openheart

Assessing the severity of cardiovascular disease in 213 088 patients with coronary heart disease: a retrospective cohort study

Salwa S Zghebi ,[1,2] Mamas A Mamas ,[2,3] Darren M Ashcroft,[1,4,5,6] Martin K Rutter,[7,8] Harm VanMarwijk,[9] Chris Salisbury,[10] Christian D Mallen,[11] Caroline A Chew-Graham,[11] Nadeem Qureshi ,[12] Stephen F Weng,[13] Tim Holt,[14] Iain Buchan,[2,15,16] Niels Peek,[5,6,16] Sally Giles,[2,5] David Reeves,[1,2,17] Evangelos Kontopantelis [1,2,16]

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/openhrt-2020-001498).

To cite: Zghebi SS, Mamas MA, Ashcroft DM, *et al*. Assessing the severity of cardiovascular disease in 213 088 patients with coronary heart disease: a retrospective cohort study. *Open Heart* 2021;**8**:e001498. doi:10.1136/openhrt-2020-001498

Received 29 October 2020
Revised 22 January 2021
Accepted 5 February 2021

For numbered affiliations see end of article.

Correspondence to
Dr Salwa S Zghebi; salwa.zghebi@manchester.ac.uk

ABSTRACT

Objective Most current cardiovascular disease (CVD) risk stratification tools are for people without CVD, but very few are for prevalent CVD. In this study, we developed and validated a CVD severity score in people with coronary heart disease (CHD) and evaluated the association between severity and adverse outcomes.

Methods Primary and secondary care data for 213 088 people with CHD in 398 practices in England between 2007 and 2017 were used. The cohort was randomly divided into training and validation datasets (80%/20%) for the severity model. Using 20 clinical severity indicators (each assigned a weight=1), baseline and longitudinal CVD severity scores were calculated as the sum of indicators. Adjusted Cox and competing-risk regression models were used to estimate risks for all-cause and cause-specific hospitalisation and mortality.

Results Mean age was 64.5±12.7 years, 46% women, 16% from deprived areas, baseline severity score 1.5±1.2, with higher scores indicating a higher burden of disease. In the training dataset, 138 510 (81%) patients were hospitalised at least once, and 39 944 (23%) patients died. Each 1-unit increase in baseline severity was associated with 41% (95% CI 37% to 45%, area under the receiver operating characteristics (AUROC) curve=0.79) risk for 1 year for all-cause mortality; 59% (95% CI 52% to 67%, AUROC=0.80) for cardiovascular (CV)/diabetes mortality; 27% (95% CI 26% to 28%) for any-cause hospitalisation and 37% (95% CI 36% to 38%) for CV/diabetes hospitalisation. Findings were consistent in the validation dataset.

Conclusions Higher CVD severity score is associated with higher risks for any-cause and cause-specific hospital admissions and mortality in people with CHD. Our reproducible score based on routinely collected data can help practitioners better prioritise management of people with CHD in primary care.

Key questions

What is already known about this subject?

- The majority of current cardiovascular disease (CVD) risk stratification tools are for people without CVD with very few tools available for people with prevalent CVD.
- Coronary heart disease (CHD) is the most common type of CVD and a leading cause of death globally. In the UK, CHD is responsible for one death around every 8 min.
- It is estimated that nearly 2.3 million people are living with CHD in the UK.
- The importance of assessing disease severity in people with CHD is well recognised, but validated CVD severity measures derived from routinely collected health records are lacking, as are applications of such measures in primary care settings.

What does this study add?

- We developed a new CVD severity score incorporating 20 severity indicators using patients' anonymised routinely collected electronic health records.
- In people with CHD, a 1-unit higher level of the severity score was linked to up to 59% significantly higher risk of hospital admission or death.

How might this impact on clinical practice?

- We demonstrate the utility and validity of a CVD-specific severity measure in people with CHD using routinely collected data.
- Our severity measure has potential applications directly relevant to clinical practice and risk stratification which informsadvanced decision making and provides a reproducible algorithm to other conditions managed in primary care.

INTRODUCTION

Cardiovascular disease (CVD) is the leading cause of death globally[1] and accounts for more than one in four UK deaths.[2] Coronary heart disease (CHD) is the most common CVD, accounting for nearly 9.5 million deaths worldwide in 2016.[2 3] Around 15.5 million people had CHD in the USA by 2016, and

2.3 million people in the UK by 2018 at a prevalence of 3%.[2 4]

Most currently available prospective cardiovascular (CV) risk stratification tools are for people without known CVD,[5] including QRISK and Framingham scores[6 7] with very few tools available to help assess the disease severity in people with existing CVD. In the context of this paper, we adopt the definition of severity of clinical conditions as the manifestation of the progression of underlying disease processes with implications on healthcare resources utilisation, multimorbidity and mortality.[8 9]

To our knowledge, no established CVD severity scores for primary care patients with CHD exist, and previous cohort studies are sparse,[10] with the majority of literature based on clinical trials of different sizes or using various sources of data mainly captured in secondary care facilities. Such attempts either focused primarily on existing scores/indices (such as SYNTAX,[11] Gensini,[12] the Duke CAD Prognostic Severity Index[13] and CAD-RADS[14]); the prevalence of multivessel disease; or the degree of coronary stenosis and/or lesions.[13–19] However, existing scores are not designed for primary care settings and only subserve a small minority of patients. While other scores would need invasive interventions that may not be indicated (or in minority of patients) and therefore resources needed for such information would be limited and not routinely available in primary care settings. CV-specific severity measures derived from routine clinical records of CVD progression are needed and could support practitioners to provide better clinical management as well as help healthcare policy makers and planners in developing services and allocating resources.

Since all of the above approaches rely on data that are not necessarily available in routine primary care health records for all patients, they are not useful for informing decisions at a primary care or public health levels based on identifying patients at risk of adverse outcomes. Currently available routinely collected electronic health records (EHRs) provide a platform for developing disease severity indices that are informative in stratifying CHD populations.

We, therefore, developed a severity score derived from routine EHR in UK primary care to stratify CHD populations in terms of CVD severity as a means of risk stratification. We aimed to: (1) develop and internally validate baseline and longitudinal CV severity scores in individuals with CHD and (2) assess what the score adds to the predictive value of sociodemographic variables for the risks for all-cause and cause-specific hospitalisation and mortality outcomes.

METHODS

Data source and patient population

In this retrospective cohort study, we used the GOLD database of the Clinical Practice Research Datalink (CPRD). The CPRD is one of the world's largest EHR databases providing anonymised medical data (including demographics, tests, diagnoses, referrals and prescriptions) and is broadly representative of the UK population.[20 21] CPRD provides data linkage to additional datasets and disease registries. We used the following linkages: Hospital Episodes Statistics Admitted Patient Care(HES APC), Office for National Statistics (ONS) cause-specific mortality data and index of multiple deprivation (IMD). The IMD used in our study is recorded at the level of the patient's residential postcode in England and is a score calculated as the weighted sum of 37 individual indicators organised across seven domains of deprivation: access to housing and services, crime, employment, education, income, finance and living environment.[22] Theemployment and income deprivation domains contribute the most weight to the overall index.

Patients with CHD (defined as patients with ≥1 CHD code listed in online supplemental table S1) aged ≥35 years and registered in linked general practices in England were identified between 1 March 2007 and 31 March 2017. The validity of CVD diagnoses in CPRD data has been acknowledged previously.[23] For each patient, the index date was defined as the earliest CHD diagnosis date. Patients were followed up until the earliest date of: developing an outcome; leaving the general practice; study end (31 March 2017); or death. By definition of multiple event models (as in Poisson models), developing the outcome of interest was not a censor point for those analyses. The final cohort of eligible patients was randomly split into training (80%) and validation (20%) datasets. The 20% split of the dataset was used to replicate the analyses performed in the training dataset as a validation.

Severity scores

A scoping review of indicators and markers of disease severity in people with CHD combined by the team's clinical expert opinion on CVD severity was used to identify clinically relevant CV severity indicators in people with CHD. A total of 20 CV indicators were used: hypertension; hyperlipidaemia; proteinuria/albuminuria; end-stage renal disease; peripheral vascular disease; stable angina; cardiac arrest; atrial fibrillation/supraventricular tachycardia; myocardial infarction/acute coronary syndrome; heart valve disease; endocarditis; myocarditis; cardiomyopathy; pericardial disease; ventricular tachycardia/fibrillation; congestive heart failure; CV procedures; transient ischaemic attack or stroke; diabetes; and pacemaker/defibrillator use. The Read codes for severity indicators recorded in CPRD were identified using the (pcdsearch) Stata command.[24]

Based on the timing of severity indicators, the severity score was calculated as the sum of indicators (each assigned weight=1) recorded at preindex (on/before first CHD diagnosis date, ie, baseline severity) or postindex (after first CHD diagnosis, ie, longitudinal severity) windows (online supplemental table S2). For preindex scores, indicators recorded in three look-back windows were considered: ever before

(unlimited look-back window), up to 10 and up to 5 years before index. This aimed to investigate the effects of varying the length of the preindex record on the model fitness in order to identify the optimal look-back window for prediction of future adverse outcomes but simultaneously considering the data quality that improved in recent years. For postindex scores, indicators recorded annually in years 1–10 after index date were considered, each combined with each of the three look-back windows. Postindex scores aimed to assess the trends of CV severity over time and how the risk for adverse outcomes change up to 10 years after CHD diagnosis.

Covariates

Age at baseline, gender, socioeconomic status (IMD 2015 quintiles 1–5 or unknown) and ethnicity (white, black, Asian, mixed, other or unknown).

Outcomes

Primary outcome was all-cause mortality. Secondary outcomes were: clustered CV/ diabetes-related mortality; any-cause hospitalisation; clustered CV/diabetes-related hospitalisation; and aggregated any-cause hospitalisation and mortality.

Data analyses

Cox proportional hazards regression models were fitted to estimate HRs and 95% CIs to assess the relationship between the calculated severity score and outcomes in the training dataset, with the inclusion of sociodemographic covariates. We developed both single event and multiple failure-time events models. The single event models were used to assess the risk for 1-year, 3-year, 5-year and 10-year for each of all-cause mortality and clustered CV/diabetes-related mortality. We experimented with different prediction horizons (1–10 years) to determine how the risks for adverse outcomes change over time after CHD diagnosis. A sensitivity analysis was conducted with the outcome being 1-year all-cause mortality excluding events in first 30 days, as these events may be related to the index event. Multiple failure-time events models were fitted, using the Breslow method to handle tied failures, for the risk of recurrent all-cause hospitalisations. Poisson regression models were used to estimate the unadjusted and adjusted incidence rate ratios (IRRs) and 95% CIs for the association between severity score in a given year and the number of all-cause hospital admissions in the following year annually for 1–10 years after index date. Competing risk analysis was conducted to estimate the subhazard ratio (SHR) and 95% CIs for the risk for 1-year any-cause hospitalisation and 1-year CV/diabetes-related hospitalisation while accounting for deaths as a competing event. Single event Cox models were used to assess the risk for the 1–10 year aggregated any-cause hospitalisation and mortality outcome. Likelihood ratio (LR) tests were fitted to assess the statistical significance of adding each of the developed severity scores (models 2, 3 or 4) to the demographics only model (model 1) in improving the models fit for predicting the outcomes. We also modelled the unlimited severity score divided into four categories: no severity (score=0) as a referent group; low severity (score=1–2); moderate severity (score=3–4); and high severity (score ≥5) to assess the strength of associations between the score and outcomes. Kaplan-Meier survivor function plots for hospitalisation, and mortality outcomes were fitted using severity score categories. All fitted models per outcome are summarised in online supplemental table S3.

The severity scores' calibration was tested using three methods: Somer's D[25]; comparing the survival curves for a given risk group[26]; and comparing the observed and predicted survival probabilities in prognostic groups derived by the severity score's cut points.[27] Poisson regression, multiple event regression and competing risk models' goodness of fit was assessed using the Akaike information criterion (AIC), where smaller AIC indicates a better fit of the data than larger AIC.[28] The predictive value of the single event survival models was assessed using Gönen and Heller's K concordance statistic (C-statistic), a measure of the area under a receiver operating characteristics (AUROC) curve for censored data.[29] C-statistic ranges between 0 and 1, where value close to 1 indicates an accurate model with high separation of subjects with different outcomes.[30 31] Hence, AUROCs are reported for all models except for the three aforementioned models where it was not possible to calculate them. Given the need to use two different postestimation measures (AIC and C-statistic), both were estimated and reported. The proportional hazards assumption was assessed using Schoenfeld residuals. All analyses were replicated in the 20% split of the data as a validation. Data were analysed using Stata software V.15.[32] The study is reported according to the RECORD checklist.

Patient and public involvement and engagement (PPIE)

We invited patients with CHD to a PPIE meeting. The participants agreed on the importance and the relevance of the study and suggested the need to raise the awareness about disease severity and to further highlight the fact that it involves several body organs and other conditions. Their perceptions about disease severity and indicators of increased disease severity varied between 'not thought about disease severity before' to listing a few indicators they considered relevant, such as declined physical function. The participants shared their views on approaches for disseminating the results via general practices and online social media outlets. We plan to disseminate the study findings widely to patient communities via local heart centres and general practices, and our social media platforms.

RESULTS

Overall, 213 088 patients with CHD were included (training dataset: n=1 70 395, validation dataset: n=42 693). Mean (±SD) age was 64.5±12.7 years; 46%

Table 1 Baseline characteristics of the included patients with coronary heart disease

Characteristic	Full cohort n=213 088	Training dataset n=170 395 (80%)	Validation dataset n=42 693 (20%)
Age (years), mean(±SD)	64.5 (±12.7)	64.5 (±12.7)	64.5 (±12.8)
Gender (female)	98 041 (46.0)	78 444 (46.0)	19 597 (46.0)
Number of general practices	398	398	395
Mean follow-up (years), mean (±SD)	9.4 (±6.0)	9.4 (±6.0)	9.4 (±6.0)
Ethnicity			
White	189 272 (88.8)	151 356 (88.83)	37 916 (88.81)
Black	2017 (0.95)	1626 (0.95)	391 (0.92)
Asian	4933 (2.32)	3940 (2.31)	993 (2.33)
Mixed	596 (0.28)	485 (0.28)	111 (0.26)
Other	1649 (0.77)	1299 (0.76)	350 (0.82)
Unknown	14 621 (6.86)	11 689 (6.86)	2932 (6.87)
Levels of social deprivation (IMD quintiles)			
Q1 (least deprived)	45 719 (21.5)	36 770 (21.6)	8949 (21.0)
Q2	49 251 (23.1)	39 472 (23.2)	9779 (22.9)
Q3	44 543 (20.9)	35 474 (20.8)	9069 (21.2)
Q4	39 032 (18.3)	31 141 (18.3)	7891 (18.5)
Q5 (most deprived)	34 412 (16.2)	27 435 (16.1)	6977 (16.3)
Unknown	131 (0.1)	103 (0.1)	28 (0.1)
Severity indicators at baseline			
Hypertension	109 455 (51.4)	87 422 (51.3)	22 033 (51.6)
Hyperlipidaemia	33 309 (15.6)	26 553 (15.6)	6756 (15.8)
Diabetes	22 763 (10.7)	18 079 (10.6)	4684 (11.0)
Proteinuria/albuminuria	4299 (2.0)	3401 (2.0)	898 (2.1)
End-stage renal disease (ESRD)	623 (0.3)	490 (0.3)	133 (0.3)
Peripheral vascular disease (PVD)	7220 (3.4)	5793 (3.4)	1427 (3.3)
Stable angina	30 838 (14.5)	24 667 (14.5)	6171 (14.5)
Cardiac arrest	1180 (0.6)	919 (0.5)	261 (0.6)
AF/SVT	17 810 (8.4)	14 270 (8.4)	3540 (8.3)
Myocardial infarction/ACS	38 451 (18.0)	30 715 (18.0)	7736 (18.1)
Heart valve disease	3587 (1.7)	2891 (1.7)	696 (1.6)
Endocarditis	292 (0.1)	235 (0.1)	57 (0.1)
Myocarditis	157 (0.1)	118 (0.1)	39 (0.1)
Cardiomyopathy	1105 (0.5)	886 (0.5)	219 (0.5)
Pericardial disease	682 (0.3)	534 (0.3)	148 (0.3)
Ventricular tachycardia/fibrillation	675 (0.3)	519 (0.3)	156 (0.4)
Cardiovascular procedures	10 270 (4.8)	8248 (4.8)	2022 (4.7)
TIA/stroke	18 783 (8.8)	15 053 (8.9)	3730 (8.7)
Pacemaker or defibrillator use	2460 (1.2)	1972 (1.2)	488 (1.1)
Congestive heart failure	10 888 (5.1)	8645 (5.1)	2243 (5.3)

All data are presented as count (%) unless otherwise stated.
ACS, acute coronary syndrome; AF/SVT, atrial fibrillation/supraventricular tachycardia; IMD, Index of multiple deprivation; TIA, transient ischaemic attack.

were women; 89% white; 16% from deprived areas (table 1). The ever before (unlimited) severity score ranged between 0 and 10 (mean±SD: 1.5±1.2), the 10-year before score between 0 and 10 (1.4±1.1) and the 5-year score between 0 and9 (1.2±1.0) (figure 1). The event rates show an increasing event rate with 1-unit

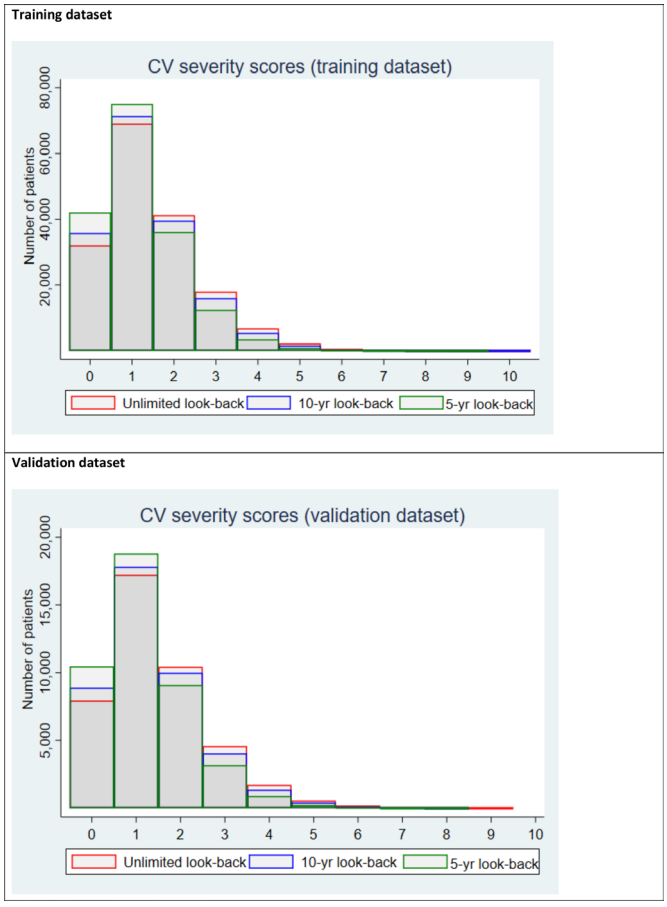


Figure 1 Distribution of baseline cardiovascular (CV) severity scores in the training and validation datasets.

increase in baseline unlimited severity score (online supplemental table S4).

