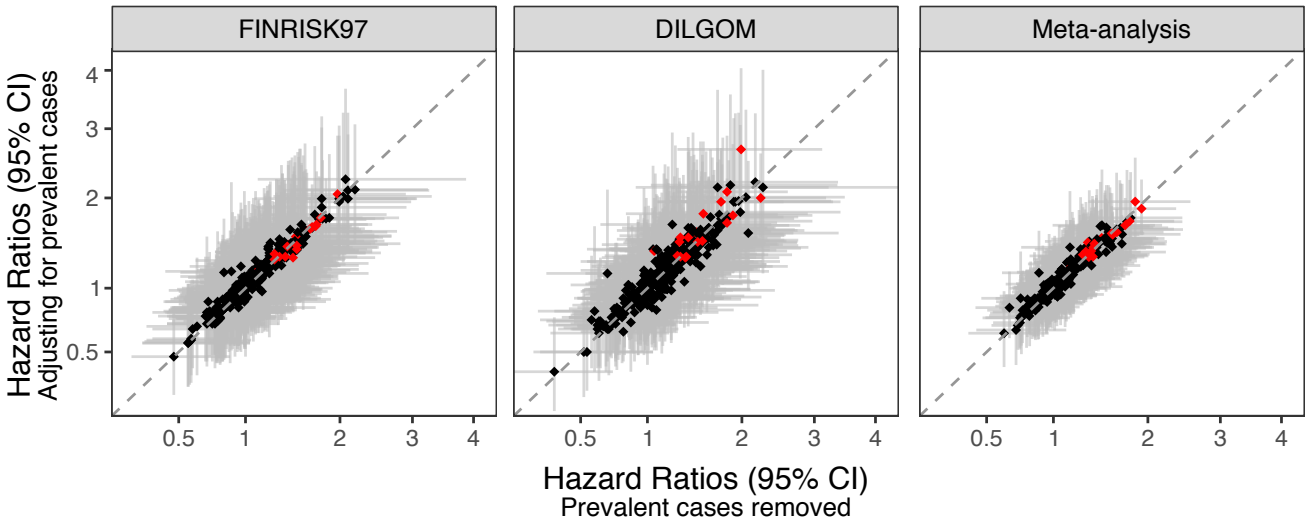

Figure S2. Sensitivity analysis to prevalent disease. Each panel compares the hazard ratios (diamonds) calculated excluding prevalent cases (x-axes) to hazard ratios calculated when adjusting for prevalent cases as a covariate (y-axes) in DILGOM, FINRISK97, and meta-analysis. Horizontal and vertical grey bars indicate the 95% confidence intervals for each outcome excluding prevalent cases (horizontal bars, x-axes) and adjusting for prevalent cases (vertical grey bars, y-axes) respectively. The diagonal dashed line shows y=x, the location where hazard ratios would fall if their estimates were identical in both models. Data are shown on a square root scale. Red diamonds indicate the significantly associated outcomes shown in **Figure 1** and **Figure S1**. Diamonds above the dashed diagonal line and > 1 indicate hazard ratios which are stronger in the model in which prevalent cases are adjusted. Cox proportional hazard models adjusted for prevalent cases were only analysed for > 20 incident cases (N=356 electronic health records in both DILGOM and FINRISK97) as the statistical models became unstable with lower number of incident cases adjusted for prevalent disease. Adjustment for prevalent cases was performed where there were > 10 cases of the respective electronic health records in the 10 years prior to sample collection.