All-cause mortality

Overall, 39 944 deaths occurred in 170 395 patients (23%), of which 1988 (1%) deaths occurred in the first year (of which, 544 events occurred within 30 days) and 24 130 (14%) by the 10th year after index. Higher levels of the severity score was positively associated with increasing risk for all-cause mortality (figure 2A,B). For each one-unit increase of the ever before (unlimited) severity score, the risks for both 1-year and 3-year all-cause mortality increased by 41% (1-year adjusted HR 1.41 (95% CI 1.37 to 1.45, AUROC=0.7912); 3-year HR: 1.41 (95% CI 1.39 to 1.43, AUROC=0.7882), 5 years by 39% (HR: 1.39, 95% CI 1.37 to 1.40, AUROC=0.7872) and 10 years by 35% (HR: 1.35, 95% CI 1.34 to 1.36, AUROC=0.7849). In comparison, the sociodemographics-only model (model 1) had AUROC of 0.7865 for 1-year all-cause mortality, indicating that adding the severity score slightly improved the models predictive value (LR test p<0.0001). The 1–10 year postindex scores showed similar results for risk of all-cause mortality (online supplemental table S5). The sensitivity analysis of excluding deaths in the first 30 days showed similar findings as the primary analysis.

CV/diabetes-related mortality

Each one unit increase of the unlimited severity score was associated with significantly higher risks at 1, 3, 5 and 10 years: HRs: 1.59 (95% CI 1.52 to 1.67, AUROC=0.8030); 1.61 (95% CI 1.56 to 1.65, AUROC=0.8041); 1.60 (95% CI 1.56 to 1.63, AUROC=0.8024); 1.57 (95% CI 1.55 to 1.60, AUROC=0.8010), respectively (table 2). For 1-year CV/diabetes-related mortality, adding the severity score improved the models predictive value (LR test p<0.0001) in comparison to model 1 (AUROC=0.7962). The 1–10 year postindex scores showed similar trends (online supplemental table S6).

All-cause hospitalisation

Overall, 138 510 (81% of patients) admissions occurred in 170 395 individuals, of which 43 023 (25%) and 127 358 (75%) occurred within 1 and 10 years after index, respectively. Higher severity showed a greater risk for future all-cause hospitalisation (figure 2E,F). Multiple failure analysis showed an increased risk for recurrent hospitalisation for one-unit increase in score (HR: 1.33, 95% CI 1.29 to 1.37) (online supplemental table S7) . For Poisson regression, unadjusted IRRs for the count of next year's hospitalisations ranged between 1.43 (95% CI 1.43 to 1.44) in 1 year after index to 1.37 (95% CI 1.35 to 1.40) in 10 years after index. When also adjusted for covariates, the model fit improved marginally with IRRs ranging between 1.39 (95% CI 1.39 to 1.40) and 1.37 (95% CI 1.34 to 1.40) for the same period (online supplemental table S8). The competing risks analysis showed each one-unit increase of the ever severity score was associated with 27% higher risks for 1-year any-cause hospitalisation (SHR: 1.27 (95% 1.26 to 1.28)) improving the predictive value provided by regression models only including sociodemographic variables – model 1 (LR test p<0.0001) (table 3).

CV/diabetes-related hospitalisation

Overall, 30 282 (18% of patients) events occurred within 1 year in 170 395 patients with CHD. For 1-year CV/diabetes-related admissions outcome, each one-unit increase in the ever before severity score was associated with SHR: 1.37 (95% CI 1.36 to 1.38) improving the predictive value provided by model 1 (LR test p<0.0001) (table 3), and it performed better than the any-cause admissions model.

Aggregated any-cause hospitalisation and mortality

Each one-unit increase in ever before severity score was associated with increased risks by 27% (26%–28%, AUROC=0.6271) at 1 year. Similar trends were observed at 3, 5 and 10 years after index (online supplemental table S9).

A summary of the estimated AIC and AUROC for fitted models is presented in table 4. For models where it was possible to estimate both AIC and AUROC, a summary is plotted in online supplemental figure S1, and there was a trend consistently showing improved model performance predicting cause-specific outcomes

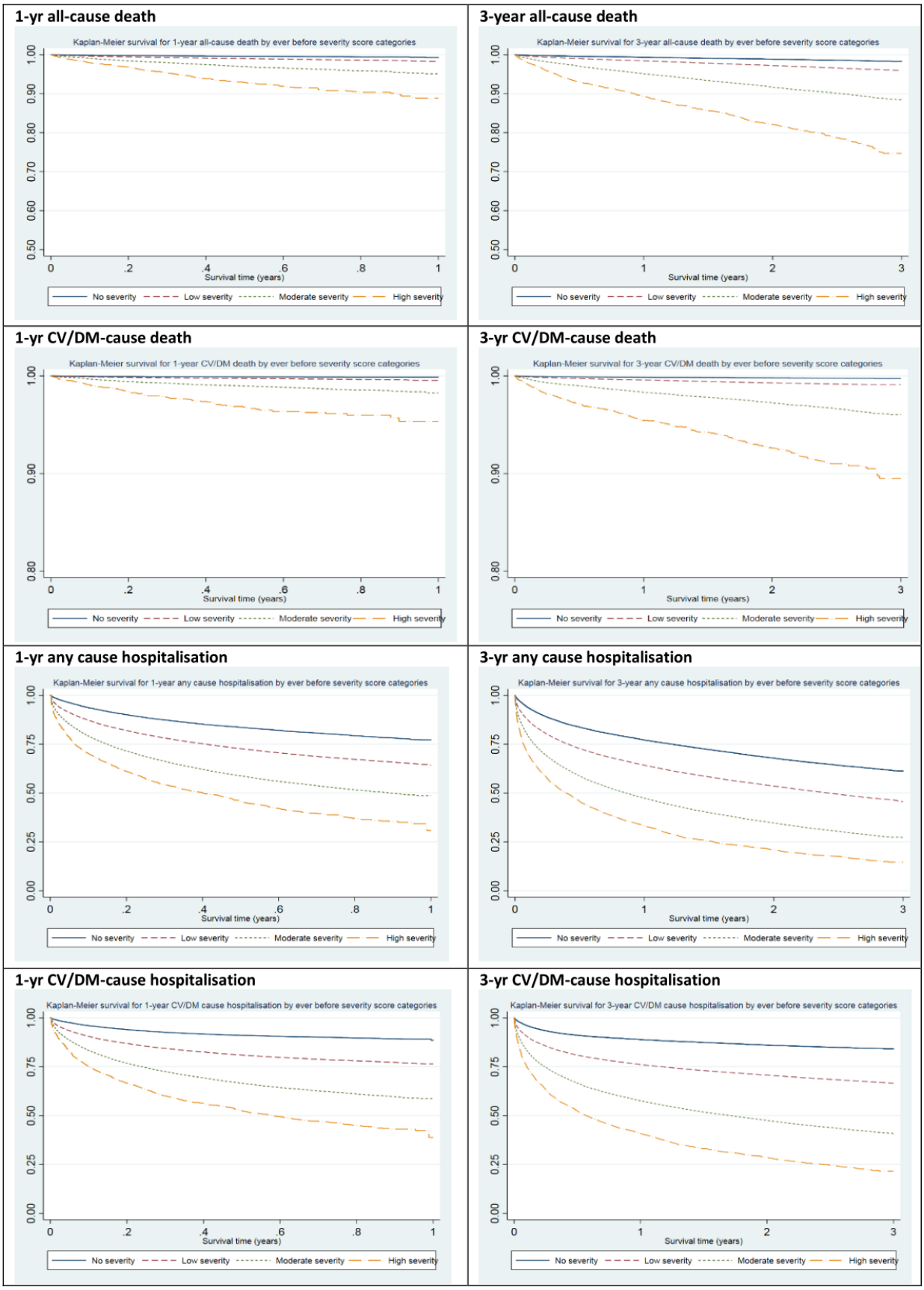


Figure 2 Kaplan-Meier survivor plots for adverse outcomes by CV severity score categories – training dataset. CV, cardiovascular; DM, diabetes mellitus. The survival probability scale (Y-axis) for 1-year and 3-year mortality was curtailed for improved differentiation of survival plots.

Table 2 Survival models for 1, 3, 5 and 10-year CV/diabetes-related mortality using baseline scores—HR (95% CI)—training dataset

	Predictor(s)	1 year	3 years	5 years	10 years
Model 1	Age	1.13 (1.12 to 1.14)	1.13 (1.13 to 1.14)	1.13 (1.13 to 1.13)	1.13 (1.12 to 1.13)
	Gender (F)	0.74 (0.63 to 0.87)	0.70 (0.64 to 0.77)	0.70 (0.65 to 0.75)	0.71 (0.68 to 0.75)
	IMD (vs least deprived)				
	Q5 (most deprived)	1.33 (1.02 to 1.73)	1.47 (1.27 to 1.71)	1.46 (1.30 to 1.64)	1.58 (1.46 to 1.72)
	Ethnicity (vs white)				
	Black	2.80 (1.44 to 5.43)	2.08 (1.35 to 3.21)	1.79 (1.24 to 2.58)	1.45 1.08 to 1.95)
	Asian	1.97 (1.16 to 3.35)	1.80 (1.31 to 2.48)	1.72 (1.33 to 2.22)	1.38 1.13 to 1.70)
	Mixed	–	1.81 (0.75 to 4.35)	1.57 (0.75 to 3.30)	0.81 0.39 to 1.70)
	Other	1.81 (0.81 to 4.05)	1.40 (0.83 to 2.37)	1.30 (0.85 to 2.00)	1.03 (0.73 to 1.45)
	Unknown	1.17 (0.83 to 1.66)	1.22 (1.01 to 1.48)	1.21 (1.04 to 1.41)	1.20 (1.07 to 1.33)
	AUROC	0.7962	0.7961	0.7935	0.7902
	AIC	13 382.17	42 896.17	69 852.12	1 38 058.2
Model 2	Ever before score	1.59 (1.52 to 1.67)	1.61 (1.56 to 1.65)	1.60 (1.56 to 1.63)	1.57 (1.55 to 1.60)
	Age	1.12 (1.11 to 1.13)	1.12 (1.11 to 1.12)	1.12 (1.11 to 1.12)	1.12 (1.11 to 1.12)
	Gender (F)	0.81 (0.69 to 0.95)	0.76 (0.70 to 0.83)	0.76 (0.71 to 0.82)	0.76 (0.72 to 0.80)
	IMD (vs least deprived)				
	Q5 (most deprived)	1.33 (1.02 to 1.73)	1.48 (1.28 to 1.72)	1.47 (1.31 to 1.65)	1.59 (1.46 to 1.72)
	Ethnicity (vs white)				
	Black	2.26 (1.16 to 4.39)	1.67 (1.08 to 2.57)	1.43 (0.99 to 2.07)	1.16 0.87 to 1.56)
	Asian	1.65 (0.97 to 2.80)	1.49 (1.09 to 2.06)	1.43 (1.11 to 1.84)	1.15 0.94 to 1.42)
	Mixed	–	1.81 (0.75 to 4.36)	1.57 (0.75 to 3.29)	0.79 0.37 to 1.65)
	Other	1.73 (0.77 to 3.87)	1.35 (0.80 to 2.29)	1.27 (0.83 to 1.96)	1.05 0.74 to 1.49)
	Unknown	1.49 (1.06 to 2.12)	1.59 (1.31 to 1.93)	1.54 (1.32 to 1.80)	1.48 (1.32 to 1.65)
	AUROC	0.8030	0.8041	0.8024	0.8010
	AIC	13 091.66	41 971.42	68 465.3	AIC=1 35 778.9
Model 3	10-year before score	1.59 (1.51 to 1.67)	1.59 (1.55 to 1.64)	1.58 (1.54 to 1.62)	1.55 (1.53 to 1.58)
	Age	1.12 (1.11 to 1.13)	1.12 (1.12 to 1.13)	1.12 (1.12 to 1.12)	1.12 (1.12 to 1.12)
	Gender (F)	0.81 (0.69 to 0.95)	0.75 (0.69 to 0.83)	0.75 (0.70 to 0.81)	0.75 (0.71 to 0.79)
	IMD (vs least deprived)				
	Q5 (most deprived)	1.32 (1.01 to 1.73)	1.47 (1.27 to 1.71)	1.46 (1.30 to 1.64)	1.57 (1.45 to 1.71)
	Ethnicity (vs white)				
	Black	2.13 (1.19 to 4.50)	1.72 (1.11 to 2.65)	1.48 (1.03 to 2.14)	1.20 1.89 to 1.65)
	Asian	1.74 (1.02 to 2.95)	1.58 (1.15 to 2.17)	1.51 (1.17 to 1.95)	1.20 0.98 to 1.47)
	Mixed	–	1.81 (0.75 to 4.36)	1.57 (0.75 to 3.30)	0.79 0.38 to 1.66)
	Other	1.73 (0.77 to 3.86)	1.38 (0.81 to 2.33)	1.28 (0.83 to 1.97)	1.05 0.74 to 1.48)
	Unknown	1.47 (1.04 to 2.08)	1.55 (1.28 to 1.88)	1.51 (1.29 to 1.75)	1.45 (1.30 to 1.61)
	AUROC	0.8032	0.8039	0.800	0.8000
	AIC	13 133.67	42 115.1	68 675.2	1 36 128.1

Continued

Table 2 Continued

	Predictor(s)	1 year	3 years	5 years	10 years
Model 4	5-year before score	1.55 (1.46 to 1.65)	1.56 (1.51 to 1.62)	1.56 (1.52 to 1.60)	1.52 (1.49 to 1.55)
	Age	1.12 (1.11 to 1.13)	1.13 (1.12 to 1.13)	1.12 (1.12 to 1.13)	1.12 (1.12 to 1.12)
	Gender (F)	0.79 (0.67 to 0.93)	0.74 (0.68 to 0.81)	0.74 (0.69 to 0.79)	0.74 (0.71 to 0.78)
	IMD (vs least deprived)				
	Q5 (most deprived)	1.32 (1.01 to 1.73)	1.48 (1.27 to 1.71)	1.46 (1.30 to 1.64)	1.58 (1.45 to 1.71)
	Ethnicity (vs White)				
	Black	2.33 (1.20 to 4.54)	1.78 (1.16 to 2.75)	1.57 (1.08 to 2.26)	1.27 (0.95 to 1.71)
	Asian	1.78 (1.05 to 3.03)	1.62 (1.18 to 2.22)	1.55 (1.20 to 2.00)	1.24 1.01 to 1.52)
	Mixed	–	1.83 (0.76 to 4.41)	1.58 (0.75 to 3.31)	0.79 0.38 to 1.66)
	Other	1.75 (0.78 to 3.92)	1.40 (0.82 to 2.36)	1.30 (0.85 to 2.00)	1.06 0.75 to 1.49)
	Unknown	1.41 (1.00 to 2.00)	1.48 (1.22 to 1.80)	1.45 (1.25 to 1.69)	1.40 (1.26 to 1.56)
	AUROC	0.8024	0.8030	0.8010	0.7983
	AIC	13 203.79	42 321.24	68 960.76	1 36 620.3

AIC, Akaike information criterion; AUROC, area under a receiver operating characteristics curve; CV, cardiovascular; DM, diabetes mellitus; IMD, Index of multiple deprivation; SHR, subhazard ratio.

over corresponding all-cause outcomes. When categorised, higher severity category levels were associated with increasing risks of hospitalisation and mortality (table 5 and figure 2). Severity score-only models using the training and validation dataset were also fitted and the AUROCs were up to 0.70 as summarised in online supplemental table S10. Models without IMD quintiles are summarised in online supplemental table S11. The performed calibration tests showed good calibration of the severity scores (online supplemental tables S12 and S13, figure 1 vs online supplemental figures S2 and S3). Testing for proportional hazards indicated the assumptions held true (online supplemental figures S4–S8). The validation dataset findings were all similar to those in the training dataset (online supplemental tables S14–S18 and figures S9–S12). The study methods and main findings are outlined in summary online supplemental figures S13 and S14.

DISCUSSION

Main findings

In this long-term retrospective cohort study, we present a contemporary and validated scoring system grading CVD severity in people with CHD. Our developed baseline and longitudinal severity scores provide important prognostic information for all-cause and cause-specific hospitalisation and mortality events in people with CHD that had marginal but statistically significant better predictive value in comparison with that provided by models only including sociodemographic variables. Each one-unit increase in disease severity was associated with elevated risks for all-cause mortality by 41%, CV/diabetes mortality by 59% and any cause hospitalisation by 27%.

Comparison with other studies

A few observational studies have assessed disease severity in people with CHD using routine primary care EHRs, while some studies used data derived from secondary or tertiary care settings to assess severity of CHD for various research questions.[19 33–35] However, the majority of prior studies assessing the severity of CHD were reporting risk scores based on the anatomical severity and characteristics of CAD, and they are used to assess the prognosis following revascularisation interventions, for example, SYNTAX and Gensini scores,[11 12] but do not provide information for the majority of patients with CHD not undergoing these interventions.

Some symptom-based tools were reportedly used to categorise disease severity.[36–38] However, the majority of people with CHD are asymptomatic, which may limit the application of such tools to the wider population of patients with CHD in clinical practice. Other studies classified CHD severity either by the CHD onset type (myocardial infarction, unstable or stable angina categories),[39–41] or the number of hospitalisation events.[42]

One observational study based on primary care data in Italy, estimated the positive predictive value for automated identification and severity assessment of four chronic conditions, including CHD.[10] The disease severity in 300 people with CHD was categorised into five levels based on the evidence of presence/absence of heart failure and coronary angioplasty. They reported a good agreement score (Cohen's kappa=0.69) between the automated algorithm and the general practitioner's assessment on the CHD severity level.