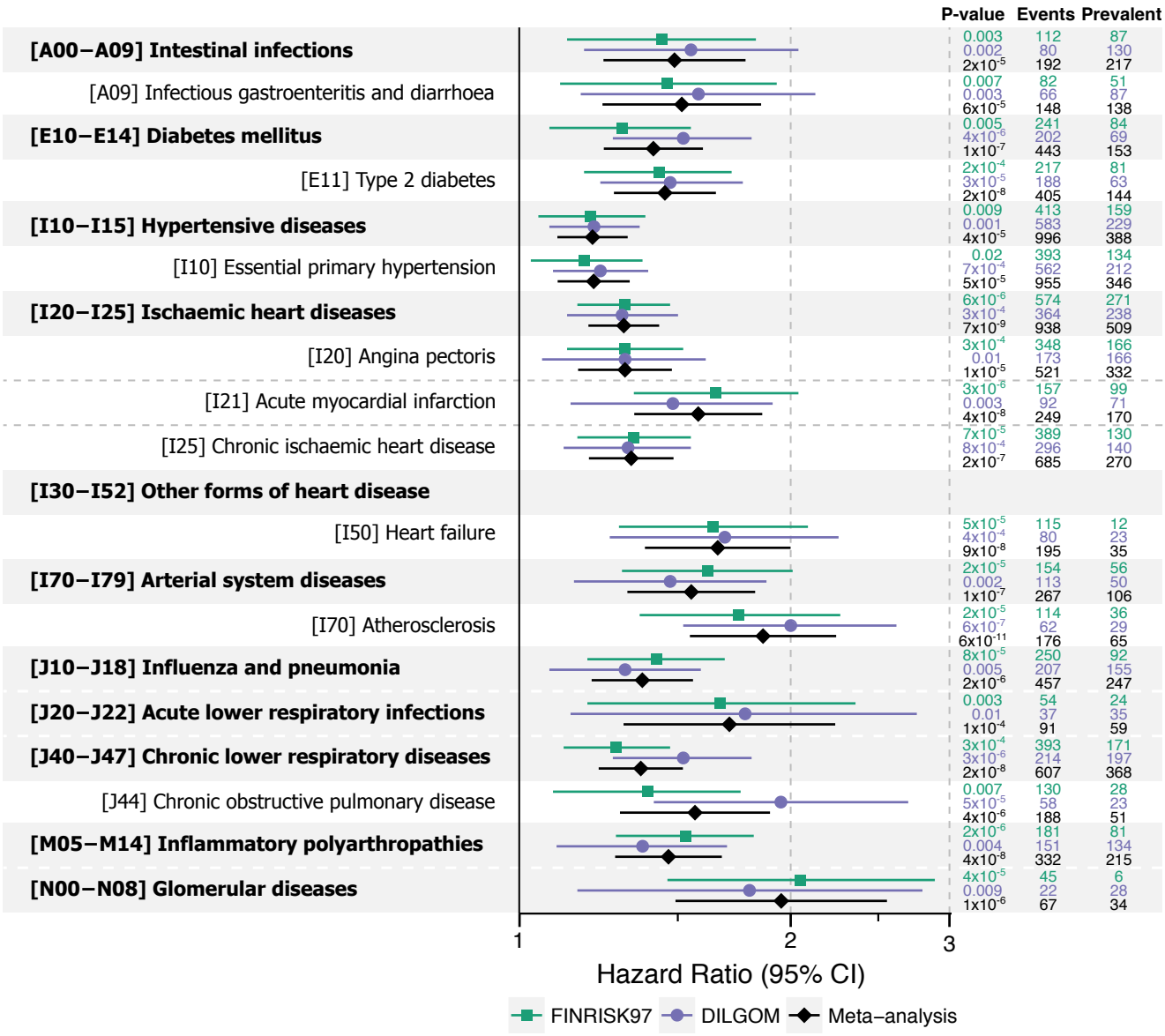

Figure S3. 8-year disease risk for GlycA and when adjusting for prevalent cases. Cox proportional hazard ratios (diamonds) and 95% confidence intervals for the first diagnosis occurrence (hospitalisation or mortality) conferred per standard deviation increase of GlycA in FINRISK97, DILGOM, and in meta-analysis. Diagnosis data were analysed for 356 outcomes with > 20 events in both cohorts over a matched 8-year follow-up period and adjustment for prevalent cases was performed where there were > 10 prevalent cases. The figure shows all ICD categories and codes where the association GlycA was nominally significant ($P < 0.05$) and the association in meta-analysis was significant after multiple testing correction ($P < 1.2\times10^{-4}$; adjusting for the 356 tested outcomes). HRs and 95% CIs are shown on a natural logarithm scale. Alphanumeric codes in the square brackets indicate the ICD10 codes/categories for each outcome. Models were fit using age as the time scale and adjusting for sex, smoking status, BMI, blood pressure, alcohol consumption, biomarkers for all-cause mortality identified alongside GlycA (citrate, albumin, and VLDL particle size), and prevalent cases of the outcome within 10 years prior to sample collection. Hazard ratios are detailed in **Table S3**.