In our study, the new CV severity score was developed in a larger cohort, and it included heart failure and coronary procedures besides 18 additional severity indicators. Our score can be more applicable to a broader

Table 3 Competing risk analysis models for 1-year any-cause hospitalisation and 1-year CV/DM-related hospitalisation (competed by all-cause death) using baseline scores – training dataset

		1-year any-cause hospitalisation	1-year CV/DM hospitalisation
	Predictor(s)	**SHR (95% CI)**	**SHR (95% CI)**
Model 1	Age	1.02 (1.02 to 1.02)	1.02 (1.02 to 1.02)
	Gender (F)	0.68 (0.67 to 0.70)	0.57 (0.55 to 0.58)
	IMD (vs least deprived)		
	Q5 (most deprived)	1.07 (1.04 to 1.11)	1.05 (1.01 to 1.09)
	Ethnicity (vs white)		
	Black	0.93 (0.84 to 1.03)	0.74 (0.64 to 0.84)
	Asian	1.42 (1.34 to 1.50)	1.53 (1.43 to 1.63)
	Mixed	0.97 (0.81 to 1.17)	1.05 (0.85 to 1.30)
	Other	1.01 (0.91 to 1.13)	1.0 (0.87 to 1.14)
	Unknown	0.22 (0.21 to 0.24)	0.22 (0.20 to 0.24)
	AIC	995 523.4	704 241.7
Model 2	Ever before score	1.27 (1.26 to 1.28)	1.37 (1.36 to 1.38)
	Age	1.01 (1.01 to 1.01)	1.01 (1.01 to 1.01)
	Gender (F)	0.71 (0.70 to 0.73)	0.60 (0.59 to 0.62)
	IMD (vs least deprived)		
	Q5 (most deprived)	1.05 (1.02 to 1.08)	1.02 (0.98 to 1.05)
	Ethnicity (vs white)		
	Black	0.88 (0.79 to 0.97)	0.68 (0.57 to 0.78)
	Asian	1.33 (1.25 to 1.40)	1.39 (1.31 to 1.49)
	Mixed	0.95 (0.79 to 1.14)	1.03 (0.84 to 1.27)
	Other	1.00 (0.90 to 1.12)	0.96 (0.84 to 1.10)
	Unknown	0.24 (0.22 to 0.26)	0.24 (0.22 to 0.27)
	AIC	992 032.8	699 582.5
Model 3	10-year before score	1.28 (1.27 to 1.29)	1.39 (1.37 to 1.40)
	Age	1.01 (1.01 to 1.01)	1.01 (1.01 to 1.01)
	Gender (F)	0.72 (0.70 to 0.73)	0.60 (0.59 to 0.62)
	IMD (vs least deprived)		
	Q5 (most deprived)	1.05 (1.02 to 1.08)	1.01 (0.98 to 1.05)
	Ethnicity (vs white)		
	Black	0.88 (0.80 to 0.97)	0.68 (0.60 to 0.78)
	Asian	1.34 (1.26 to 1.42)	1.40 (1.32 to 1.51)
	Mixed	0.96 (0.80 to 1.15)	1.04 (0.84 to 1.28)
	Other	1.00 (0.90 to 1.12)	0.96 (0.84 to 1.10)
	Unknown	0.24 (0.22 to 0.26)	0.24 (0.22 to 0.27)
	AIC	992 096.3	699 635.1
Model 4	5-year before score	1.30 (1.29 to 1.31)	1.41 (1.40 to 1.43)
	Age	1.01 (1.01 to 1.01)	1.01 (1.01 to 1.01)
	Gender (F)	0.71 (0.70 to 0.73)	0.60 (0.59 to 0.62)
	IMD (vs least deprived)		
	Q5 (most deprived)	1.05 (1.02 to 1.08)	1.02 (0.98 to 1.05)
	Ethnicity (vs white)		
	Black	0.89 (0.80 to 0.98)	0.69 (0.60 to 0.79)
	Asian	1.36 (1.28 to 1.44)	1.44 (1.35 to 1.54)

Continued

Table 3 Continued

Predictor(s)	1-year any-cause hospitalisation SHR (95% CI)	1-year CV/DM hospitalisation SHR (95% CI)
Mixed	0.96 (0.80 to 1.16)	1.04 (0.85 to 1.29)
Other	1.00 (0.90 to 1.12)	0.97 (0.85 to 1.11)
Unknown	0.24 (0.22 to 0.26)	0.24 (0.22 to 0.27)
AIC	992 244.2	699 853.5

AIC, Akaike information criterion; CV, cardiovascular; DM, diabetes mellitus; IMD, Index of multiple deprivation; SHR, subhazard ratio.

population of patients with CHD than existing scores. Current scores either focus on small and highly selected groups of patients undergoing coronary procedures (eg, SYNTAX)[11] or define CHD severity in an overall simplistic approach by syndrome that does not take into account the close pathophysiological links between some of the included CV conditions thereby possibly reducing clinical relevance. In addition, we included clinically relevant severity indicators (such as diabetes which contributes to CHD severity[42]), and we evaluated the association between severity score and health outcomes. CV severity indicators may need revising in a few years as newer tests and measures become available in primary care setting. Therefore, future studies can include additional severity indicators, subject to their availability and well recording in primary care data, such as the coronary calcium score, ankle-brachial index test, B-type natriuretic peptide (BNP) or N-terminal pro BNP (NT-pro-BNP) levels and high-sensitivity C reactive protein (hs-CRP) levels. The inclusion of social deprivation data highlights the advantage of used EHRs driven from national healthcare systems, such as the NHS, as patients represent all social levels unlike what would be recorded from private medical systems. While social deprivation levels may not be directly compared with other populations, understanding the underlying domains and the allocation of patients into categories of least deprived versus most deprived may allow for a rough comparison with other populations as appropriate.

Potential benefits to clinical practice

People with CHD are mainly managed in primary care settings. Our severity measure is based on medical data routinely collected in general practice visits, which indicates its potential usefulness in risk stratification of people with CHD. The score calculation method can be first implemented as a simple table (online supplemental table S2) to enable clinicians estimate patient's CVD severity at baseline and over time. This can help identify people with CHD at a greater risk for adverse health outcomes, which informs advanced decision making. On a wider context, our algorithm is reproducible for other long-term conditions managed in primary care.

Table 4 Summary of AIC and AUROCs of fitted Cox and Poisson regression models – training dataset

Predictors/ model	1-year all-cause mortality	1-year CV/ diabetes-related mortality	Any-cause hospitalisation: 1-year hospitalisation (single event)*	Any-cause hospitalisation: Recurrent event	Any-cause hospitalisation: Poisson† (count in first year)	1-year CV/ diabetes-related hospitalisation*	1-year aggregated any hospitalisation or mortality
Demographics-only model							
Model 1	AUROC=0.7865 (AIC=43 050)	AUROC=0.7962 (AIC=13 382)	AIC=995 523	AIC=2.27e+07	AIC=477 990	AIC=704 241	AUROC=0.6055 (AIC=1 008 481)
Severity score+demographics models							
Model 2 (model 1+ever before severity score)	AUROC=0.7912 (AIC=42 586.52)	AUROC=0.8030 (AIC=13 091.66)	AIC=992 032.8	AIC=2.26e+07	AIC=456 825	AIC=699 582.5	AUROC=0.6271 (AIC=1 004 865)
Model 3 (model 1+10-year severity score)	AUROC=0.7912 (AIC=42 649.12)	AUROC=0.8032 (AIC=13 133)	AIC=992 096.3	AIC=2.26e+07	AIC=457 753	AIC=699 635	AUROC=0.6270 (AIC=1 004 939)
Model 4 (model 1 + 5-year severity score)	AUROC=0.7910 (AIC=42 732.13)	AUROC=0.8024 (AIC=13 203.79)	AIC=992 244.2	AIC=2.27e+07	AIC=460 109	AIC= 699 853.5	AUROC=0.6265 (AIC=1 005 101)

*Competing risk analysis.
†Adjusted for age, gender and IMD only.
AIC, Akaike information criterion; AUROC, area under a Receiver Operating Characteristics curve; CV, cardiovascular; DM, diabetes mellitus.

Table 5 Adjusted 1-year and 3-year HR or SHR (95% CI) for mortality and hospitalisation outcomes by the cardiovascular severity score category

	Training dataset				Validation dataset			
Outcome	No severity	Low severity	Moderate severity	High severity	No severity	Low severity	Moderate severity	High severity
1-year all-cause mortality	Reference	1.49 (1.22 to 1.81)	3.03 (2.46 to 3.72)	5.64 (4.44 to 7.15)	Reference	1.25 (0.87 to 1.80)	2.62 (1.80 to 3.82)	4.76 (3.06 to 7.40)
	AUROC=0.7899				**AUROC=0.7881**			
3-year all-cause mortality	Reference	1.40 (1.26 to 1.54)	2.90 (2.61 to 3.22)	5.34 (4.71 to 6.04)	Reference	1.20 (1.00 to 1.45)	2.17 (1.78 to 2.63)	4.63 (3.67 to 5.84)
	AUROC=0.7863				**AUROC=0.7857**			
1-year CV/diabetes mortality	Reference	1.58 (1.06 to 2.35)	4.09 (2.73 to 6.11)	9.38 (6.04 to 14.54)	Reference	1.12 (0.58 to 2.18)	3.32 (1.69 to 6.49)	6.53 (3.05 to 13.97)
	AUROC=0.8003				**AUROC=0.7951**			
3-year CV/diabetes mortality	Reference	1.95 (1.53 to 2.48)	5.46 (4.28 to 6.96)	11.75 (9.00 to 15.33)	Reference	1.14 (0.78 to 1.68)	2.96 (2.00 to 4.38)	8.53 (5.54 to 13.12)
	AUROC=0.8038				**AUROC=0.7953**			
1-year any hospitalisation*	Reference	1.62 (1.57 to 1.67)	2.47 (2.38 to 2.56)	3.47 (3.27 to 3.68)	Reference	1.65 (1.55 to 1.76)	2.51 (2.33 to 2.70)	3.31 (2.94 to 3.73)
	AIC=992 385.1				**AIC=2 19 586.1**			
3-year any hospitalisation*	Reference	1.41 (1.38 to 1.44)	2.10 (2.04 to 2.15)	2.85 (2.73 to 2.99)	Reference	1.41 (1.35 to 1.47)	2.05 (1.95 to 2.15)	2.72 (2.48 to 2.97)
	AIC=1 965 741				**AIC=4 33 770.1**			
1-year CV/diabetes hospitalisation*	Reference	2.11 (2.02 to 2.20)	3.71 (3.54 to 3.89)	5.51 (5.15 to 5.91)	Reference	2.15 (1.98 to 2.34)	3.73 (3.39 to 4.10)	5.26 (4.59 to 6.04)
	AIC=6 99 933.1				**AIC=1 55 656.5**			
3-year CV/diabetes hospitalisation *	Reference	2.03 (1.97 to 2.10)	3.79 (3.66 to 3.92)	5.63 (5.34 to 5.92)	Reference	2.04 (1.92 to 2.17)	3.73 (3.48 to 4.00)	5.45 (4.93 to 6.03)
	AIC=1 262 175				**AIC=280 208.9**			

Severity score categories: no severity: score=0 (referent category); low severity: score=1–2; moderate severity: score=3–4; high severity: score ≥5.
All results are adjusted for age, gender, IMD and ethnicity (model 2).
*SHR: subhazard ratio estimated for the competing risk models for risk for hospitalisation.
AIC, Akaike information criterion; AUROC, area under a Receiver Operating Characteristics curve; CV, cardiovascular; IMD, index of multiple deprivation.

Study strengths and limitations

The strengths of our study include: first, we analysed a large cohort of patients with CHD to develop and validate the severity scores, derived from a high-quality EHR database. Second, our models were based on baseline and longitudinal severity scores and included important sociodemographic variables, including social deprivation and ethnicity. Third, we compared the added predictive value of the developed score in comparison with that provided by models only including sociodemographic variables in all outcomes. In addition to all-cause mortality, which allows for a broad perspective of the burden of CHD, and hospital admissions, our measured outcomes also included CV and diabetes-related events. Fourth, we used longer term follow-up different from the available 30-day and 6-month risk scores. Finally, we invited people with CHD who provided their feedback on different aspects of the study.

Our study has several limitations. First, there is a risk of misclassifying the identified cases and severity indicators. However, the high validity of CVD diagnoses using CPRD data has been reported previously.[23 43] Second, other important severity indicators may have been missed since they are not available or routinely recorded in primary care, such as NT-pro-BNP levels or ankle-brachial index. However, using routinely available data allows the creation of a tool that can be applied to primary care and relevant research with EHRs. Third, by the nature of the cohort design, we missed non-survivors (people who died due to the first event). Fourth, as our validation was based on replicated analyses in a separate dataset (internal validation), future study is needed for external validation in an independent database before reporting the complete clinical utility and implications of our score. However, we observed very similar results when we compared our approach to two additional validation analyses based on postestimation from training dataset using CV mortality outcome.[26 44] Fifth, although the selection of binary weighting system is practical for replication of the score in clinical practice, future studies examining the risks of these indicators considering their different levels of severity, that is, as severity-weighted indicators are required. Finally, generalisability to other healthcare systems and/or other ethnic groups may be limited, but we believe a similar algorithm can be used in those circumstances given that the severity indicators are collected in routine primary care visits.

CONCLUSIONS

While CHD is associated with multiple morbidities and a leading cause of mortality worldwide,[3] severity measures for CHD based on primary care data are limited and needed. This study provides a contemporary measure of CVD severity derived by routine primary care EHRs for people with CHD, which showed high predictive value of hospitalisation and death outcomes. Our findings indicate that an increase in CVD severity in adult

people with CHD was associated with higher risks for all-cause and CV-specific hospital admissions and mortality outcomes. There is underused informative longitudinal, multimorbid structure in routine clinical records and our paper focus on the wider CV spectrum around CHD. Disease-specific severity tools have direct impact on clinical practice, by stratifying care according to disease severity, and can help inform service planning and risk stratification for precision medicine. Future research on external validation of the severity score is needed before reporting its complete clinical utility and implications.

Author affiliations
[1]NIHR School for Primary Care Research, Centre for Primary Care and Health Services Research, Manchester Academic Health Science Centre (MAHSC), The University of Manchester, Manchester, UK
[2]Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre (MAHSC), The University of Manchester, Manchester, UK
[3]Keele Cardiovascular Research Group, Centre for Prognosis Research, School of Primary, Community and Social Care, Keele University, Stoke-on-Trent, UK
[4]Division of Pharmacy and Optometry, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre (MAHSC), The University of Manchester, Manchester, UK
[5]NIHR Greater Manchester Patient Safety Translational Research Centre, The University of Manchester, Manchester, UK
[6]NIHR Manchester Biomedical Research Centre, Manchester Academic Health Science Centre (MAHSC), Manchester, UK
[7]Division of Diabetes, Endocrinology and Gastroenterology, School of Medical Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre (MAHSC), The University of Manchester, Manchester, UK
[8]Diabetes, Endocrinology and Metabolism Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre (MAHSC), Manchester, UK
[9]Division of Primary Care and Public Health, Brighton and Sussex Medical School, University of Brighton, Brighton, UK
[10]Centre for Academic Primary Care, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
[11]School of Primary, Community and Social Care, Faculty of Medicine and Health Sciences, Keele University, Staffordshire, UK
[12]Primary Care Stratified Medicine (PRISM) Research Group, Division of Primary Care, School of Medicine, University of Nottingham, Nottingham, UK
[13]Statistical Decision Sciences, Cardiovascular and Metabolism, Janssen Research and Development, High Wycombe, UK
[14]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
[15]Institute of Population Health Sciences, University of Liverpool, Liverpool, UK
[16]Division of Informatics, Imaging and Data Sciences, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre (MAHSC), The University of Manchester, Manchester, UK
[17]Centre for Biostatistics, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre (MAHSC), The University of Manchester, Manchester, UK

Twitter Mamas A Mamas @mmamas1973

Acknowledgements The authors would like to thank Mr Manoj Mistry and Mrs Angela Ruddock for their invaluable insight codeveloping and cofacilitating the patient and public involvement and engagement meeting and all participants for their time and important feedback. This study is based in part on data from the Clinical Practice Research Datalink (CPRD) obtained under licence from the UK Medicines and Healthcare products Regulatory Agency (MHRA). The study protocol was approved by the CPRD Independent Scientific Advisory Committee (ISAC). The data are provided by patients and collected by the NHS as part of their care and support. These data included ONS and HES data. The interpretation and conclusions contained in this study are those of the author/s alone. 'Copyright © (2018), re-used with the permission of The Health & Social Care Information Centre. All rights reserved'. Some of the findings were presented at the National Institute for Health Research School for Primary Care Research Showcase, November 2019, London, UK; and the Health Services Research UK Conference, July 2020 (Online).

Contributors EK, MM and SZ designed the study. SZ extracted and analysed the data and drafted the manuscript. MM, DR and EK critically revised the initial versions, and all authors contributed to interpretation of data and revised the paper for important intellectual content. All authors agreed on the final version of the paper before submission. SZ is the guarantor of the paper.

Funding This study is funded by the National Institute for Health Research (NIHR) School for Primary Care Research (NIHR SPCR – grant number 331).

Disclaimer The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The funder of the study had no role in study design, data collection, data analysis, data interpretation or writing of the report. The lead author had full access to the data in the study, takes responsibility for its integrity and the data analysis and had final responsibility for the decision to submit for publication.

Competing interests DMA reports research grants from Abbvie, Almirall, Celgene, Eli Lilly, Novartis, UCB and the Leo Foundation and is funded by the NIHR Greater Manchester Patient Safety Translational Research Centre, the NIHR School for Primary Care Research, and the NIHR Manchester Biomedical Research Centre. MKR has received educational grant support from MSD and Novo Nordisk; has modest stock ownership in GSK; and has consulted for Roche. CS reports grants from NIHR School for Primary Care Research during the conduct of the study, is partially supported by NHS ARC West and is an NIHR Senior Investigator. CDM is funded by the NIHR Collaborations for Leadership in Applied Health Research and Care West Midlands, the NIHR School for Primary Care Research and a NIHR Research Professorship in General Practice (NIHR-RP-2014-04-026). He has provided support for a Bristol Myer Squibb non-pharmacological Atrial Fibrillation study (funding to School). NQ reports grants from the NIHR SPCR and NIHR HTA, during the conduct of the study. SFW is an employee of Janssen R NIHR Senior Investigator grant. NP's time was partially funded by the NIHR Manchester Biomedical Research Centre. Other coauthors have no disclosures.

Patient consent for publication Not required.

Ethics approval The study was approved by the Independent Scientific Advisory Committee (ISAC) for MHRA Database Research (protocol number: 17_168). Generic ethical approval for observational research using CPRD with approval from ISAC has been granted by a Health Research Authority (HRA) Research Ethics Committee (East Midlands—Derby, REC reference number 05/MRE04/87).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Clinical code lists are available from clinicalcodes.org. Electronic health records are, by definition, considered sensitive data in the UK by the Data Protection Act and cannot be shared via public deposition because of information governance restriction in place to protect patient confidentiality. Access to data is available only once approval has been obtained through the individual constituent entities controlling access to the data. The primary care data can be requested via application to the Clinical Practice Research Datalink, secondary care data can be requested via application to the Hospital Episode Statistics from the UK Health and Social Care Information Centre, and mortality data are available by application to the UK Office for National Statistics.