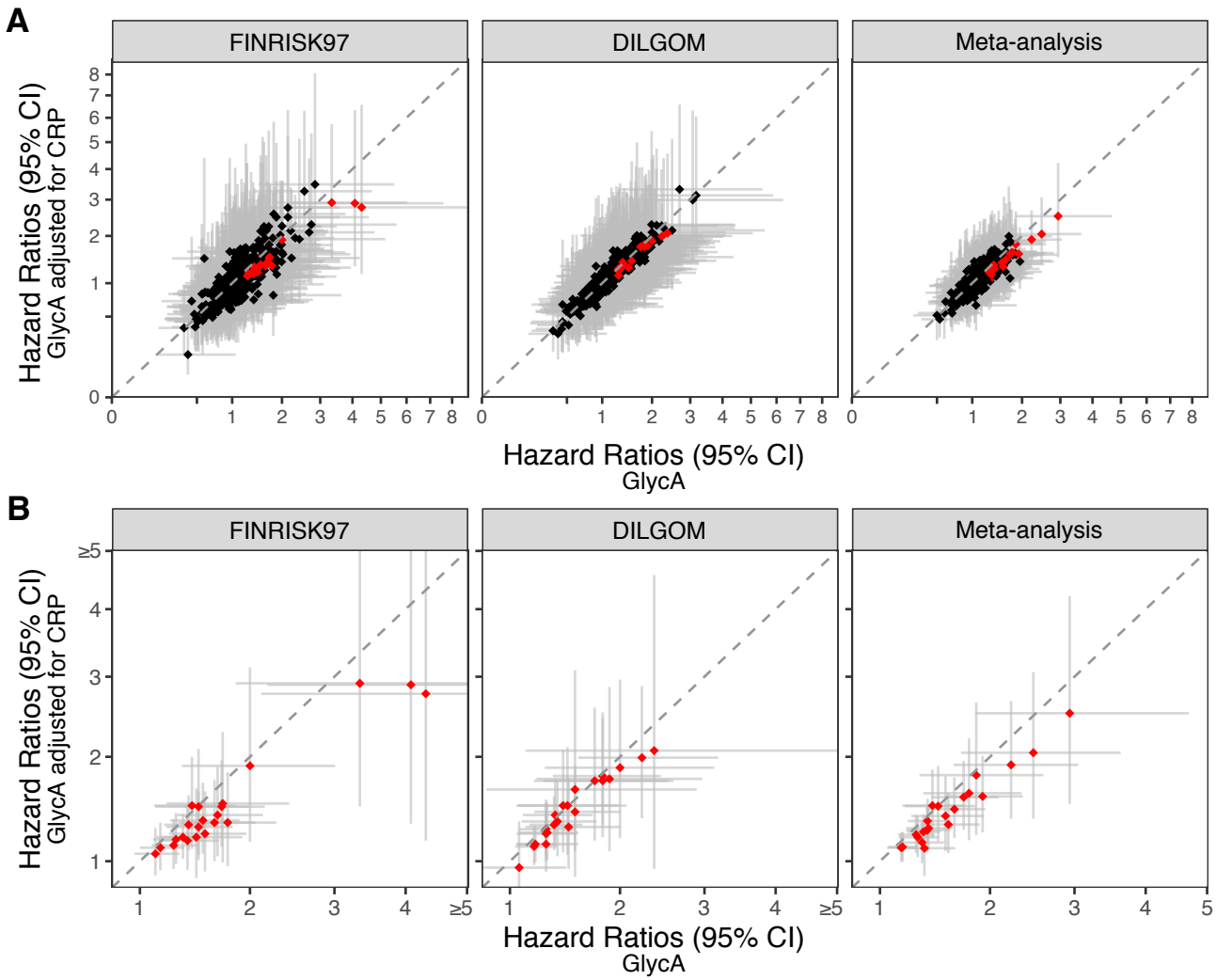

Figure S4. Sensitivity analysis to C-reactive protein. Each plot compares the hazard ratios (diamonds) conferred per standard deviation increase of GlycA (x-axes) to hazard ratios conferred per standard deviation increase of GlycA when adjusting for CRP (y-axes) in DILGOM, FINRISK97, and meta-analysis. Horizontal and vertical grey bars indicate the 95% confidence intervals for each hazard ratio for GlycA (horizontal bars, x-axes) and for GlycA adjusted for CRP (vertical grey bars, y-axes) respectively. The diagonal dashed line shows y=x, the location where hazard ratios would fall if their estimates were identical in both models. Diamonds above the dashed diagonal line and > 1 indicate hazard ratios which are stronger after adjusting GlycA for CRP. Data are shown on a square root scale. Red diamonds indicate the significantly associated outcomes shown in **Figure 1** and **Figure S1**. Panel **A)** shows all EHRs while **B)** shows just the significant outcomes. 95% confidence intervals in **B)** with a lower limit ≤ 0.8 or an upper limit ≥ 5 are truncated. Models were fit using age as the time scale and adjusting for sex, smoking status, BMI, blood pressure, alcohol consumption and biomarkers for all-cause mortality identified alongside GlycA: citrate, albumin, and VLDL particle size. Prevalent cases within 10 years prior to sample collection were excluded for each association test.