ORCID iDs
Salwa S Zghebi http://orcid.org/0000-0002-7978-1094
Mamas A Mamas http://orcid.org/0000-0001-9241-8890
Nadeem Qureshi http://orcid.org/0000-0003-4909-0644
Evangelos Kontopantelis http://orcid.org/0000-0001-6450-5815

REFERENCES

1 World Health Organization (WHO) media centre. Cardiovascular diseases (CVDs). Secondary Cardiovascular diseases (CVDs) [Fact sheet], 2017. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

2 British Heart Foundation. BHF UK Factsheet November 2018. Secondary BHF UK Factsheet, 2018. Available: https://www.bhf.org.uk/what-we-do/our-research/heart-statistics
3 Sanchis-Gomar F, Perez-Quilis C, Leischik R, *et al*. Epidemiology of coronary heart disease and acute coronary syndrome. *Ann Transl Med* 2016;4:256.
4 Bhatnagar P, Wickramasinghe K, Wilkins E, *et al*. Trends in the epidemiology of cardiovascular disease in the UK. *Heart* 2016;102:1945–52.
5 Ginghina C, Bejan I, Ceck CD. Modern risk stratification in coronary heart disease. *J Med Life* 2011;4:377–86.
6 Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ* 2017;357:j2099.
7 Wilson PW, D'Agostino RB, Levy D, *et al*. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97:1837–47.
8 Zghebi SS, Rutter MK, Ashcroft DM, *et al*. Using electronic health records to quantify and stratify the severity of type 2 diabetes in primary care in England: rationale and cohort study design. *BMJ Open* 2018;8:e020926.
9 Gambert S. Disease Severity. In: Gellman MD, Turner JR, eds. *Encyclopedia of behavioral medicine*. New York, NY: Springer, 2013.
10 Gini R, Schuemie MJ, Mazzaglia G, *et al*. Automatic identification of type 2 diabetes, hypertension, ischaemic heart disease, heart failure and their levels of severity from Italian general practitioners' electronic medical records: a validation study. *BMJ Open* 2016;6:e012413.
11 Syntax Score Working Group. The SYNTAX score. secondary the SYNTAX score, 2016. Available: http://www.syntaxscore.com/
12 Gensini GG. A more meaningful scoring system for determining the severity of coronary heart disease. *Am J Cardiol* 1983;51:606.
13 Vivekananthan DP, Blackstone EH, Pothier CE, *et al*. Heart rate recovery after exercise is a predictor of mortality, independent of the angiographic severity of coronary disease. *J Am Coll Cardiol* 2003;42:831–8.
14 Cury RC, Abbara S, Achenbach S, *et al*. Coronary Artery Disease - Reporting and Data System (CAD-RADS): An Expert Consensus Document of SCCT, ACR and NASCI: Endorsed by the ACC. *JACC Cardiovasc Imaging* 2016;9:1099–113.
15 Smith LR, Harrell FE, Rankin JS, *et al*. Determinants of early versus late cardiac death in patients undergoing coronary artery bypass graft surgery. *Circulation* 1991;84:245–53.
16 Helmy H, Abdel-Galeel A, Taha Kishk Y, *et al*. Correlation of corrected QT dispersion with the severity of coronary artery disease detected by SYNTAX score in non-diabetic patients with STEMI. *Egypt Heart J* 2017;69:111–7.
17 Langroudi TF, Haybar H, Parsa SA, *et al*. The severity of coronary artery disease was not associated with non-alcoholic fatty liver disease in a series of 264 non-diabetic patients who underwent coronary angiography. *Rom J Intern Med* 2018;56:167–72.
18 Bouzidi N, Sahli S, Khlifi L. Lipid profile variables and prediction of the severity of coronary artery disease in Tunisian type 2 diabetic patients. *Int J Pharm Sci Rev Res* 2017.
19 Srinivasan MP, Kamath PK, Bhat NM, *et al*. Manipal diabetes coronary artery severity score. *Diabetes Metab Syndr* 2017;11 Suppl 1:S33–7.
20 Herrett E, Gallagher AM, Bhaskaran K, *et al*. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.
21 Kontopantelis E, Stevens RJ, Helms PJ, *et al*. Spatial distribution of clinical computer systems in primary care in England in 2016 and implications for primary care electronic medical record databases: a cross-sectional population study. *BMJ Open* 2018;8:e020738.
22 Department for Communities and Local Government. The English indices of deprivation 2015. Secondary the English indices of deprivation, 2015. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/465791/English_Indices_of_Deprivation_2015_-_Statistical_Release.pdf
23 Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the general practice research database: a systematic review. *Br J Gen Pract* 2010;60:e128–36.
24 Olier I, Springate DA, Ashcroft DM, *et al*. Modelling conditions and health care processes in electronic health records: an application to severe mental illness with the clinical practice research Datalink. *PLoS One* 2016;11:e0146715.
25 Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics* 2005;61:92–105.
26 Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol* 2013;13:33.
27 Royston P. Tools for checking calibration of a Cox model in external validation: prediction of population-averaged survival curves based on risk groups. *Stata J* 2015;15:275–91.
28 StataCorp LLC. Stata Base Reference Manual - Release 15. Secondary Stata Base Reference Manual - Release 15 2017, 2017. Available: https://www.stata.com/manuals/r.pdf
29 Royston P, Lambert PC. Prognostic models. In: *Flexible parametric survival analysis using Stata: beyond the Cox model*. Texas, USA: Stata Press, 2011: 152–3.
30 Caetano SJ, Sonpavde G, Pond GR. C-statistic: A brief explanation of its construction, interpretation and limitations. *Eur J Cancer* 2018;90:130–2.
31 StataCorp. Stata 15 survival analysis reference manual. Secondary Stata 15 survival analysis reference manual, 2017. Available: https://www.stata.com/manuals/st.pdf
32 StataCorp LLC. Stata statistical software. Secondary stata statistical software, 2018. Available: https://www.stata.com/
33 Srinivasan M, Bhat N, Kamath P, *et al*. Risk factors for complex and severe coronary artery disease in type 2 diabetes mellitus. *J Cardiovasc Dis Res* 2017;8:19–23.
34 Özcan C, Deleskog A, Schjerning Olsen A-M, *et al*. Coronary artery disease severity and long-term cardiovascular risk in patients with myocardial infarction: a Danish nationwide register-based cohort study. *Eur Heart J Cardiovasc Pharmacother* 2018;4:25–35.
35 Shenoy R, Bhat R, Srinivasan M. Correlation between insulin resistance and severity of coronary artery disease in non-diabetes. *Asian J Pharm Clin Res* 2016;9:331–3.
36 Mancini GBJ, Boden WE, Brooks MM, *et al*. Impact of treatment strategies on outcomes in patients with stable coronary artery disease and type 2 diabetes mellitus according to presenting angina severity: a pooled analysis of three federally-funded randomized trials. *Atherosclerosis* 2018;277:186–94.
37 Schopfer DW, Regan M, Heidenreich PA, *et al*. Depressive symptoms, cardiac disease severity, and functional status in patients with coronary artery disease (from the heart and soul study). *Am J Cardiol* 2016;118:1287–92.
38 Sudzinova A, Nagyova I, Studencan M, *et al*. Roma coronary heart disease patients have more medical risk factors and greater severity of coronary heart disease than non-Roma. *Int J Public Health* 2013;58:409–15.
39 Liu C, Yang Y, Peng D, *et al*. Hyperhomocysteinemia as a metabolic disorder parameter is independently associated with the severity of coronary heart disease. *Saudi Med J* 2015;36:839–46.
40 Katsouras CS, Karabina SA, Tambaki AP, *et al*. Serum lipoprotein(a) concentrations and apolipoprotein(a) isoforms: association with the severity of clinical presentation in patients with coronary heart disease. *J Cardiovasc Risk* 2001;8:311–7.
41 Chotchaeva FR, Balatskiy AV, Samokhodskaya LM. Association between T-cadherin gene (CDH13) variants and severity of coronary heart disease manifestation. *Int J Clin Exp Med* 2016;9:4059–8353.
42 Javidi D, Gharaei B, Fateh S. The contribution of diabetes mellitus per se to the severity of coronary artery disease. *Cardiovasc J S Afr* 2005;16:246–8.
43 Maru S, Koch GG, Stender M, *et al*. Antidiabetic drugs and heart failure risk in patients with type 2 diabetes in the U.K. primary care setting. *Diabetes Care* 2005;28:20–6.
44 Newson RB. Comparing the Predictive Powers of Survival Models Using Harrell's C or Somers' D. *Stata J* 2010;10:339–58.

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

SUPPLEMENTAL MATERIAL

Zghebi SS, *et al. Open Heart* 2021; 8:e001498. doi: 10.1136/openhrt-2020-001498

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- *Zghebi et al. 2020*

Table S1 inclusion codes for CHD

	Medical code	Read code	Description	Medical code	Read code	Description	Medical code	Read code	Description
1.	240	G3...00	Ischaemic heart disease	29902	G330z00	Angina decubitus NOS	102943	8Hkl.00	Referral to cardiac rehabilitation service by secondary care
2.	241	G30..00	Acute myocardial infarction	30171	G5...00	Other forms of heart disease	103046	G210z00	Malignant hypertensive heart disease NOS
3.	1021	5543.00	Coronary arteriograph.abnormal	30330	G309.00	Acute Q-wave infarct	103655	187..00	Frequency of angina
4.	1204	G30..14	Heart attack	30421	G30..13	Cardiac rupture following myocardial infarction (MI)	103932	8CMP.00	Coronary heart disease care plan
5.	1344	G340.12	Coronary artery disease	30963	1J61.00	Suspected ischaemic heart disease	104675	8F97.00	Cardiac rehabilitation programme completed
6.	1414	G33z300	Angina on effort	31464	G21z.00	Hypertensive heart disease NOS	105216	14AW.00	H/O acute coronary syndrome
7.	1430	G33..00	Angina pectoris	32272	G38..00	Postoperative myocardial infarction	105250	G341111	Mural cardiac aneurysm
8.	1431	G311.13	Unstable angina	32450	G33z400	Ischaemic chest pain	105479	G39..00	Coronary microvascular disease
9.	1490	G5z..00	Heart disease NOS	32526	14AA.00	H/O: heart disease NOS	105615	G01yz00	Other acute rheumatic heart disease NOS
10.	1537	662..11	Heart disease monitoring	32854	G30B.00	Acute posterolateral myocardial infarction	105938	G211z00	Benign hypertensive heart disease NOS
11.	1655	G340.11	Triple vessel disease of the heart	34207	9Ob4.00	Coronary heart disease monitoring 2nd letter	106812	G383.00	Postoperative transmural myocardial infarction unspec site
12.	1676	G3z..00	Ischaemic heart disease NOS	34328	G311300	Refractory angina	107574	8T04.00	Referral to Angina Plan self-management programme
13.	1677	G30..15	MI - acute myocardial infarction	34329	9Ob5.00	Coronary heart disease monitoring 3rd letter	107967	661M000	Angina self-management plan agreed
14.	1678	G308.00	Inferior myocardial infarction NOS	34633	G34y.00	Other specified chronic ischaemic heart disease	108056	8IEY.00	Referral to Angina Plan self-management programme declined
15.	1792	G3...13	IHD - Ischaemic heart disease	34803	G30y.00	Other acute myocardial infarction	109035	Gyu3500	[X]Subsequent myocardial infarction of other sites
16.	1811	G5yz.00	Other heart disease NOS	34952	32B..00	ECG: Q wave	110535	8F98.00	Cardiac rehabilitation programme offered
17.	2155	G341000	Ventricular cardiac aneurysm	35119	G501.00	Post infarction pericarditis	110634	Gyu4000	[X]Other specified pulmonary heart diseases
18.	2491	G30..12	Coronary thrombosis	35277	9Ob1.00	Refuses coronary heart disease monitoring	25583	G574011	Cardiac arrest-ventricular fibrillation
19.	3468	662..00	Cardiac disease monitoring	35287	322Z.00	ECG: myocardial ischaemia NOS	25814	9Ob3.00	Coronary heart disease monitoring 1st letter

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

	Medical code	Read code	Description	Medical code	Read code	Description	Medical code	Read code	Description
20.	3704	G307.00	Acute subendocardial infarction	35373	9Ob0.00	Attends coronary heart disease monitoring	25842	G33z.00	Angina pectoris NOS
21.	3999	G340000	Single coronary vessel disease	35382	5533.00	Angiocardiography abnormal	26863	G33z600	New onset angina
22.	4017	G32..00	Old myocardial infarction	35674	14A3.00	H/O: myocardial infarct <60	26972	3234.00	ECG:posterior/inferior infarct
23.	4656	G311.11	Crescendo angina	35713	G34yz00	Other specified chronic ischaemic heart disease NOS	26973	3222.00	ECG:shows myocardial ischaemia
24.	5221	44H3.00	Cardiac enzymes abnormal	36193	G5y..00	Other specified heart disease	26975	3233.00	ECG: antero-septal infarct.
25.	5254	G340100	Double coronary vessel disease	36423	G36..00	Certain current complication follow acute myocardial infarct	27484	G341.11	Cardiac aneurysm
26.	5387	G301.00	Other specified anterior myocardial infarction	36523	G311.00	Preinfarction syndrome	27951	G31..00	Other acute and subacute ischaemic heart disease
27.	5413	G340.00	Coronary atherosclerosis	36609	G342.00	Atherosclerotic cardiovascular disease	27977	G31yz00	Other acute and subacute ischaemic heart disease NOS
28.	6331	G341.00	Aneurysm of heart	36854	G332.00	Coronary artery spasm	28138	G34..00	Other chronic ischaemic heart disease
29.	6336	14A5.00	H/O: angina pectoris	37657	G362.00	Ventric septal defect/curr comp fol acut myocardal infarctn	28554	G33zz00	Angina pectoris NOS
30.	7320	G343.00	Ischaemic cardiomyopathy	37990	8F90.00	Cardiac rehabilitation - phase 1	28736	G30y000	Acute atrial infarction
31.	7347	G311100	Unstable angina	37991	8F91.00	Cardiac rehabilitation - phase 2	29300	662K300	Angina control - worsening
32.	7696	G33z200	Syncope anginosa	38609	G351.00	Subsequent myocardial infarction of inferior wall	29421	G344.00	Silent myocardial ischaemia
33.	7783	323..00	ECG: myocardial infarction	39449	G312.00	Coronary thrombosis not resulting in myocardial infarction	29553	G366.00	Thrombosis atrium,auric append&vent/curr comp foll acute MI
34.	8246	322..00	ECG: myocardial ischaemia	39500	9Ob8.00	Coronary heart disease monitoring check done	29643	G303.00	Acute inferoposterior infarction
35.	8516	8F9..00	Cardiac rehabilitation	39546	Gyu3000	[X]Other forms of angina pectoris	29758	G30X.00	Acute transmural myocardial infarction of unspecif site
36.	8568	G37..00	Cardiac syndrome X	39655	G311.12	Impending infarction	37908	9Ob6.00	Coronary heart disease monitoring verbal invitation
37.	8935	G302.00	Acute inferolateral infarction	39693	G31y200	Subendocardial ischaemia	102914	8IE3.00	Referral to cardiac rehabilitation programme declined
38.	9276	G31y000	Acute coronary insufficiency	39904	3232.00	ECG: old myocardial infarction	66388	G33z000	Status anginosus
39.	9413	G31y.00	Other acute and subacute ischaemic heart disease	40399	14A4.00	H/O: myocardial infarct >60	67087	G341100	Other cardiac wall aneurysm
40.	9507	G307000	Acute non-Q wave infarction	40429	G301000	Acute anteroapical infarction	68357	G31y100	Microinfarction of heart
41.	9555	G33z500	Post infarct angina	40624	ZL62200	Referral to cardiac rehabilitation nurse	68401	Gyu3200	[X]Other forms of acute ischaemic heart disease
42.	10109	G....13	Heart diseases	41032	8F92.00	Cardiac rehabilitation - phase 3	68748	G38z.00	Postoperative myocardial infarction, unspecified

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- *Zghebi et al. 2020*

	Medical code	Read code	Description	Medical code	Read code	Description	Medical code	Read code	Description
43.	10127	8H7v.00	Referral to cardiac rehabilitation nurse	41179	G5yyz00	Other ill-defined heart disease NOS	68849	G01z.00	Acute rheumatic heart disease NOS
44.	10260	6A4..00	Coronary heart disease review	41221	G30y200	Acute septal infarction	68979	Gyu5.00	[X]Other forms of heart disease
45.	10562	G307100	Acute non-ST segment elevation myocardial infarction	41677	G341z00	Aneurysm of heart NOS	69474	G365.00	Rupture papillary muscle/curr comp fol acute myocard infarct
46.	11048	G331.11	Variant angina pectoris	41835	G384.00	Postoperative subendocardial myocardial infarction	69776	SP00300	Mechanical complication of coronary bypass
47.	11648	8B3k.00	Coronary heart disease medication review	42104	32E4.00	ECG: S-T depression	70160	9Ob9.00	Coronary heart disease monitoring telephone invite
48.	11983	G311500	Acute coronary syndrome	45476	14AL.00	H/O: Treatment for ischaemic heart disease	71046	G41yz00	Other chronic pulmonary heart disease NOS
49.	12139	G300.00	Acute anterolateral infarction	45809	G350.00	Subsequent myocardial infarction of anterior wall	72562	G353.00	Subsequent myocardial infarction of other sites
50.	12229	G30X000	Acute ST segment elevation myocardial infarction	45960	8B27.00	Antianginal therapy	91774	G341300	Acquired atrioventricular fistula of heart
51.	12804	G33z700	Stable angina	46017	G30yz00	Other acute myocardial infarction NOS	95550	8H2V.00	Admit ischaemic heart disease emergency
52.	12986	G331.00	Prinzmetal's angina	13250	G....12	Cardiac diseases	97001	44p2.00	Cardiac troponin positive
53.	13185	662K.00	Angina control	46166	G35X.00	Subsequent myocardial infarction of unspecified site	99991	Gyu3600	[X]Subsequent myocardial infarction of unspecified site
54.	13187	662N.00	CHD monitoring	46227	32B2.00	ECG: Q wave abnormal	100139	14AT.00	History of myocardial infarction
55.	46112	G380.00	Postoperative transmural myocardial infarction anterior wall	46276	G381.00	Postoperative transmural myocardial infarction inferior wall	101164	8LF..00	Coronary angiography planned
56.	13566	G30..11	Attack - heart	46565	8F93.00	Cardiac rehabilitation - phase 4	101373	8L41.00	Coronary angioplasty planned
57.	13571	G30..16	Thrombosis - coronary	47637	Gyu3300	[X]Other forms of chronic ischaemic heart disease	102447	8Hkk.00	Referral to cardiac rehabilitation programme
58.	14658	G30z.00	Acute myocardial infarction NOS	47798	9Ob2.00	Coronary heart disease monitoring default	19067	ZL22200	Under care of cardiac rehabilitation nurse
59.	14782	662K200	Angina control - improving	48981	66f1.00	Cardiovascular disease interim monitoring	19185	66f..00	Cardiovascular disease monitoring
60.	14897	G301z00	Anterior myocardial infarction NOS	50372	14AH.00	H/O: Myocardial infarction in last year	19250	8I3a.00	Cardiac rehabilitation declined
61.	14898	G305.00	Lateral myocardial infarction NOS	51043	ZRBN.00	Duke's coronary artery disease score	19542	662K000	Angina control - good
62.	15349	662Kz00	Angina control NOS	52517	Gyu3.00	[X]Ischaemic heart diseases	19655	G311.14	Angina at rest
63.	15373	662K100	Angina control - poor	52705	3236.00	ECG: lateral infarction	19744	8I37.00	Coronary heart disease monitoring refused
64.	15661	G310.11	Dressler's syndrome	54251	G311z00	Preinfarction syndrome NOS	19827	3213111	Positive exercise ECG test

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

	Medical code	Read code	Description	Medical code	Read code	Description	Medical code	Read code	Description
65.	15754	G34z.00	Other chronic ischaemic heart disease NOS	54535	G33z100	Stenocardia	20001	G1z..00	Chronic rheumatic heart disease NOS
66.	15782	G41z.00	Chronic pulmonary heart disease NOS	55137	G311011	MI - myocardial infarction aborted	20095	G330.00	Angina decubitus
67.	16173	G21zz00	Hypertensive heart disease NOS	55401	3235.00	ECG: subendocardial infarct	20416	G3...12	Atherosclerotic heart disease
68.	16408	G32..11	Healed myocardial infarction	57062	14AJ.00	H/O: Angina in last year	21844	G31y300	Transient myocardial ischaemia
69.	16657	ZV7B011	[V]Screening for ischaemic heart disease (IHD)	59032	323Z.00	ECG: myocardial infarct NOS	22383	G3y..00	Other specified ischaemic heart disease
70.	17133	G30A.00	Mural thrombosis	59189	G363.00	Ruptur cardiac wall w'out haemopericard/cur comp fol ac MI	23078	G34y100	Chronic myocardial ischaemia
71.	17307	G311200	Angina at rest	59193	G341200	Aneurysm of coronary vessels	23098	ZV57900	[V]Cardiac rehabilitation
72.	17464	G32..12	Personal history of myocardial infarction	59854	G1yzz00	Other rheumatic heart disease NOS	23579	G310.00	Postmyocardial infarction syndrome
73.	17681	662Z.00	Cardiac disease monitoring NOS	59940	G364.00	Ruptur chordae tendinae/curr comp fol acute myocard infarct	23708	G361.00	Atrial septal defect/curr comp folow acut myocardial infarct
74.	17689	G30..17	Silent myocardial infarction	60664	44H3000	Cardiac enzymes abnormal - first set	23892	G304.00	Posterior myocardial infarction NOS
75.	17872	G301100	Acute anteroseptal infarction	61166	G21z000	Hypertensive heart disease NOS without CCF	24126	G360.00	Haemopericardium/current comp folow acut myocard infarct
76.	18118	G311400	Worsening angina	61670	889A.00	Diab mellit insulin-glucose infus acute myocardial infarct	24540	G34y000	Chronic coronary insufficiency
77.	18125	G330000	Nocturnal angina	62270	32B3.00	ECG: Q wave pathological	24783	G3...11	Arteriosclerotic heart disease
78.	18134	182A.00	Chest pain on exertion	62626	G30y100	Acute papillary muscle infarction	19044	ZLA2200	Seen by cardiac rehabilitation nurse
79.	18135	6A2..00	Coronary heart disease annual review	62718	G21z100	Hypertensive heart disease NOS with CCF	66285	32BZ.00	ECG: Q wave NOS
80.	18150	9Ob..00	Coronary heart disease monitoring administration	63467	G306.00	True posterior myocardial infarction	18889	G34z000	Asymptomatic coronary heart disease
81.	18218	Z677.00	Cardiac rehabilitation class	63538	32B1.00	ECG: Q wave normal	65533	G40z.00	Acute pulmonary heart disease NOS
82.	18842	G35..00	Subsequent myocardial infarction						

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- *Zghebi et al. 2020*

Table S2 Simple illustration of pre-index and post-index CVD score calculation in people with CHD

Patient ID	CHD diagnosis	Diagnosis of severity indicator	Pre-index severity indicator eligibility at			Post-index severity indicator eligibility at post-index (e.g. 3-year post-index) +		
			5-yr look-back window	10- yr look-back window	Unlimited look-back window	5-yr look-back window#	10- yr look-back window^	Unlimited look-back window&
0001	11/07/2007	PVD on 13/10/1992	N	N	**Y**	N	N	**Y**
		TIA on 01/03/1994	N	N	**Y**	N	N	**Y**
		Stroke on 22/08/2001	N	**Y**	**Y**	N	**Y**	**Y**
		DM on 06/12/2005	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
		MI on 18/03/2010	-	-	-	**Y**	**Y**	**Y**
	Score / window		**1**	**2**	**4**	**2**	**3**	**5**
0002	19/08/2011	CABG on 28/01/2014	-	-	-	**Y**	**Y**	**Y**
	Score / window		**0**	**0**	**0**	**1**	**1**	**1**
0003	03/02/2015	HT on 27/05/2002	N	N	**Y**	N	N	**Y**
		ESRD on 14/12/2016	-	-	-	**Y**	**Y**	**Y**
	Score / window		**0**	**0**	**1**	**1**	**1**	**2**

These data are based on fictional patient IDs and presented for illustrative purpose only.
this covers 3 years after index and 2 years before index i.e. total of 5 years window.
^ this covers 3 years after index and 7 years before index i.e. total of 10 years window.
& this covers 3 years after index and unlimited window before index.

CABG: coronary artery bypass graft; **DM**: diabetes mellitus; **ESRD**: end stage renal disease; **HT**: hypertension; **MI**: myocardial infarction; **PVD**: peripheral vascular disease; **TIA**: transient ischaemic attack.

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

So, we computed a total of 33 pre-index and post-index severity scores using different study windows for each patient:

- Three pre-index severity scores for computing overall severity using the three look-back windows:
  - Unlimited look-back window
  - 10-year look-back window
  - 5-year look-back window

- 30 post-index severity scores, based on combining each of windows 1-3 above with post-index windows of length of 1-10 years. For example, for the 1-year post-index window:
  - 1-year post-index window combined with unlimited look-back window
  - 1-year post-index window combined with 10-year look-back window
  - 1-year post-index window combined with 5-year look-back window
  - And so on for the 2-10 years post-index windows.

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Table S3 A total of 212 Cox and Poisson models were fitted for primary and secondary outcomes.

		All-cause death	Clustered CV- or Diabetes-related death	Any cause hospitalisation	Clustered CV/Diabetes-hospitalisation	Aggregated any hospitalisation or mortality	Recurrent hospitalisation (multiple event)
Baseline CVD scores (Cox models)							
	Model 1 (Age, gender, ethnicity, and IMD)	√ √ √ √	√ √ √ √	√ √ √ √	√ √ √ √	√ √ √ √	√
Model 2	Model 1 + Unlimited score	√ √ √ √	√ √ √ √	√ √ √ √	√ √ √ √	√ √ √ √	√
	Model 1 + 10-year before score	√ √ √ √	√ √ √ √	√ √ √ √	√ √ √ √	√ √ √ √	√
	Model 1 + 5-year before score	√ √ √ √	√ √ √ √	√ √ √ √	√ √ √ √	√ √ √ √	√
Baseline CVD score categories (Cox models)		√ √	√ √	√ √	√ √	-	-
Moving CVD scores (Cox models)							
	Model 1 (Age, gender, ethnicity, and IMD) between index+1year to index+10 years	√ √ √ √ √ √ √ √ √ √	√ √ √ √ √ √ √ √ √ √	-	-	-	-
Model 2	Model 1 + Unlimited scores between index+1year to index+10 years	√ √ √ √ √ √ √ √ √ √	√ √ √ √ √ √ √ √ √ √	-	-	-	-
	Model 1 + 10-year before scores between index+1year to index+10 years	√ √ √ √ √ √ √ √ √ √	√ √ √ √ √ √ √ √ √ √	-	-	-	-
	Model 1 + 5-year before scores between index+1year to index+10 years	√ √ √ √ √ √ √ √ √ √	√ √ √ √ √ √ √ √ √ √	-	-	-	-
Moving CVD scores (Poisson models)							
	Model 1 (Age, gender, ethnicity, and IMD) at index+1year to index+10 years	-	-	√ √ √ √ √ √ √ √ √ √	-	-	-
Model 2	Model 1 + Unlimited scores at index+1year to index+10 years	-	-	√ √ √ √ √ √ √ √ √ √	-	-	-
	Model 1 + 10-year before scores at index+1year to index+10 years	-	-	√ √ √ √ √ √ √ √ √ √	-	-	-
	Model 1 + 5-year before scores at index+1year to index+10 years	-	-	√ √ √ √ √ √ √ √ √ √	-	-	-

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- Zghebi et al. 2020

Table S4 Absolute event rates (%) per CV severity score strata – Training and validation datasets

Unlimited score	1-year all-cause mortality	10-year all-cause mortality	1-year CV_DM mortality	10-year CV_DM mortality	1-year any-cause hospitalisation	10-year any-cause hospitalisation	1-year CV_DM hospitalisation	10-year CV_DM hospitalisation
Training dataset								
0	0.3	6.2	0.09	1.3	14.9	62.8	7.9	27.0
1	0.7	10.9	0.17	2.5	22.3	73.0	14.7	43.6
2	1.3	16.7	0.47	4.7	29.1	79.5	21.4	59.5
3	2.3	23.3	0.89	7.7	35.7	84.6	28.3	71.2
4	3.9	32.3	1.46	11.4	41.8	87.0	34.9	78.3
5	5.7	39.0	2.17	14.9	47.3	89.5	40.5	84.7
6	7.0	45.2	4.10	19.0	50.9	87.9	44.6	82.6
7	9.6	50.0	5.15	23.7	54.5	92.1	51.7	94.8
8	14.3	62.9	10.53	31.6	71.4	94.3	62.9	89.5
9	11.1	88.9	0.00	33.3	77.8	88.9	55.6	83.3
10	33.3	100.0	0.00	0.0	66.7	66.7	66.7	100.0
Overall	**1.2**	**14.2**	**0.36**	**3.7**	**25.2**	**74.7**	**17.8**	**48.0**
Validation dataset								
0	0.4	6.1	0.13	1.3	14.8	63.3	7.8	25.4
1	0.7	10.6	0.20	2.2	22.2	72.4	14.5	42.0
2	1.3	16.8	0.35	3.9	29.4	79.6	22.0	57.3
3	2.4	24.2	0.83	7.7	36.1	85.1	28.7	70.7
4	4.7	31.1	2.03	10.5	41.8	86.2	34.4	77.1
5	5.3	38.2	2.36	13.3	44.9	89.8	38.4	80.5
6	9.7	57.8	3.24	23.2	48.6	85.9	43.2	82.7
7	8.7	41.3	4.35	26.1	65.2	95.7	63.0	95.7
8	16.7	66.7	8.33	16.7	58.3	91.7	41.7	83.3
9	0.0	66.7	0.00	66.7	100.0	100.0	66.7	100.0
Overall	**1.2**	**14.2**	**0.41**	**3.6**	**25.3**	**74.7**	**17.9**	**47.9**

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Table S5 Cox regression models using moving CV severity scores for all-cause mortality HR (95% CI) – for 2-9 yrs windows models only adjusted HRs for severity scores are presented for simplicity – training dataset

Al-cause death		Index+1 year	Index+2 years	Index+3 years	Index+4 years	Index+5 years	Index+6 years	Index+7 years	Index+8 years	Index+9 years	Index+10 years
Age Gender IMD, ethnicity Only	**Model 1**	**AUROC**=0.7744 AIC=848,222.2	**AUROC**= 0.7722 AIC= 803,473.9	**AUROC**=0.7697 AIC= 745,513.2	**AUROC**=0.7669 AIC=691,389.6	**AUROC**=0.7645 AIC= 638,300.5	**AUROC**=0.7622 AIC= 581,226.6	**AUROC**=0.7598 AIC=526,475.7	**AUROC**= 0.7572 AIC=470,098.2	**AUROC**= 0.7552 AIC=415,919.8	**AUROC**=0.7531 AIC=362,958.9
+Ever before score	**Model 2**	• Score 1.28 (1.27; 1.29) • Age 1.11 (1.11; 1.11) • Gender (F) 0.73 (0.72; 0.75) • IMD (Q5) 1.58 (1.53; 1.63) • Race (Asian) 0.90 (0.81; 0.98) **AUROC**= 0.7822 AIC= 844,464.5	1.25 (1.24; 1.26) **AUROC**=0.7802 AIC= 800,104.8	1.24 (1.23; 1.25) **AUROC**=0.7776 AIC= 742,514.6	1.23 (1.22; 1.24) **AUROC**=0.7747 AIC= 688,677	1.22 (1.21; 1.23) **AUROC**=0.7722 AIC=635,824.1	1.21 (1.21; 1.22) **AUROC**=0.7701 AIC= 578,906.6	1.21 (1.20; 1.22) **AUROC**=0.7676 AIC=524,377.8	1.21 (1.20; 1.22) **AUROC**=0.7654 AIC= 468,110.3	1.21 (1.20; 1.22) **AUROC**=0.7640 AIC=414,083.5	• Score 1.21 (1.20; 1.22) • Age 1.10 (1.10; 1.11) • Gender (F) 0.76 (0.73; 0.78) • IMD (Q5) 1.51 (1.44; 1.59) • Race (Asian) 0.74 (0.64; 0.87) **AUROC**=0.7620 AIC= 361,310.1
+10 yrs before score	**Model 3**	• Score 1.27 (1.25; 1.27) • Age 1.11 (1.11; 1.11) • Gender (F) 0.73 (0.71; 0.74) • IMD (Q5) 1.57 (1.52; 1.63) • Race (Asian) 0.91 (0.83; 1.00) **AUROC**=0.7811 AIC= 845,152.5	1.23 (1.22; 1.24) **AUROC**=0.7789 AIC= 800,781.2	1.22 (1.21 1.23) **AUROC**=0.7762 AIC= 743,106.9	1.21 (1.20; 1.22) **AUROC**=0.7733 AIC= 689,221.3	1.20 (1.19; 1.21) **AUROC**=0.7707 AIC=636,353	1.20 (1.19; 1.20) **AUROC**=0.7686 AIC=579,404.8	1.19 (1.18; 1.20) **AUROC**=0.7660 AIC=524,842.3	1.19 (1.18; 1.20) **AUROC**=0.7637 AIC= 468,527.3	1.19 (1.18; 1.20) **AUROC**=0.7621 AIC=414,480.9	• Score 1.19 (1.18; 1.20) • Age 1.10 (1.10; 1.11) • Gender (F) 0.75 (0.72; 0.77) • IMD (Q5) 1.52 (1.44; 1.59) • Race (Asian) 0.76 (0.65; 0.89) **AUROC**=0.7599 AIC= 361,717.9

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

+5 yrs before score	**Model 4**	• Score 1.23 (1.22; 1.24) • Age 1.11 (1.11; 1.11) • Gender (F) 0.72 (0.71; 0.74) • IMD (Q5) 1.57 (1.52; 1.63) • Race (Asian) 0.92 (0.84; 1.01) **AUROC**=0.7794 AIC= 846,039.7	1.20 (1.19; 1.21) **AUROC**=0.7771 AIC= 801,618.9	1.19 (1.18; 1.20) **AUROC**= 0.7744 AIC= 743,886.8	1.18 (1.17; 1.19) **AUROC**=0.7713 AIC=689,968.4	1.16 (1.15; 1.17) **AUROC**=0.7685 AIC=637,088.9	1.16 (1.15; 1.18) **AUROC**=0.7655 AIC=580,275	1.16 (1.15; 1.18) **AUROC**=0.7627 AIC= 525,672.3	1.16 (1.15; 1.18) **AUROC**=0.7602 AIC=469,352.6	1.17 (1.15; 1.18) **AUROC**=0.7583 AIC=415,256.6	• Score 1.16 (1.15; 1.17) • Age 1.11 (1.10; 1.11) • Gender (F) 0.71 (0.69; 0.73) • IMD (Q5) 1.54 (1.47; 1.62) • Race (Asian) 0.80 (0.68; 0.93) **AUROC**=0.7561 AIC=362,395.7

Models were limited to patients contributing to each post-index window

AIC: Akaike information criterion; **AUROC**: area under a Receiver Operating Characteristics; **IMD**: index for multiple deprivation.

Q5: most deprived IMD quintile vs. least deprived IMD quintile. Race: Asian vs. White.

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Table S6 Cox regression models using moving CV severity scores for CV/DM-related mortality HR (95% CI) – for 2-9 yrs windows models only adjusted HRs for severity scores are presented for simplicity – training dataset

CV/DM death		Index+1 year	Index+2 years	Index+3 years	Index+4 years	Index+5 years	Index+6 years	Index+7 years	Index+8 years	Index+9 years	Index+10 years
Age Gender IMD, ethnicity Only	**Model 1**	**AUROC**=0.7844 AIC=219,806.1	**AUROC**= 0.7823 AIC= 205,935.6	**AUROC**=0.7789 AIC= 190,182	**AUROC**=0.7759 AIC=176,601.2	**AUROC**=0.7739 AIC=163,333	**AUROC**=0.7711 AIC= 149,588.6	**AUROC**=0.7686 AIC= 135,792.3	**AUROC**= 0.7657 AIC= 121,937.1	**AUROC**=0.7624 AIC=107,696.1	**AUROC**=0.7600 AIC= 93,469.8
+Ever before score	**Model 2**	• Score 1.50 (1.48; 1.52) • Age 1.12 (1.11; 1.12) • Gender (F) 0.77 (0.74; 0.80) • IMD (Q5) 1.55 (1.45; 1.65) • Race (Asian) 1.02 (0.86; 1.22) **AUROC**= 0.7984 AIC= 216,823.2	1.47 (1.45; 1.49) **AUROC**=0.7974 AIC=203,092	1.45 (1.43; 1.47) **AUROC**=0.7945 AIC= 187,508.5	1.44 (1.43; 1.46) **AUROC**=0.7921 AIC=174,047.4	1.43 (1.41; 1.45) **AUROC**=0.7901 AIC= 160,964.5	1.43 (1.41; 1.45) **AUROC**=0.7879 AIC= 147,333	1.42 (1.40; 1.44) **AUROC**=0.7857 AIC=133,724.7	1.42 (1.40; 1.44) **AUROC**=0.7840 AIC=119,953.9	1.43 (1.41; 1.45) **AUROC**=0.7824 AIC=105,807.8	• Score 1.42 (1.40; 1.44) • Age 1.11 (1.10; 1.11) • Gender (F) 0.83 (0.78; 0.88) • IMD (Q5) 1.40 (1.27; 1.54) • Race (Asian) 0.71 (0.52; 0.97) **AUROC**=0.7803 AIC= 918,30.52
+10 yrs before score	**Model 3**	• Score 1.48 (1.46; 1.50) • Age 1.12 (1.11; 1.12) • Gender (F) 0.76 (0.73; 0.79) • IMD (Q5) 1.54 (1.44; 1.64) • Race (Asian) 1.05 (0.88; 1.25) **AUROC**=0.7968 AIC= 217,291.4	1.44 (1.42; 1.46) **AUROC**=0.7954 AIC= 203,554.9	1.43 (1.41; 1.45) **AUROC**=0.7924 AIC= 187,954.1	1.42 (1.40; 1.44) **AUROC**=0.7899 AIC= 174,460.1	1.40 (1.38; 1.42) **AUROC**=0.7876 AIC= 161,387.1	1.39 (1.38; 1.42) **AUROC**=0.7853 AIC= 147,746.5	1.39 (1.37; 1.41) **AUROC**=0.7828 AIC= 134,101.6	1.39 (1.37; 1.41) **AUROC**=0.7810 AIC=120,299.3	1.40 (1.38; 1.42) **AUROC**=0.7789 AIC=106,162.1	• Score 1.38 (1.36; 1.41) • Age 1.11 (1.11; 1.11) • Gender (F) 0.81 (0.77; 0.86) • IMD (Q5) 1.40 (1.27; 1.54) • Race (Asian) 0.74 (0.54; 1.01) **AUROC**=0.7764 AIC= 92,174.03

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

+5 yrs before score	**Model 4**	• Score 1.44 (1.41; 1.46) • Age 1.12 (1.12; 1.12) • Gender (F) 0.75 (0.72; 0.78) • IMD (Q5) 1.54 (1.44; 1.64) • Race (Asian) 1.08 (0.90; 1.28) **AUROC**=0.7942 AIC= 217,938.5	1.40 (1.38; 1.42) **AUROC**=0.7924 AIC= 204,191.6	1.38 (1.36; 1.40) **AUROC**=0.7892 AIC= 188,572.5	1.37 (1.35; 1.39) **AUROC**=0.7863 AIC=175,067.9	1.35 (1.33; 1.37) **AUROC**=0.7837 AIC=161,964.9	1.35 (1.33; 1.38) **AUROC**=0.7793 AIC= 148,485.2	1.36 (1.33; 1.38) **AUROC**=0.7761 AIC=134,828.3	1.37 (1.35; 1.40) **AUROC**=0.7738 AIC=120,991	1.37 (1.35; 1.40) **AUROC**=0.7709 AIC=106,834.5	• Score 1.36 (1.34; 1.39) • Age 1.11 (1.11; 1.12) • Gender (F) 0.73 (0.69; 0.78) • IMD (Q5) 1.43 (1.30; 1.57) • Race (Asian) 0.82 (0.61; 1.12) **AUROC**=0.7685 AIC= 92,737.82

Models were limited to patients contributing to each post-index window.

AIC: Akaike information criterion; **AUROC**: area under a Receiver Operating Characteristics; **IMD**: index for multiple deprivation.

Q5: most deprived IMD quintile vs. least deprived IMD quintile. Race: Asian vs. White.