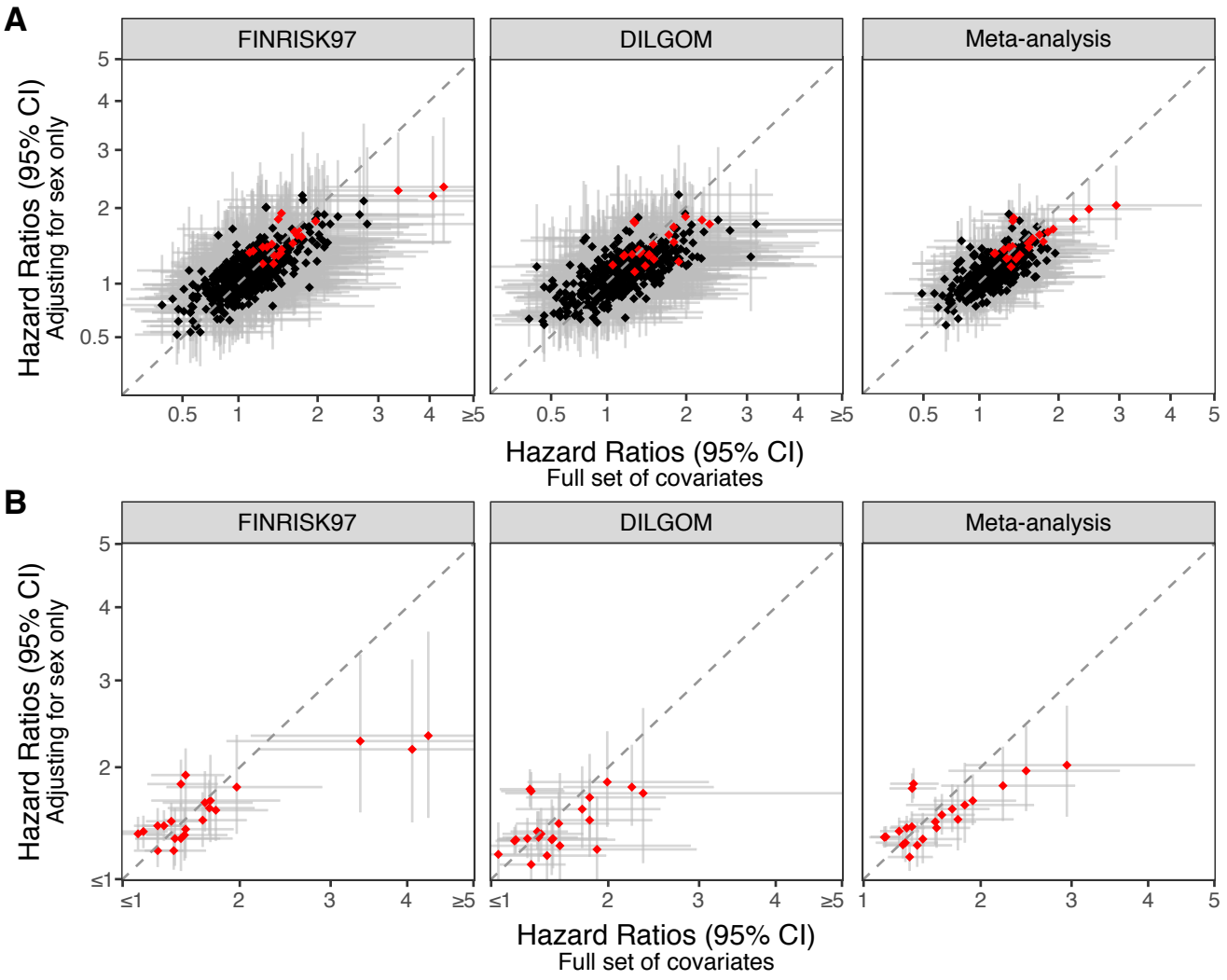

Figure S5. Sensitivity analysis to models with minimal adjustment. Each panel compares the hazard ratios (diamonds) conferred per standard deviation increase of GlycA adjusting for the full set of covariates (x-axes) to hazard ratios conferred per standard deviation increase of GlycA when adjusting for sex only (y-axes) in DILGOM, FINRISK97, and meta-analysis. The full set of covariates used in Cox proportional hazards models shown on the x-axis were sex, smoking status, BMI, blood pressure, alcohol consumption and biomarkers for all-cause mortality identified alongside GlycA: citrate, albumin, and VLDL particle size. In both models age was used as the time-scale and prevalent cases within 10 years prior to sample collection were excluded for each association test. Horizontal and vertical grey bars indicate the 95% confidence intervals for each hazard ratio for GlycA when adjusting for the full set of covariates (horizontal bars, x-axes) and for GlycA adjusting for sex only (vertical grey bars, y-axes) respectively. The diagonal dashed line shows y=x, the location where hazard ratios would fall if their estimates were identical in both models. Data are shown on a square root scale. Red diamonds indicate the significantly associated outcomes shown in **Figure 1** and **Figure S1**. Panel **A)** shows all EHRs while **B)** shows just the significant outcomes. 95% confidence intervals in **A)** with an upper limit ≥ 5 are truncated. 95% confidence intervals in **B)** with a lower limit ≤ 1 or an upper limit ≥ 5 are truncated.

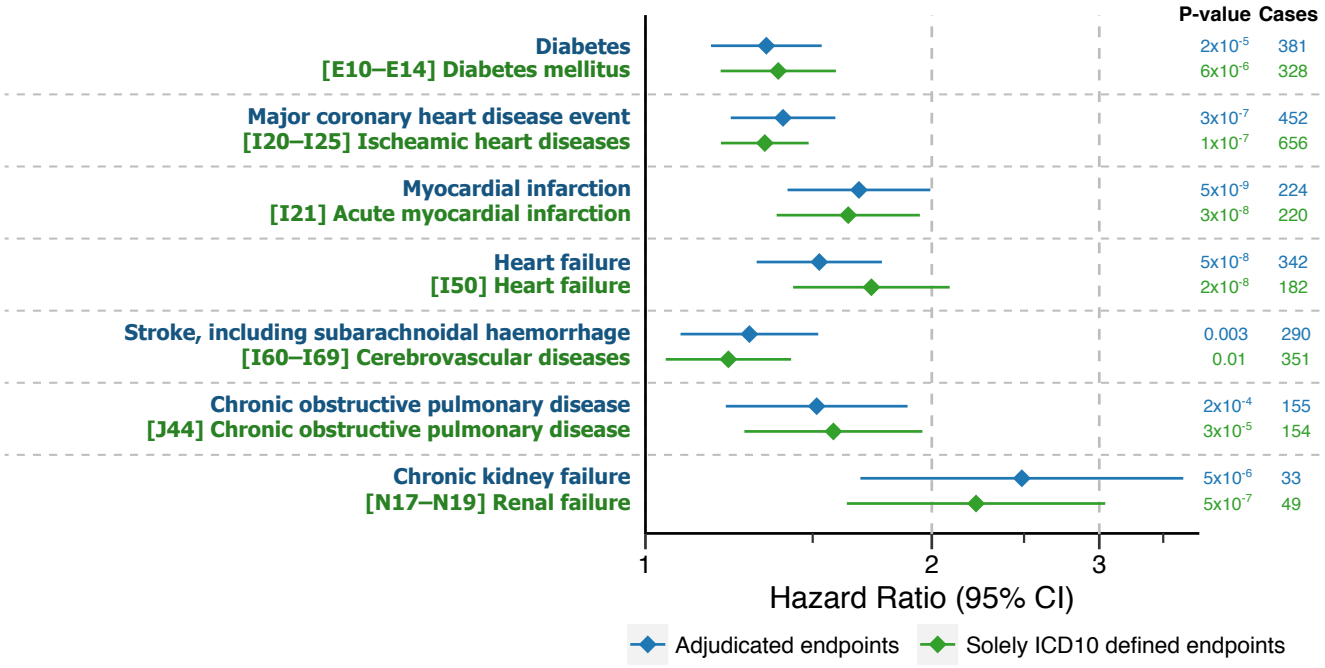

Figure S6. Sensitivity analysis using broader registry information for major common diseases. Cox proportional hazard ratios (diamonds) and 95% confidence intervals for the first diagnosis occurrence (**Methods**) conferred per standard deviation increase of GlycA in meta-analysis of FINRISK97 and DILGOM. The figure shows selected major common diseases from the broader registry information compared to their closest equivalent outcome (ICD10 code or category) in the EHR scan. Models were fit using age as the time scale and adjusting for sex, smoking status, BMI, blood pressure, alcohol consumption, and biomarkers for all-cause mortality identified alongside GlycA: citrate, albumin, and VLDL particle size. Prevalent cases detected using the same registry information were excluded from the analysis. Hazard ratios for the adjudicated endpoint definitions are detailed in **Table S4**.

Tables

Table S1. Listing of outcomes analysed for an association with GlycA. NB this table is provided as a separate Excel file.

Table S2: Electronic health record associations between GlycA and 8-year disease risk. Details of the Cox proportional hazard models shown in **Figure 1** and **Figure S1**. Models were fit with age as time scale, adjusting for sex, smoking status, BMI, blood pressure, alcohol consumption, and biomarkers for all-cause mortality identified alongside GlycA: citrate, albumin, and VLDL particle size. Diagnosis data was obtained from a matched 8-year follow-up period of the DILGOM and FINRISK97 cohorts. Prevalent cases within 10 years prior to sample collection were excluded from association tests between GlycA and each outcome. Alphanumeric codes within the square brackets indicate the ICD10 categories or codes for each outcome. Meta-analysis results were obtained from an inverse variance weighted fixed effects meta-analysis of the DILGOM and FINRISK97 cohorts. HR: the hazard ratio conferred per standard deviation increase of GlycA levels. SE: standard error of the hazard ratio estimate. 95% CI: 95% confidence interval of the hazard ratio.