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- Zghebi et al. 2020

Table S7 Multiple events Cox regression models using baseline CV severity scores for any cause hospital admissions HR (95% CI) using Breslow option – training dataset*

Any hosp_breslow Multiple failure	Predictor(s)	HR (95% CI)	AIC
Model 1 w/o severity score	**Age**	1.03 (1.02; 1.03)	2.27e+07
	Gender (F)	0.80 (0.77; 0.83)	
	Patient-level IMD		
	• Q1 (least deprived)	Referent	
	• Q5 (most deprived)	1.37 (1.29; 1.46)	
	• Unknown	0.98 (0.79; 1.22)	
	Ethnicity		
	• White	Referent	
	• Asian	1.56 (1.25; 1.94)	
Model 2	**Ever before score**	1.3.1 (1.28; 1.34)	2.26e+07
	Age	1.02 (1.02; 1.02)	
	Gender (F)	0.84 (0.81; 0.87)	
	Patient-level IMD		
	• Q1 (least deprived)	Referent	
	• Q5 (most deprived)	1.33 (1.25; 1.42)	
	• Unknown	1.00 (0.80; 1.25)	
	Ethnicity		
	• White	Referent	
	• Asian	1.45 (1.17; 1.79)	
Model 3	**10 yrs before score**	1.31 (1.27; 1.34)	
	Age	1.01 (1.02; 1.02)	2.26e+07
	Gender (F)	0.83 (0.80; 0.87)	
	Patient-level IMD		
	• Q1 (least deprived)	Referent	
	• Q5 (most deprived)	1.33 (1.25; 1.42)	
	• Unknown	0.99 (0.79; 1.23)	
	Ethnicity		
	• White	Referent	
	• Asian	1.46 (1.18; 1.81)	
Model 4	**5 yrs before score**	1.31 (1.27; 1.34)	2.27e+07
	Age	1.02 (1.02; 1.02)	
	Gender (F)	0.83 (0.80; 0.86)	
	Patient-level IMD		
	• Q1 (least deprived)	Referent	
	• Q5 (most deprived)	1.34 (1.25; 1.42)	
	• Unknown	0.97 (0.79; 1.22)	
	Ethnicity		
	• White	Referent	
	• Asian	1.48 (1.20; 1.84)	

*based on using the longest hospital stay in case of multiple same-day admissions

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Table S8 Poisson regression models using moving severity scores vs. the count of <u>any cause hospitalisation</u> admission(s) in the following year (IRR (95% CI)) - for 2-9 yrs windows models only IRRs for severity scores are presented for simplicity – training dataset

Any apc hosp		Index+1 year	Index+2 years	Index+3 years	Index+4 years	Index+5 years	Index+6 years	Index+7 years	Index+8 years	Index+9 years	Index+10 years
Age Gender IMD, ethnicity Only	**Model 1**	**AIC**=477,990.9	**AIC**=430,060.8	**AIC**=412,502	**AIC**=399,811.4	**AIC**=390,006.7	**AIC**=383,434.3	**AIC**=369,291.8	**AIC**=345,273.2	**AIC**=325,745.9	**AIC**=309,306
+Ever before score	**Model 2**	• Score 1.38 (1.38; 1.39) • Age 1.01 (1.01; 1.01) • Gender (F) 0.84 (0.83; 0.85) • IMD (Q5) 1.30 (1.27; 1.33) **AIC**=456,825	1.23 (1.22; 1.23) **AIC=**221,331.4	1.17 (1.16; 1.18) **AIC=**144,601.4	1.13 (1.12; 1.15) **AIC=**115,227.6	1.13 (1.12; 1.14) **AIC=**94,969.21	1.15 (1.14; 1.17) **AIC=**81,186.23	1.17 (1.15; 1.19) **AIC=**65,785.7	1.20 (1.18; 1.22) **AIC=**55,033.83	1.25 (1.23; 1.27) **AIC=**45,422.61	• Score 1.19 (1.16; 1.21) • Age 1.01 (1.00; 1.01) • Gender (F) 1.13 (1.07; 1.19) • IMD (Q5) 1.07 (0.97; 1.17) **AIC=**37,918.6
+10 yrs before score	**Model 3**	• Score 1.39 (1.39; 1.40) • Age 1.01 (1.01; 1.01) • Gender (F) 0.84 (0.83; 0.85) • IMD (Q5) 1.30 (1.27; 1.32) **AIC**=457,753.5	1.23 (1.22; 1.24) **AIC=**221,537.4	1.18 (1.17; 1.19) **AIC=**144,624.4	1.14 (1.13; 1.16) **AIC=**115,206.7	1.15 (1.13; 1.16) **AIC=**94,920.38	1.18 (1.17; 1.20) **AIC=**81,078.77	1.22 (1.20; 1.24) **AIC=**65,587.48	1.28 (1.26; 1.30) **AIC=**54,718.66	**1.38 (1.35; 1.40)** **AIC= 44,921.51**	• Score 1.37 (1.34; 1.40) • Age 1.01 (1.00;1.01) • Gender (F) 1.15 (1.09; 1.22) • IMD (Q5) 1.04 (0.95; 1.14) AIC=37,389.42
+5 yrs before score	**Model 4**	• Score 1.40 (1.39; 1.41) • Age 1.01 (1.01; 1.01) • Gender (F) 0.83 (0.82; 0.84) • IMD (Q5) 1.31 (1.28; 1.33) **AIC**=460,109	1.22 (1.21; 1.23) **AIC=**221,985.3	1.18 (1.17; 1.19) **AIC=**144,758.2	1.14 (1.13; 1.16) **AIC=**115,260.8	1.17 (1.15; 1.19) **AIC=**94,855.22	1.32 (1.30; 1.34) AIC=80,588.38	1.47 (1.44; 1.49) **AIC**=64,630.33	1.57 (1.54; 1.60) **AIC**=53,757.77	1.72 (1.69; 1.76) **AIC**=43,891.78	• Score 1.71 (1.66; 1.75) • Age 1.01 (1.01; 1.12) • Gender (F) 1.06 (1.00; 1.17) • IMD (Q5) 1.03 (0.94; 1.13) AIC=36,749.45

Models were limited to patients contributing to each post-index window.

AIC: Akaike information criterion; **AUROC**: area under a Receiver Operating Characteristics; **IMD**: index for multiple deprivation.

Q5: most deprived IMD quintile vs. least deprived IMD quintile. Race: Asian vs. White.

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- Zghebi et al. 2020

Table S9 Survival models for 1, 3, 5, and 10-year aggregated any hospitalisation and all-cause mortality outcome using baseline scores HR (95% CI) – training dataset

any hosp or all-death	Predictor(s)	1-year	3-year	5-year	10-year
Model 1 w/o severity score	**Age, gender, deprivation, ethnicity only**	**AUROC**= 0.6055 **AIC**= 1,008,481	**AUROC**= 0.6115 **AIC**= 1,995,512	**AUROC**= 0.6140 **AIC**= 2,424,369	**AUROC**= 0.6169 **AIC**= 2,915,837
Model 2		1.27 (1.26; 1.28)	1.24 (1.23; 1.25)	1.22 (1.22; 1.23)	1.21 (1.20; 1.21)
	Ever before score	**AUROC**= 0.6271 **AIC**= 1,004,865	**AUROC**= 0.6288 **AIC**= 1,990,204	**AUROC**= 0.6294 **AIC**= 2,418,915	**AUROC**= 0.6307 **AIC**= 2,910,327
	Age	1.01 (1.01; 1.01)	1.02 (1.02; 1.02)	1.02 (1.02; 1.02)	1.02 (1.02; 1.02)
	Gender (F)	0.72 (0.70; 0.73)	0.80 (0.79; 0.81)	0.83 (0.82; 0.84)	0.85 (0.84; 0.86)
	IMD (vs. least deprived)				
	• Q5 (most deprived)	1.05 (1.02; 1.09)	1.14 (1.12; 1.17)	1.17 (1.14; 1.19)	1.18 (1.16; 1.20)
	Ethnicity (vs. White)				
	• Black	0.89 (0.80; 0.98)	0.97 (0.90; 1.04)	1.01 (0.95; 1.08)	1.02 (0.96; 1.08)
	• Asian	1.33 (1.26; 1.41)	1.28 (1.23; 1.33)	1.26 (1.21; 1.31)	1.23 (1.19; 1.28)
	• Mixed	0.96 (0.80; 1.15)	1.02 (0.90; 1.16)	1.08 (0.96; 1.21)	1.02 (0.92; 1.14)
	• Other	1.02 (0.92; 1.14)	0.98 (0.91; 1.06)	0.98 (0.91; 1.05)	0.95 (0.89; 1.02)
	• Unknown	0.26 (0.24; 0.28)	0.25 (0.24; 0.26)	0.25 (0.24; 0.26)	0.26 (0.25; 0.26)
Model 3		1.28 (1.27; 1.29)	1.24 (1.24; 1.25)	1.22 (1.22; 1.23)	1.21 (1.20; 1.21)
	10 yrs before score	**AUROC**= 0.6270 **AIC**= 1,004,939	**AUROC**= 0.6281 **AIC**= 1,990,505	**AUROC**= 0.6286 **AIC**= 2,419,291	**AUROC**= 0.6296 **AIC**= 2,910,771
	Age	1.01 (1.01; 1.01)	1.02 (1.02; 1.02)	1.02 (1.02; 1.02)	1.02 (1.02; 1.02)
	Gender (F)	0.72 (0.70; 0.73)	0.80 (0.79; 0.81)	0.83 (0.82; 0.84)	0.84 (0.84; 0.85)
	IMD (vs. least deprived)				
	• Q5 (most deprived)	1.05 (1.02; 1.09)	1.14 (1.12; 1.17)	1.16 (1.14; 1.19)	1.18 (1.16; 1.20)
	Ethnicity (vs. White)				
	• Black	0.89 (0.80; 0.99)	0.98 (0.91; 1.05)	1.02 (0.96; 1.09)	1.03 (0.97; 1.09)
	• Asian	1.34 (1.27; 1.42)	1.29 (1.24; 1.34)	1.27 (1.22; 1.32)	1.24 (1.19; 1.29)
	• Mixed	0.96 (0.80; 1.16)	1.02 (0.90; 1.16)	1.08 (0.96; 1.21)	1.02 (0.92; 1.14)
	• Other	1.02 (0.92; 1.14)	0.99 (0.91; 1.07)	0.98 (0.92; 1.06)	0.95 (0.89; 1.02)
	• Unknown	0.26 (0.24; 0.28)	0.25 (0.24; 0.26)	0.25 (0.24; 0.26)	0.26 (0.25; 0.26)
Model 4		1.30 (1.29; 1.31)	1.25 (1.24; 1.26)	1.23 (1.22; 1.23)	1.21 (1.20; 1.21)
	5 yrs before score	**AUROC**= 0.6265 **AIC**= 1,005,101	**AUROC**= 0.6266 **AIC**= 1,991,076	**AUROC**= 0.6268 **AIC**= 2,419,997	**AUROC**=0.6277 **AIC**= 2,911,543
	Age	1.01 (1.01; 1.01)	1.02 (1.02; 1.02)	1.02 (1.02; 1.02)	1.02 (1.02; 1.02)
	Gender (F)	0.71 (0.70; 0.73)	0.79 (0.78; 0.80)	0.82 (0.81; 0.83)	0.84 (0.83; 0.85)
	IMD (vs. least deprived)				
	• Q5 (most deprived)	1.05 (1.02; 1.09)	1.14 (1.12; 1.17)	1.17 (1.14; 1.19)	1.18 (1.16; 1.20)
	Ethnicity (vs. White)				
	• Black	0.90 (0.81; 0.99)	0.98 (0.92; 1.06)	1.03 (0.97; 1.10)	1.04 (0.98; 1.10)
	• Asian	1.36 (1.29; 1.44)	1.31 (1.25; 1.36)	1.28 (1.23; 1.34)	1.25 (1.21; 1.30)
	• Mixed	0.97 (0.81; 1.16)	1.03 (0.90; 1.17)	1.08 (1.00; 1.21)	1.02 (0.92; 1.14)
	• Other	1.03 (0.92; 1.14)	0.99 (0.92; 1.07)	0.99 (0.92; 1.06)	0.96 (0.90; 1.02)
	• Unknown	0.26 (0.24; 0.28)	0.25 (0.24; 0.28)	0.25 (0.24; 0.26)	0.25 (0.25; 0.26)

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- Zghebi et al. 2020

Table S10 - Summary of AUROCs of fitted severity score-only Cox regression models – training and validation datasets

AUROC (Gonen and Heller's K)	Training dataset				Validation dataset			
Outcome	Ever before score (Unlimited)	10-year before score	5-year before score	Unlimited score categories	Ever before score (Unlimited)	10-year before score	5-year before score	Unlimited score categories
1-year all-cause mortality	0.6446	0.6367	0.6241	0.6391	0.6445	0.6333	0.6230	0.6307
3-year all-cause mortality	0.6424	0.6333	0.6188	0.6335	0.6399	0.6311	0.6442	0.6203
5-year all-cause mortality	0.6380	0.6288	0.6146	0.6276	0.6406	0.6315	0.6172	0.6269
10-year all-cause mortality	0.6646	0.6203	0.6054	0.6185	0.6340	0.6248	0.6092	0.6194
1-year CV/diabetes mortality	0.6662	0.6584	0.6429	0.6538	0.6617	0.6564	0.6416	0.6383
3-year CV/diabetes mortality	0.6669	0.6582	0.6430	0.6645	0.6671	0.6602	0.6452	0.6347
5-year CV/diabetes mortality	0.6646	0.6558	0.6413	0.6599	0.6690	0.6609	0.6453	0.6448
10-year CV/diabetes mortality	0.6598	0.6501	0.6345	0.6518	0.6616	0.6540	0.6376	0.6412
1-year aggregated any hospitalisation or mortality	0.5845	0.5832	0.5801	0.5809	-	-	-	-
3-year aggregated any hospitalisation or mortality	0.5829	0.5801	0.5746	0.5747	-	-	-	-
5-year aggregated any hospitalisation or mortality	0.5812	0.5779	0.5715	0.5718	-	-	-	-
10-year aggregated any hospitalisation or mortality	0.5799	0.5760	0.5692	0.5695	-	-	-	-

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Table S11 f of AIC and AUROCs of fitted Cox regression models without IMD – training dataset

Model	1-year all-cause mortality	3-year all-cause mortality	5-year all-cause mortality	10-year all-cause mortality	1-year CV/diabetes-related mortality	3-year CV/diabetes-related mortality	5-year CV/diabetes-related mortality	10-year CV/diabetes-related mortality	1-year aggregated any hospitalisation or mortality	3-year aggregated any hospitalisation or mortality	5-year aggregated any hospitalisation or mortality	10-year aggregated any hospitalisation or mortality
Model 1a	AUROC=0.7860 (AIC=43,059)	AUROC=0.7821 (AIC=155,632.1)	AUROC=0.7808 (AIC= 265,700.5)	AUROC=0.7776 (AIC= 533,430.3)	AUROC=0.7958 (AIC=13,382)	AUROC=0.7955 (AIC=42,933.69)	AUROC=0.7930 (AIC= 69,906.25)	AUROC=0.7899 (AIC= 138,197)	AUROC=0.6051 (AIC= 1,008,511)	AUROC=0.6107 (AIC=1,995, 732)	AUROC=0.6131 (AIC= 2,424,698)	AUROC=0.6160 (AIC=2,916, 288)
Model 2a	AUROC= 0.7906 (AIC=42,597 .87)	AUROC=0.7875 (AIC= 154,095.5)	AUROC=0.7866 (AIC= 263,462)	AUROC=0.7843 (AIC=530,087.5)	AUROC= 0.8024 (AIC= 13,091.68)	AUROC=0.8033 (AIC= 42,008.19)	AUROC=0.8016 (AIC=68,518 .67)	AUROC= 0.8004 (AIC= 135,915.5)	AUROC=0.6269 (AIC= 1,004,876)	AUROC=0.6283 (AIC=1,990, 366)	AUROC=0.6288 (AIC= 2,419,175)	AUROC=0.6301 (AIC=2,910, 691)
Model 3a	AUROC= 0.7907 (AIC=42,659 .83)	AUROC=0.7872 (AIC= 154,364.7)	AUROC=0.7861 (AIC=263,843.8)	AUROC=0.7836 (AIC= 530,648.4)	AUROC=0.8026 (AIC=13,129)	AUROC=0.8031 (AIC= 42,150.3)	AUROC=0.8013 (AIC= 68,725.07)	AUROC=0.7995 (AIC= 136,260.7)	AUROC=1004949 (AIC= 1,004,949)	AUROC=0.6276 (AIC=1,990, 664)	AUROC=0.6280 (AIC=2,419, 548)	AUROC=0.6290 (AIC=2,911, 132)
Model 4a	AUROC= 0.7905 (AIC=42,742 .73)	AUROC=0.7866 (AIC= 154,696.2)	AUROC=0.7854 (AIC=264,316.9)	AUROC=0.7824 (AIC=531,402.4)	AUROC= 0.8018 (AIC= 13,199.54)	AUROC=0.8023 (AIC=42,357 .11)	AUROC=0.8003 (AIC=69,012. 6)	AUROC=0.7978 (AIC=136,754.8)	AUROC=0.6263 (AIC= 1,005,112)	AUROC=0.6261 (AIC= 1,991,239)	AUROC=0.6262 (AIC=2,420, 258)	AUROC=0.6271 (AIC=2,911, 909)

*Competing risk analysis; $ adjusted for age, gender, IMD only.

Model 1a: age, gender, and ethnicity.

Model 2a: Model 1a + ever before severity score.

Model 3a: Model 1a + 10-year severity score

Model 4a: Model 1a + 5-year severity score

AIC: Akaike information criterion; **AUROC**: area under a Receiver Operating Characteristics curve; **CV**: cardiovascular.

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- Zghebi et al. 2020

Table S12 Calibration test results for baseline CVD scores in training and validation datasets using Somer's D

Somer's D	Training dataset				Validation dataset			
	Ever before score (Unlimited)	10-year before score	5-year before score	Unlimited score categories	Ever before score (Unlimited)	10-year before score	5-year before score	Unlimited score categories
1-year all-cause mortality	0.2891	0.2734	0.2482	0.2782	0.2889	0.2667	0.2459	0.2615
3-year all-cause mortality	0.2848	0.2665	0.2376	0.2671	0.2799	0.2622	0.2349	0.2407
5-year all-cause mortality	0.2760	0.2576	0.2292	0.2552	0.2812	0.2631	0.2343	0.2537
10-year all-cause mortality	0.2600	0.2407	0.2109	0.2369	0.2680	0.2496	0.2185	0.2389
1-year CV/diabetes mortality	0.3323	0.3169	0.2858	0.3077	0.3233	0.3128	0.2832	0.2765
3-year CV/diabetes mortality	0.3338	0.3165	0.2861	0.3290	0.3341	0.3205	0.2904	0.2694
5-year CV/diabetes mortality	0.3293	0.3116	0.2827	0.3198	0.3380	0.3218	0.2905	0.2895
10-year CV/diabetes mortality	0.3196	0.3003	0.2690	0.3035	0.3232	0.3080	0.2752	0.2823
1-year aggregated any hospitalisation or mortality	0.1691	0.1664	0.1601	0.1618	-	-	-	-
3-year aggregated any hospitalisation or mortality	0.1658	0.1850	0.1491	0.1495	-	-	-	-
5-year aggregated any hospitalisation or mortality	0.1624	0.1558	0.1431	0.1436	-	-	-	-
10-year aggregated any hospitalisation or mortality	0.1598	0.1521	0.1385	0.1390	-	-	-	-

Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Table S13 Calibration test results for moving CVD scores in training dataset using Somer's D

Somer's D		Index+1 year	Index+2 years	Index+3 years	Index+4 years	Index+5 years	Index+6 years	Index+7 years	Index+8 years	Index+9 years	Index+10 years
All-cause mortality	Unlimited score	0.2393	0.2324	0.2280	0.2253	0.2234	0.2246	0.2241	0.2272	0.2295	0.2305
	10-year before score	0.2176	0.2098	0.2064	0.2037	0.2010	0.2022	0.2014	0.2055	0.2071	0.2054
	5-year before score	0.1866	0.1784	0.1744	0.1708	0.1651	0.1492	0.1437	0.1423	0.1396	0.1374
CV/diabetes mortality	Unlimited score	0.3081	0.3064	0.3054	0.3066	0.3054	0.3075	0.3077	0.3128	0.3186	0.3168
	10-year before score	0.2857	0.2829	0.2822	0.2837	0.2807	0.2823	0.2825	0.2889	0.2926	0.2878
	5-year before score	0.2514	0.2476	0.2458	0.2461	0.2407	0.2161	0.2093	0.2112	0.2088	0.2047

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- *Zghebi et al. 2020*

Table S14 Survival models for 1-year all-cause mortality outcome using baseline scores HR (95% CI) - Validation dataset

1-yr all-cause death	Predictor(s)	HR (95% CI)	AUROC AIC
Model 1 **w/o severity score**	**Age, gender, deprivation, ethnicity only**	HRs Similar to below	**AUROC**= 0.7862 **AIC**= 10,023.61
Model 2	**Ever before score**	1.40 (1.32 to 1.48)	**AUROC**= 0.7908 **AIC**= 9,903.685
	Age	1.11 (1.10 to 1.12)	
	Gender (F)	0.64 (0.54 to 0.76)	
	IMD (vs. least deprived) • Q5 (most deprived)	1.33 (1.00 to 1.76)	
	Ethnicity (vs. White)		
	• Black	1.03 (0.33 to 3.20)	
	• Asian	1.63 (0.89 to 2.98)	
	• Mixed	3.44 (0.86 to 13.8)	
	• Other	0.62 (0.15 to 2.47)	
	• Unknown	1.37 (0.94 to 1.99)	
Model 3	**10 yrs before score**	1.36 (1.28 to 1.45)	**AUROC**= 0.7899 **AIC**= 9,935.482
	Age	1.11 (1.10 to 1.12)	
	Gender (F)	0.64 (0.53 to 0.76)	
	IMD (vs. least deprived)	Referent	
	• Q5 (most deprived)	1.32 (1.00 to 1.76)	
	Ethnicity (vs. White)		
	• Black	1.08 (0.35 to 3.37)	
	• Asian	1.66 (0.91 to 3.04)	
	• Mixed	3.34 (0.83 to 13.4)	
	• Other	0.64 (0.16 to 2.58)	
	• Unknown	1.31 (0.90 to 1.91)	
Model 4	**5 yrs before score**	1.38 (1.29 to 1.47)	**AUROC=** 0.7901 **AIC**= 9,946.772
	Age	1.12 (1.11 to 1.13)	
	Gender (F)	0.63 (0.53 to 0.75)	
	IMD (vs. least deprived)		
	Q5 (most deprived)	1.30 (0.98 to 1.73)	
	Ethnicity (vs. White)		
	• Black	1.09 (0.35 to 3.41)	
	• Asian	1.71 (0.94 to 3.11)	
	• Mixed	3.44 (0.86 to 13.8)	
	• Other	0.67 (0.17 to 2.71)	
	• Unknown	1.28 (0.88 to 1.86)	

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- Zghebi et al. 2020

Table S15 Survival models for 1, 3, 5, and 10-year CV/diabetes-related mortality outcome using baseline scores HR (95% CI) - Validation dataset

1-10yr CV/DM death	Predictor(s)	1-year	3-year	5-year	10-year
Model 1 w/o severity score	**Age, gender, deprivation, ethnicity only**	HR Similar to below (**AUROC**= 0.7934 (**AIC**= 3,295.584	HR Similar to below (**AUROC**= 0.7945) (**AIC**= 9,555.281)	HR Similar to below (**AUROC**= 0.7869) (**AIC**= 15,109.95)	HR Similar to below (**AUROC**= 0.7852) (**AIC**= 29,798.48)
Model 2	**Ever before score**	1.54 (1.40 to 1.69) **AUROC**=0.7993 **AIC**=3,227.852	1.59 (1.51 to 1.68) **AUROC**= 0.8018 **AIC**= 9,324.133	1.62 (1.55 to 1.70) **AUROC**= 0.7958 **AIC**= 14,734.48	1.57 (1.52 to 1.63) **AUROC**= 0.7959 **AIC**= 29,239.54
	Age	1.11 (1.10 to 1.13)	1.12 (1.10 to 1.13)	1.11 (1.10 to 1.12)	1.11 (1.10 to 1.12)
	Gender (F)	0.63 (0.47 to 0.86)	0.78 (0.65 to 0.93)	0.78 (0.67 to 0.90)	0.73 (0.66 to 0.81)
	IMD (vs. least deprived) • Q5 (most deprived)	1.21 (0.72 to 2.04)	1.27 (0.94 to 1.72)	1.28 (1.00 to 1.63)	1.37 (1.16 to 1.63)
	Ethnicity (vs. White)				
	• Black	2.11 (0.52 to 8.58)	1.55 (0.58 to 4.16)	2.16 (1.11 to 4.18)	1.30 (0.70 to 2.43)
	• Asian	0.44 (0.06 to 3.15)	0.64 (0.24 to 1.72)	0.98 (0.52 to 1.83)	0.79 (0.48 to 1.30)
	• Mixed	5.24 (0.73 to 37.5)	3.58 (0.89 to 14.4)	2.20 (0.55 to 8.84)	1.46 (0.47 to 4.21)
	• Other	0.91 (0.13 to 6.53)	1.26 (0.47 to 3.37)	1.38 (0.66 to 2.91)	0.83 (0.41 to 1.66)
	• Unknown	1.52 (0.80 to 2.90)	1.45 (0.98 to 2.14)	1.50 (1.10 to 2.03)	1.44 (1.16 to 1.78)
Model 3	**10 yrs before score**	1.55 (1.41 to 1.71) **AUROC**= 0.7989 **AIC**= 3,233.783	1.59 (1.50 to 1.69) (**AUROC**= 0.8016) **AIC**= 9,353.166	1.61 (1.53 to 1.69) **AUROC**= 0.7952 **AIC**=14,792.79	1.57 (1.51 to 1.63) **AUROC**=0.7949) **AIC=**29304.99
	Age	1.12 (1.10 to 1.13)	1.12 (1.11 to 1.13)	1.11 (0.67 to 0.89)	1.11 (1.11 to 1.12)
	Gender (F)	0.63 (0.46 to 0.86)	0.77 (0.64 to 0.93)	0.77 (0.70 to 0.81)	0.73 (0.66 to 0.80)
	IMD (vs. least deprived)	Referent	Referent	Referent	Referent
	• Q5 (most deprived)	1.21 (0.72 to 2.04)	1.28 (0.95 to 1.73)	1.28 (1.01 to 1.64)	1.37 (1.16 to 1.63)
	Ethnicity (vs. White)				
	• Black	2.30 (0.57 to 9.35)	1.69 (0.63 to 2.65)	2.04 (1.24 to 4.67)	1.43 (0.77 to 2.68)
	• Asian	0.44 (0.06 to 3.18)	0.64 (0.24 to 1.72)	0.99 (0.53 to 1.86)	0.80 (0.49 to 1.31)
	• Mixed	5.05 (0.71 to 36.2)	3.39 (0.84 to 13.6)	2.09 (0.52 to 8.36)	1.40 (0.45 to 4.35)
	• Other	0.98 (0.14 to 6.97)	1.35 (0.50 to 3.62)	1.51 (0.72 to 3.18)	0.89 (0.44 to 1.78)
	• Unknown	1.47 (0.77 to 2.79)	1.38 (0.93 to 2.04)	1.43 (1.05 to 1.94)	1.39 (1.12 to 1.72)
Model 4	**5 yrs before score**	1.54 (1.38 to 1.71) **AUROC=** 0.7989 **AIC**= 3,247.165	1.58 (1.48 to 1.69) **AUROC**=0.8014 **AIC**= 9,397.813	1.59 (1.51 to 1.68) **AUROC**= 0.7947 **AIC**= 14,863	1.54 (1.48 to 1.61) **AUROC**= 0.7935) **AIC**= 29,416.99
	Age	1.12 (1.10 to 1.14)	1.12 (1.11 to 1.13)	1.12 (1.11 to 1.12)	1.12 (1.11 to 1.12)
	Gender (F)	0.62 (0.46 to 0.85)	0.76 (0.64 to 0.92)	0.76 (0.66 to 0.88)	0.72 (0.65 to 0.80)
	IMD (vs. least deprived) Q5 (most deprived)	1.18 (0.70 to 1.98)	1.27 (0.94 to 1.71)	1.27 (1.00 to 1.62)	1.36 (1.15 to 1.61)
	Ethnicity (vs. White)				
	• Black	2.34 (0.58 to 9.52)	1.71 (0.64 to 4.59)	2.43 (1.25 to 4.71)	1.44 (0.77 to 2.68)
	• Asian	0.46 (0.06 to 3.32)	0.67 (0.25 to 1.80)	1.05 (0.56 to 1.960)	0.85 (0.52 to 1.39)
	• Mixed	5.23 (0.73 to 37.8)	3.54 (0.88 to 14.2)	2.14 (0.53 to 8.60)	1.44 (0.46 to 4.48)
	• Other	1.04 (0.15 to 7.47)	1.48 (0.55 to 3.97)	1.66 (0.79 to 3.50)	0.96 (0.48 to 1.92)
	• Unknown	1.39 (0.73 to 2.65)	1.32 (0.89 to 1.94)	1.35 (1.00 to 1.84)	1.34 (1.08 to 1.66)

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- Zghebi et al. 2020

Table S16 Competing risk analysis models for 1-year any cause or CV/DM-related hospitalisation (competed by all-cause death) using baseline scores - Validation dataset

		1-year any cause hospitalisation		1-year CV/DM hospitalisation	
1-yr hosp (competing)	**Predictor(s)**	**SHR (95% CI)**	**AIC**	**SHR (95% CI)**	**AIC**
Model 1 w/o severity score	**Age, gender, deprivation, ethnicity only**	HR Similar to below	**220,363.5**	HR Similar to below	**156,705.3**
Model 2	**Ever before score**	1.26 (1.25; 1.28)		1.36 (1.34; 1.38)	
	Age	1.01 (1.01; 1.01)		1.01 (1.00; 1.01)	
	Gender (F)	0.71 (0.69; 0.74)		0.60 (0.58; 0.63)	
	IMD (vs. least deprived)				
	Q5 (most deprived)	1.07 (1.01; 1.14)		1.04 (0.97; 1.12)	
	Ethnicity (vs. White)		**219,487.8**		**155,549.8**
	• Black	0.72 (0.57; 0.90)		0.64 (0.48; 0.84)	
	• Asian	1.13 (1.01; 1.27)		1.30 (0.98; 1.30)	
	• Mixed	1.25 (0.79; 1.14)		1.26 (0.84; 1.91)	
	• Other	0.98 (0.80; 1.21)		0.97 (0.76; 1.25)	
	• Unknown	0.25 (0.22; 0.29)		0.26 (0.22; 0.31)	
Model 3	**10 yrs before score**	1.27 (1.25; 1.29)		1.36 (1.34 to 1.39)	
	Age	1.01 (1.01; 1.01)		1.01 (1.01 to 1.01)	
	Gender (F)	0.71 (0.69; 0.74)		0.60 (0.58 to 0.63)	
	IMD (vs. least deprived)				
	Q5 (most deprived)	1.06 (1.00; 1.13)		1.03 (0.96; 1.11)	
	Ethnicity (vs. White)		**219,551**		**155,634.7**
	• Black	0.73 (0.59; 0.92)		0.66 (0.50; 0.87)	
	• Asian	1.14 (1.02; 1.28)		1.14 (0.99; 1.31)	
	• Mixed	1.25 (0.87; 1.78)		1.26 (0.83; 1.91)	
	• Other	0.98 (0.80; 1.21)		0.98 (0.76; 1.25)	
	• Unknown	0.25 (0.22; 0.29)		0.26 (0.21; 0.30)	
Model 4	**5 yrs before score**	1.29 (1.27; 1.31)		1.39 (1.36; 1.42)	
	Age	1.01 (1.01; 1.01)		1.01 (1.01; 1.01)	
	Gender (F)	0.71 (0.69; 0.74)		0.60 (0.57; 0.63)	
	IMD (vs. least deprived)				
	Q5 (most deprived)	1.06 (1.00; 1.13)		1.03 (0.95; 1.10)	
	Ethnicity (vs. White)		**219,566.1**		**155,676.2**
	• Black	0.74 (0.59; 0.93)		0.67 (0.50; 0.89)	
	• Asian	1.16 (1.04; 1.31)		1.17 (1.02; 1.35)	
	• Mixed	1.25 (0.88; 1.79)		1.26 (0.84; 1.91)	
	• Other	1.00 (0.81; 1.23)		1.00 (0.78; 1.28)	
	• Unknown	0.25 (0.22; 0.29)		0.25 (0.21; 0.30)	

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Table S17 Summary of AIC and AUROCs of fitted regression models – Validation dataset

Predictors /model	1-year all-cause mortality	1-year CV/diabetes-related mortality	1-year hospitalisation (Single event)*	1-year CV/diabetes-related hospitalisa-tion*
Demographics-only model				
Model 1	AUROC=0.7862 (AIC= 10,023.61)	AUROC= 0.7934 (AIC= 3,295.584)	AIC= 220,363.5	AIC= 156,705.3
Severity score + demographics models				
Model 2 (Model 1 + ever before severity score)	AUROC= 0.7908 (AIC= 9,903.685)	AUROC= 0.7993 (AIC= 3,227.85)	AIC= 219,487.8	AIC= 155,549.8
Model 3 (Model 1 + 10-year severity score)	AUROC= 0.7899 (AIC= 9,935.482)	AUROC= 0.7989 (AIC= 3,233.783)	AIC= 21,9551	AIC= 155,634.7
Model 4 (Model 1 + 5-year severity score)	AUROC= 0.7901 (AIC= 9,946.772)	AUROC= 0.7989 (AIC= 3,247.165)	AIC= 219,566.1	AIC= 155,676.2

*Competing risk analysis; $ adjusted for age, gender, IMD only

Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

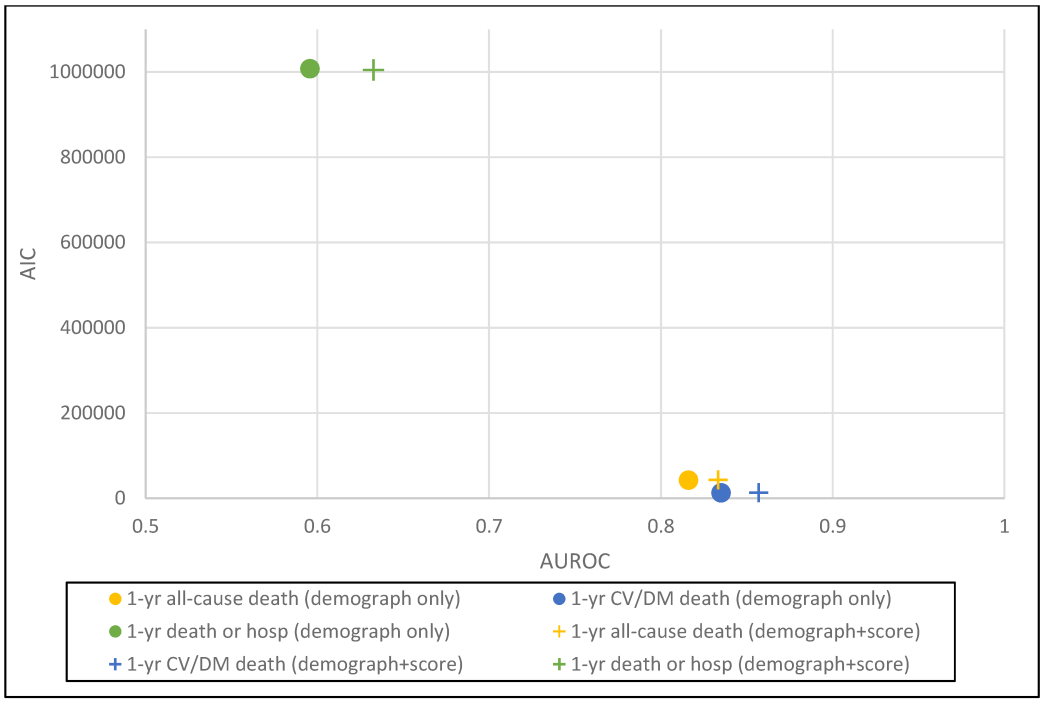
Table S18 Summary of AIC and AUROCs of fitted Cox regression models without IMD – validation dataset

Predictors /model	1-year all-cause mortality	3-year all-cause mortality	5-year all-cause mortality	10-year all-cause mortality	1-year CV/diabetes-related mortality	3-year CV/diabetes-related mortality	5-year CV/diabetes-related mortality	10-year CV/diabetes-related mortality
Model 1a	AUROC= 0.7859 (AIC= 10,025.01)	AUROC= 0.7841 (AIC= 34,719.34)	AUROC= 0.7836 (AIC= 59,255.96)	AUROC= 0.7785 (AIC= 116,868.5)	AUROC= 0.7930 (AIC= 3,291.006)	AUROC= 0.7945 (AIC= 9,551.787)	AUROC= 0.7869 (AIC= 15,110.67)	AUROC= 0.7851 (AIC= 29,815.99)
Model 2a	AUROC= 0.7905 (AIC= 9,904.208)	AUROC= 0.7888 (AIC= 34,378.47)	AUROC= 0.7890 (AIC= 58,681.04)	AUROC= 0.7851 (AIC= 115,971.3)	AUROC= 0.7990 (AIC= 3,221.004)	AUROC= 0.8018 (AIC= 9,319.359)	AUROC= 0.7959 (AIC= 14,732.54)	AUROC= 0.7958 (AIC= 29,254.44)
Model 3a	AUROC= 0.7896 (AIC= 9,936.146)	AUROC= 0.7884 (AIC= 34,443.99)	AUROC= 0.7883 (AIC= 58,788.02)	AUROC= 0.7842 (AIC= 116,120.1)	AUROC= 0.7986 (AIC= 3,226.968)	AUROC= 0.8015 (AIC= 9,350.691)	AUROC= 0.7953 (AIC= 14,791.14)	AUROC= 0.7948 (AIC= 29,319.57)
Model 4a	AUROC= 0.7899 (AIC= 9,946.966)	AUROC= 0.7882 (AIC= 34,501.83)	AUROC= 0.7879 (AIC= 58,889.06)	AUROC= 0.7831 (AIC= 116,304.7)	AUROC= 0.7987 (AIC= 32,42.228)	AUROC=0.8013 (AIC= 9,394.803)	AUROC= 0.7947 (AIC= 14,861.42)	AUROC= 0.7933 (AIC= 29,430.88)

*Competing risk analysis; $ adjusted for age, gender, IMD only.
Model 1a: age, gender, and ethnicity.
Model 2a: Model 1a + ever before severity score.
Model 3a: Model 1a + 10-year severity score
Model 4a: Model 1a + 5-year severity score
AIC: Akaike information criterion; **AUROC**: area under a Receiver Operating Characteristics curve; **CV**: cardiovascular.