Outcome	Cohort	Events	HR	SE	95% CI	P-value
[E10–E14] Diabetes mellitus	Meta-analysis	328	1.38	0.07	1.20–1.59	6×10^{-6}
	DILGOM	149	1.31	0.10	1.08–1.59	0.005
	FINRISK97	179	1.46	0.10	1.19–1.79	2×10^{-4}
[E11] Type 2 diabetes	Meta-analysis	311	1.39	0.07	1.20–1.61	7×10^{-6}
	DILGOM	145	1.30	0.10	1.07–1.58	0.009
	FINRISK97	166	1.50	0.11	1.22–1.85	2×10^{-4}
[I20–I25] Ischaemic heart diseases	Meta-analysis	656	1.34	0.05	1.20–1.48	1×10^{-7}
	DILGOM	235	1.36	0.09	1.15–1.62	5×10^{-4}
	FINRISK97	421	1.32	0.07	1.15–1.51	5×10^{-5}
[I20] Angina pectoris	Meta-analysis	421	1.38	0.07	1.21–1.57	1×10^{-6}
	DILGOM	135	1.37	0.11	1.09–1.71	0.006
	FINRISK97	286	1.38	0.08	1.18–1.62	6×10^{-5}
[I21] Acute myocardial infarction	Meta-analysis	220	1.63	0.09	1.37–1.94	3×10^{-8}
	DILGOM	83	1.54	0.13	1.19–2.01	0.001
	FINRISK97	137	1.71	0.12	1.35–2.15	6×10^{-6}
[I25] Chronic ischaemic heart disease	Meta-analysis	526	1.27	0.06	1.13–1.44	9×10^{-5}
	DILGOM	212	1.28	0.10	1.06–1.54	0.01
	FINRISK97	314	1.27	0.08	1.09–1.49	0.003
[I50] Heart failure	Meta-analysis	182	1.73	0.10	1.43–2.09	2×10^{-8}
	DILGOM	70	1.82	0.15	1.35–2.46	9×10^{-5}
	FINRISK97	112	1.67	0.12	1.31–2.13	4×10^{-5}
[I70–I79] Arterial system diseases	Meta-analysis	223	1.59	0.09	1.33–1.90	4×10^{-7}
	DILGOM	92	1.49	0.15	1.12–1.99	0.007
	FINRISK97	131	1.65	0.12	1.32–2.07	1×10^{-5}
[I70] Atherosclerosis	Meta-analysis	144	1.92	0.11	1.56–2.37	7×10^{-10}
	DILGOM	47	2.24	0.18	1.58–3.18	7×10^{-6}
	FINRISK97	97	1.77	0.13	1.37–2.29	1×10^{-5}
[J10–J18] Influenza and pneumonia	Meta-analysis	424	1.36	0.07	1.19–1.55	9×10^{-6}
	DILGOM	192	1.31	0.10	1.07–1.60	0.008
	FINRISK97	232	1.40	0.09	1.17–1.68	3×10^{-4}
[J20–J22] Acute lower respiratory infections	Meta-analysis	87	1.78	0.14	1.36–2.34	3×10^{-5}
	DILGOM	34	1.89	0.23	1.21–2.97	0.005
	FINRISK97	53	1.72	0.18	1.22–2.42	0.002
[J40–J47] Chronic lower respiratory diseases	Meta-analysis	434	1.30	0.07	1.15–1.48	6×10^{-5}
	DILGOM	127	1.39	0.12	1.09–1.76	0.007
	FINRISK97	307	1.27	0.08	1.09–1.48	0.002

Outcome	Cohort	Events	HR	SE	95% CI	P-value
[J44] Chronic obstructive pulmonary disease	Meta-analysis	154	1.58	0.11	1.27–1.96	3×10^{-5}
	DILGOM	45	1.75	0.19	1.19–2.55	0.004
	FINRISK97	109	1.50	0.13	1.15–1.94	0.003
[K70] Alcoholic liver disease	Meta-analysis	28	2.94	0.24	1.85–4.68	5×10^{-6}
	DILGOM	13	2.36	0.38	1.11–5.01	0.03
	FINRISK97	15	3.37	0.30	1.87–6.07	5×10^{-5}
[N17–N19] Renal failure	Meta-analysis	49	2.23	0.16	1.63–3.04	5×10^{-7}
	DILGOM	33	1.82	0.18	1.27–2.61	0.001
	FINRISK97	16	4.07	0.32	2.18–7.61	1×10^{-5}
[N18] Chronic renal failure	Meta-analysis	33	2.47	0.19	1.69–3.61	3×10^{-6}
	DILGOM	20	1.99	0.23	1.28–3.12	0.002
	FINRISK97	13	4.30	0.36	2.11–8.76	6×10^{-5}

Table S3. Hazard ratio details for GlycA associations and when adjusting for prevalent cases. Details of the Cox proportional hazard models shown in **Figure S3**. Models were fit with age as time scale, adjusting for sex, smoking status, BMI, blood pressure, alcohol consumption, biomarkers for all-cause mortality identified alongside GlycA (citrate, albumin, and VLDL particle size), and prevalent cases of the outcome within 10 years prior to sample collection. Diagnosis data were analysed for 356 outcomes with > 20 events in both cohorts over a matched 8-year follow-up period and adjustment for prevalent cases was performed where there were > 10 prevalent cases. Alphanumeric codes within the square brackets indicate the ICD10 categories or codes for each outcome. Meta-analysis results were obtained from an inverse variance weighted fixed effects meta-analysis of the DILGOM and FINRISK97 cohorts. #E: the number of people who were hospitalised or died from the outcome within the 8-year follow-up period. #P: the number of people who had been hospitalised from that outcome in the 10 years prior to sample collection. HR: the hazard ratio conferred per standard deviation increase of GlycA levels. SE: standard error of the hazard ratio estimate. 95% CI: 95% confidence interval of the hazard ratio.

Outcome	Cohort	#E	#P	HR	SE	95% CI	P-value
[A00-A09] Intestinal infections	Meta-analysis	192	217	1.49	0.09	1.24–1.78	2×10^{-5}
	DILGOM	80	130	1.55	0.14	1.18–2.04	0.002
	FINRISK97	112	87	1.44	0.12	1.13–1.83	0.003
[A09] Infectious gastroenteritis and diarrhoea	Meta-analysis	148	138	1.51	0.10	1.24–1.85	6×10^{-5}
	DILGOM	66	87	1.58	0.15	1.17–2.13	0.003
	FINRISK97	82	51	1.46	0.14	1.11–1.93	0.007
[E10-E14] Diabetes mellitus	Meta-analysis	443	153	1.41	0.06	1.24–1.60	1×10^{-7}
	DILGOM	202	69	1.52	0.09	1.27–1.81	4×10^{-6}
	FINRISK97	241	84	1.3	0.09	1.08–1.55	0.005
[E11] Type 2 diabetes	Meta-analysis	405	144	1.45	0.07	1.27–1.65	2×10^{-8}
	DILGOM	188	63	1.47	0.09	1.23–1.77	3×10^{-5}
	FINRISK97	217	81	1.43	0.10	1.18–1.72	2×10^{-4}
[I10-I15] Hypertensive diseases	Meta-analysis	996	388	1.21	0.05	1.10–1.32	4×10^{-5}
	DILGOM	583	229	1.21	0.06	1.08–1.36	0.001
	FINRISK97	413	159	1.2	0.07	1.05–1.38	0.009
[I10] Essential primary hypertension	Meta-analysis	955	346	1.21	0.05	1.10–1.33	5×10^{-5}
	DILGOM	562	212	1.23	0.06	1.09–1.39	7×10^{-4}
	FINRISK97	393	134	1.18	0.07	1.03–1.37	0.02
[I20-I25] Ischaemic heart diseases	Meta-analysis	938	509	1.31	0.05	1.19–1.43	7×10^{-9}
	DILGOM	364	238	1.3	0.07	1.13–1.50	3×10^{-4}
	FINRISK97	574	271	1.31	0.06	1.16–1.47	6×10^{-6}
[I20] Angina pectoris	Meta-analysis	521	332	1.31	0.06	1.16–1.48	1×10^{-5}
	DILGOM	173	166	1.31	0.11	1.06–1.61	0.01
	FINRISK97	348	166	1.31	0.07	1.13–1.52	3×10^{-4}
[I21] Acute myocardial infarction	Meta-analysis	249	170	1.58	0.08	1.34–1.86	4×10^{-8}
	DILGOM	92	71	1.48	0.13	1.14–1.91	0.003
	FINRISK97	157	99	1.65	0.11	1.34–2.04	3×10^{-6}
[I25] Chronic ischaemic heart disease	Meta-analysis	685	270	1.33	0.06	1.19–1.48	2×10^{-7}
	DILGOM	296	140	1.32	0.08	1.12–1.55	8×10^{-4}
	FINRISK97	389	130	1.34	0.07	1.16–1.55	7×10^{-5}
[I50] Heart failure	Meta-analysis	195	35	1.66	0.09	1.38–2.00	9×10^{-8}
	DILGOM	80	23	1.69	0.15	1.26–2.26	4×10^{-4}
	FINRISK97	115	12	1.64	0.12	1.29–2.09	5×10^{-5}

Outcome	Cohort	#E	#P	HR	SE	95% CI	P-value
[I70-I79] Arterial system diseases	Meta-analysis	267	106	1.55	0.08	1.32–1.83	1×10^{-7}
	DILGOM	113	50	1.47	0.12	1.15–1.88	0.002
	FINRISK97	154	56	1.62	0.11	1.30–2.01	2×10^{-5}
[I70] Atherosclerosis	Meta-analysis	176	65	1.86	0.10	1.55–2.25	6×10^{-11}
	DILGOM	62	29	2	0.14	1.52–2.62	6×10^{-7}
	FINRISK97	114	36	1.75	0.13	1.36–2.27	2×10^{-5}
[J10-J18] Influenza and pneumonia	Meta-analysis	457	247	1.37	0.07	1.20–1.56	2×10^{-6}
	DILGOM	207	155	1.31	0.10	1.08–1.59	0.005
	FINRISK97	250	92	1.42	0.09	1.19–1.69	8×10^{-5}
[J20-J22] Acute lower respiratory infections	Meta-analysis	91	59	1.71	0.14	1.31–2.24	1×10^{-4}
	DILGOM	37	35	1.78	0.23	1.14–2.76	0.01
	FINRISK97	54	24	1.67	0.17	1.19–2.36	0.003
[J40-J47] Chronic lower respiratory diseases	Meta-analysis	607	368	1.36	0.05	1.22–1.52	2×10^{-8}
	DILGOM	214	197	1.52	0.09	1.27–1.81	3×10^{-6}
	FINRISK97	393	171	1.28	0.07	1.12–1.47	3×10^{-4}
[J44] Chronic obstructive pulmonary disease	Meta-analysis	188	51	1.57	0.10	1.29–1.90	4×10^{-6}
	DILGOM	58	23	1.95	0.16	1.41–2.70	5×10^{-5}
	FINRISK97	130	28	1.39	0.12	1.09–1.76	0.007
[M05-M14] Inflammatory polyarthropathies	Meta-analysis	332	215	1.46	0.07	1.28–1.68	4×10^{-8}
	DILGOM	151	134	1.37	0.11	1.10–1.70	0.004
	FINRISK97	181	81	1.53	0.09	1.28–1.82	2×10^{-6}
[N00-N08] Glomerular diseases	Meta-analysis	67	34	1.95	0.14	1.49–2.56	1×10^{-6}
	DILGOM	22	28	1.8	0.22	1.16–2.80	0.009
	FINRISK97	45	6	2.05	0.17	1.46–2.89	4×10^{-5}

Table S4. Hazard ratio details for GlycA associations with major common diseases using more detailed registry information. Details of the Cox proportional hazard models shown in **Figure S6**. Models were fit with age as time scale, adjusting for sex, smoking status, BMI, blood pressure, alcohol consumption, and biomarkers for all-cause mortality identified alongside GlycA: citrate, albumin, and VLDL particle size. Prevalent cases were removed from the analysis. HR: the hazard ratio conferred per standard deviation increase of GlycA levels. SE: standard error of the hazard ratio estimate. 95% CI: 95% confidence interval of the hazard ratio.

Outcome	Cohort	#E	HR	SE	95% CI	P-value
Diabetes	FINRISK97	237	1.38	0.10	1.14–1.67	8×10^{-4}
	DILGOM	144	1.30	0.10	1.07–1.57	0.008
	Meta-analysis	381	1.34	0.07	1.17–1.53	2×10^{-5}
Major coronary heart disease event	FINRISK97	277	1.46	0.08	1.24–1.73	7×10^{-6}
	DILGOM	175	1.31	0.10	1.08–1.59	0.007
	Meta-analysis	452	1.4	0.06	1.23–1.58	3×10^{-7}
Myocardial infarction	FINRISK97	140	1.79	0.12	1.43–2.26	6×10^{-7}
	DILGOM	84	1.54	0.13	1.18–2.00	0.001
	Meta-analysis	224	1.68	0.09	1.41–1.99	5×10^{-9}
Heart failure	FINRISK97	242	1.56	0.09	1.30–1.86	1×10^{-6}
	DILGOM	100	1.43	0.15	1.07–1.91	0.02
	Meta-analysis	342	1.52	0.08	1.31–1.77	5×10^{-8}
Stroke, including subarachnoidal haemorrhage	FINRISK97	189	1.23	0.11	1.00–1.52	0.05
	DILGOM	101	1.39	0.14	1.05–1.83	0.02
	Meta-analysis	290	1.29	0.09	1.09–1.52	0.003
Chronic obstructive pulmonary disease	FINRISK97	111	1.49	0.13	1.14–1.93	0.003
	DILGOM	44	1.57	0.20	1.05–2.35	0.03
	Meta-analysis	155	1.51	0.11	1.21–1.89	2×10^{-4}
Chronic kidney failure	FINRISK97	12	4.57	0.40	2.08–10.06	2×10^{-4}
	DILGOM	21	2.04	0.23	1.30–3.20	0.002
	Meta-analysis	33	2.49	0.20	1.68–3.68	5×10^{-6}

Table S5. Hazard ratio details for 12-year cardiovascular mortality risk in ANGES shown in **Figure 2**. Hazard ratios indicate risk of cardiovascular mortality relative to individuals in the lowest GlycA quintile. There were 180 people in each quintile. 95% CI: 95% confidence interval. Cox proportional hazard models were adjusted with age, sex, albumin, VLDL-diameter, citrate, thrombocyte count and ejection fraction.

GlycA Quintile	Deaths	Hazard Ratio	Standard Error	95% CI	P-value
20–40%	23	1.95	0.34	1.00–3.80	0.05
40–60%	26	2.35	0.36	1.16–4.77	0.02
60–80%	26	4.87	0.35	2.45–9.65	6×10^{-6}
80–100%	41	5.00	0.38	2.38–10.48	2×10^{-5}

Table S6. CRP hazard ratio details for 12-year cardiovascular mortality risk in ANGES. Hazard ratios indicate risk of cardiovascular mortality relative to individuals in the lowest CRP quintile. 95% CI: 95% confidence interval. Cox proportional hazard models were adjusted with age, sex, albumin, VLDL-diameter, citrate, thrombocyte count and ejection fraction. Analyses were conducted using 582 individuals with complete data.

CRP Quintile	Deaths	Hazard Ratio	Standard Error	95% CI	P-value
20–40%	23	1.77	0.43	0.76–4.14	0.2
40–60%	26	3.41	0.37	1.64–7.09	0.001
60–80%	26	3.13	0.38	1.49–6.56	0.003
80–100%	41	4.62	0.35	2.31–9.23	2×10^{-5}

Table S7: Hazard ratios and standard errors for all outcomes analysed in the EHR analyses. NB this table is provided as a separate Excel file. Hazard ratios and standard errors in each sensitivity analysis are provided as additional sheets in the Excel file.