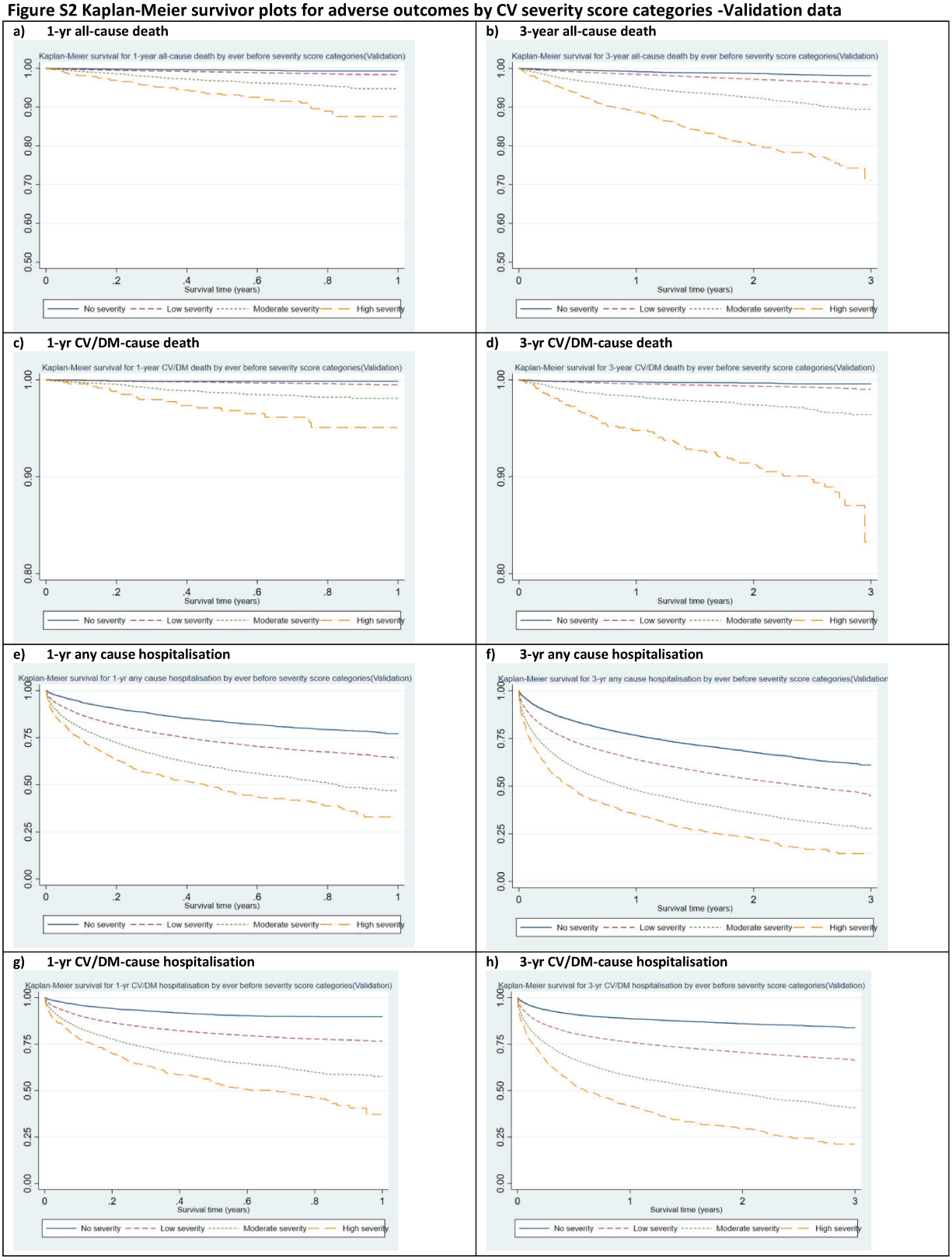
Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Figure S1 Scatter plot of estimated AIC and AUROC for models with and without severity scores for three outcomes at 1 year



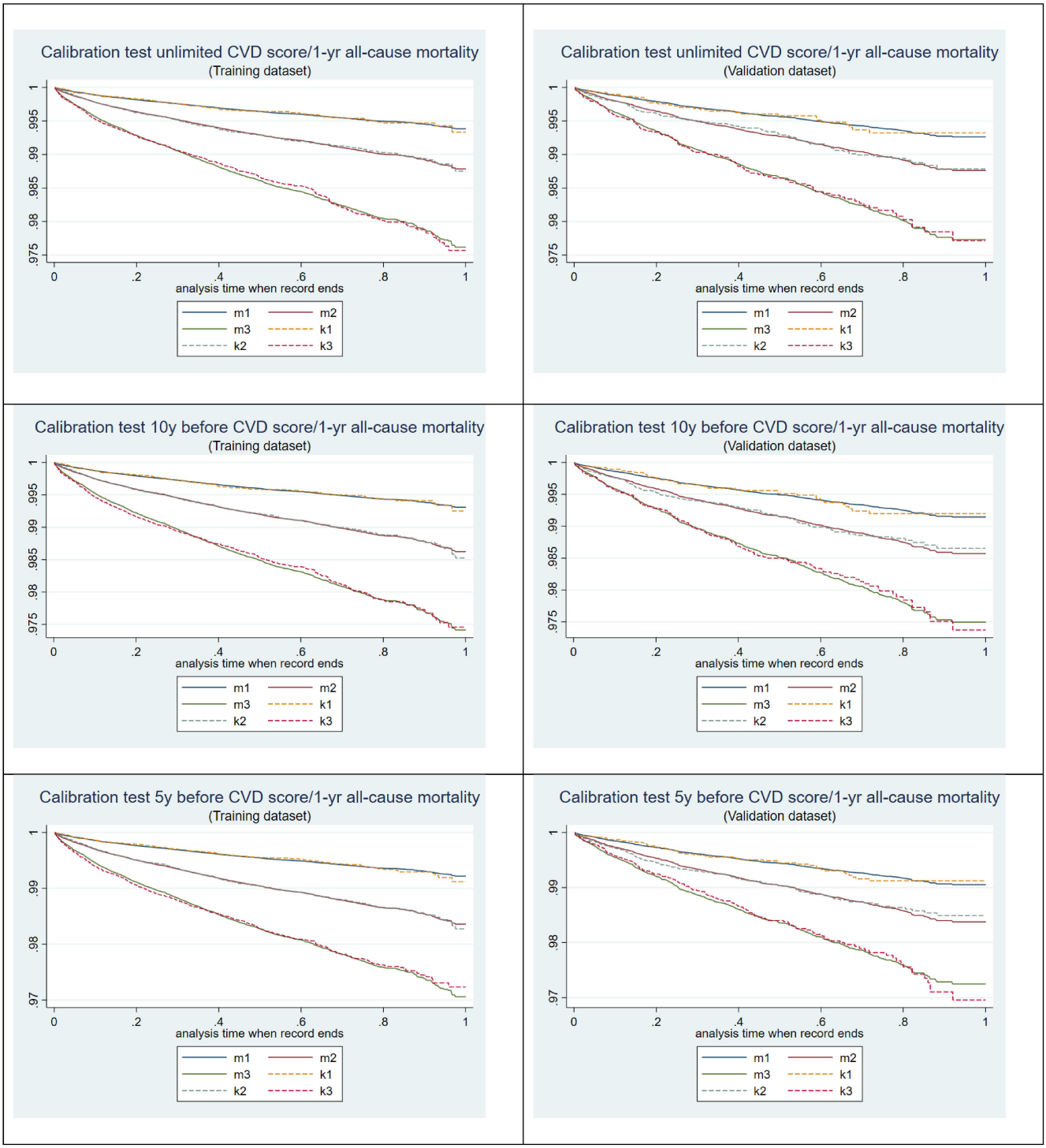
Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- *Zghebi et al. 2020*

Figure S2 Kaplan-Meier survivor plots for adverse outcomes by CV severity score categories -Validation data



Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- Zghebi et al. 2020

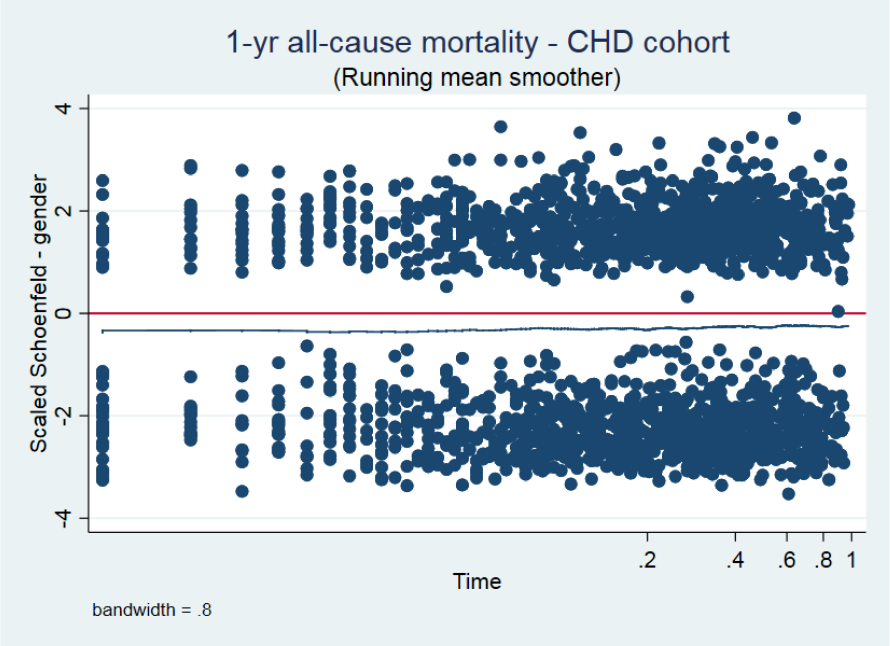
Figure S3 Calibration test by prediction of population-averaged survival probabilities for CVD scores for 1-year all-cause mortality (primary outcome) in training and validation datasets



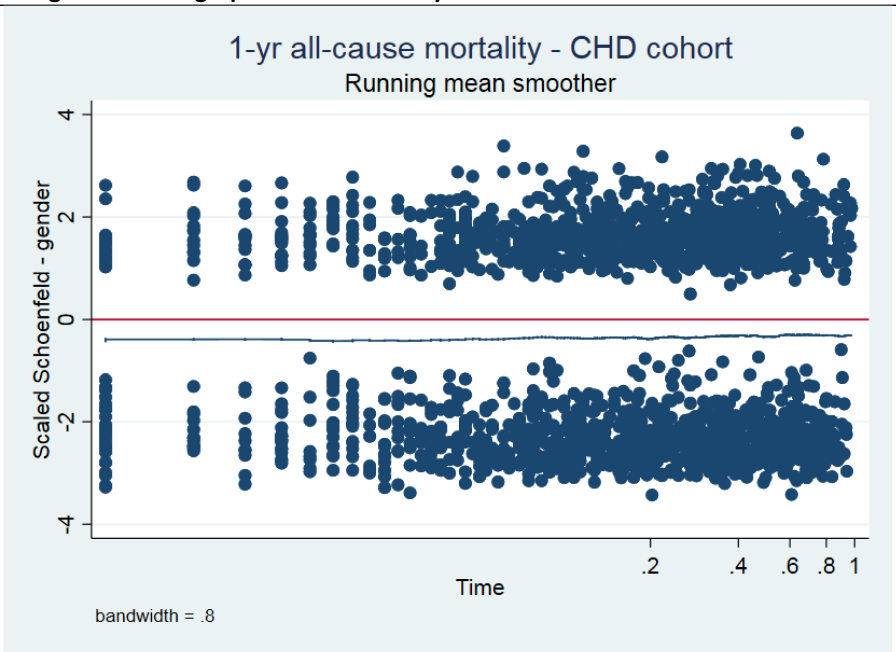
Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Figure S4 Schoenfeld residuals for testing proportional hazards of fitted survival models using unlimited CV severity score (by gender) for 1-yr all-cause mortality – training dataset

a) using unlimited CV severity score



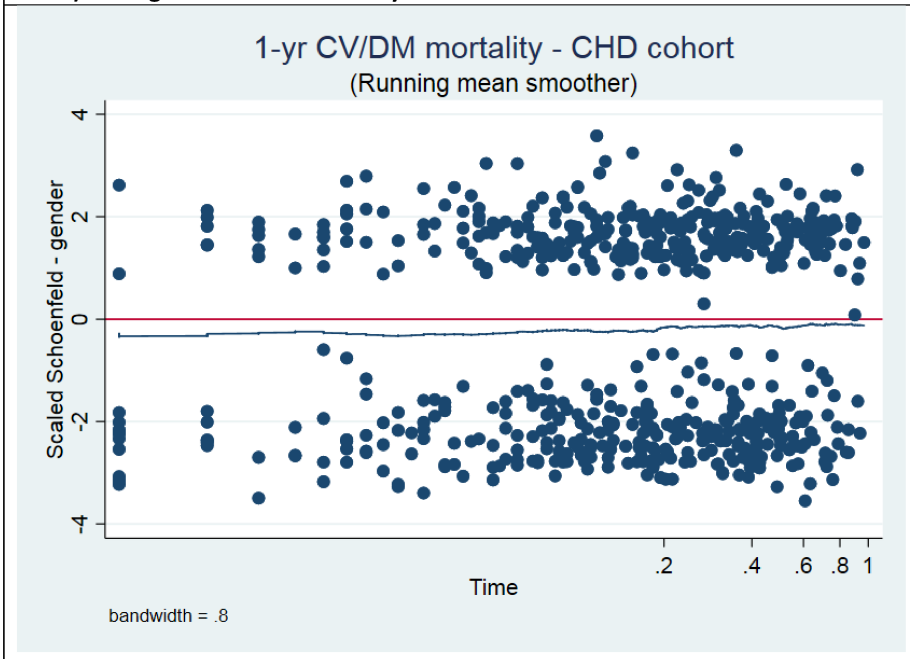
b) using socio-demographics variables only



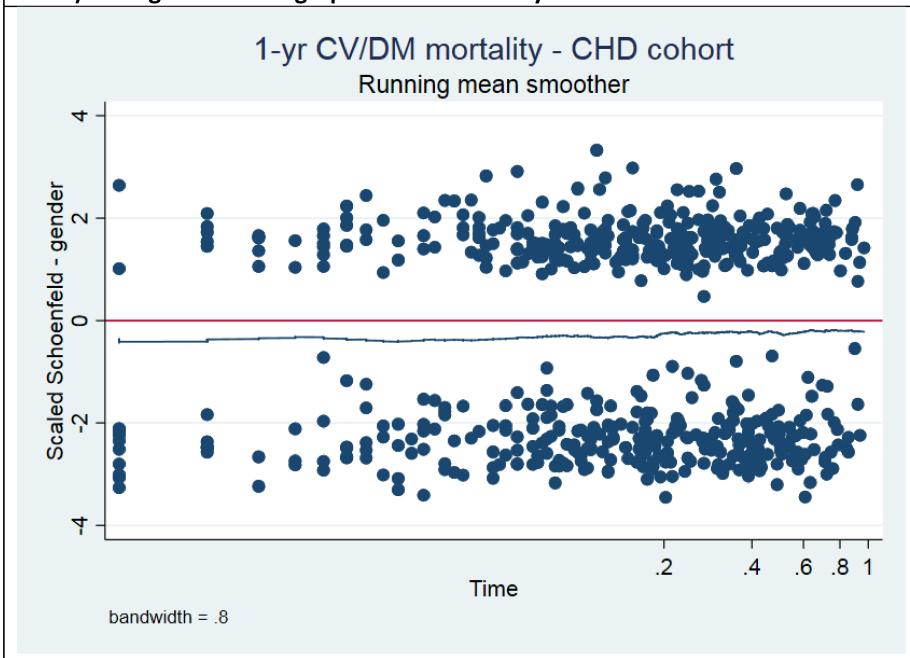
Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Figure S5 Schoenfeld residuals for testing proportional hazards of fitted survival models using unlimited CV severity score (by gender) for1-yr 1-year CV/diabetes-related mortality – training dataset

a) using unlimited CV severity score



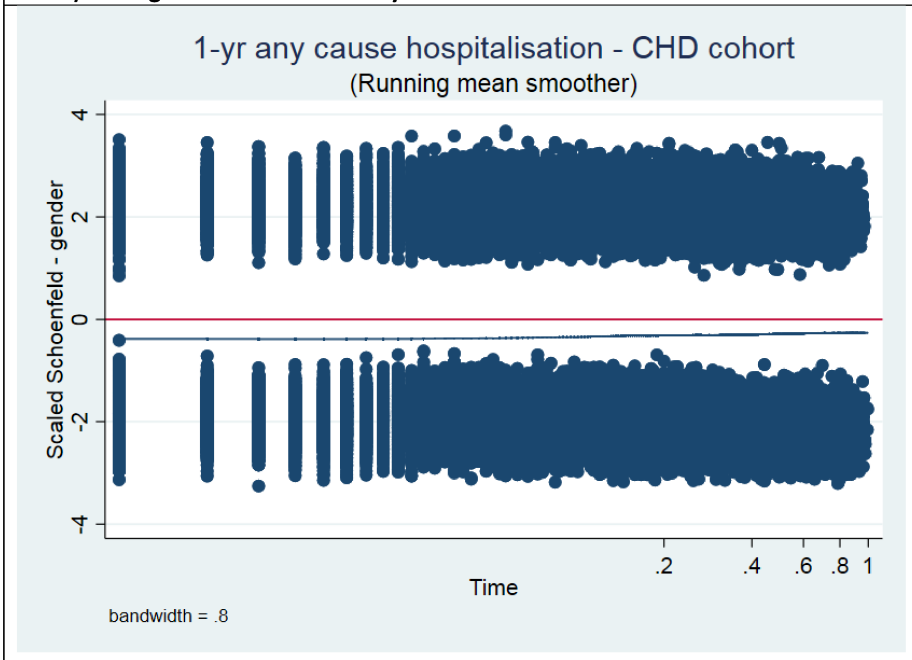
b) using socio-demographics variables only



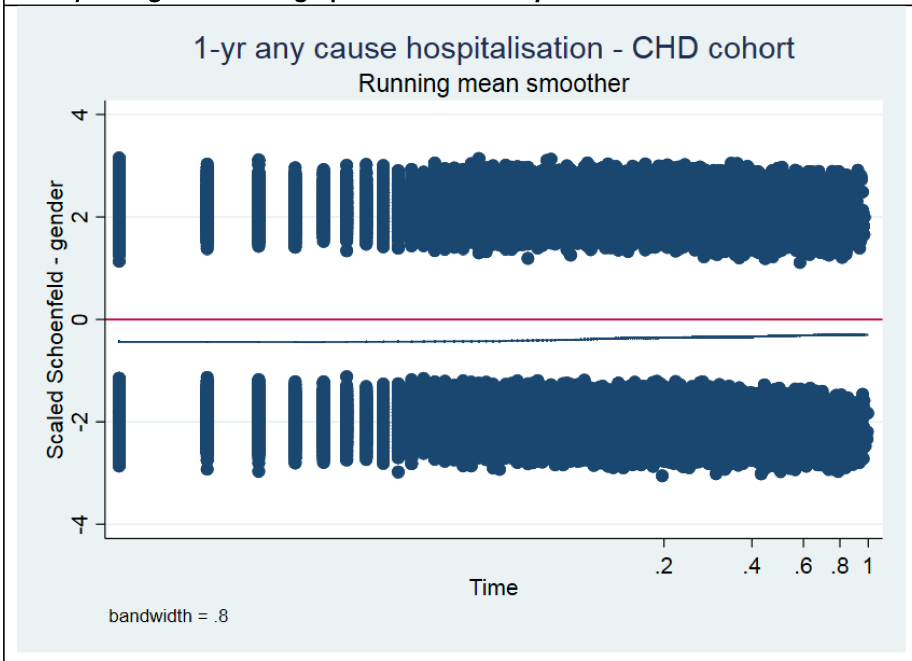
Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Figure S6 Schoenfeld residuals for testing proportional hazards of fitted survival models using unlimited CV severity score (by gender) for 1-yr any cause hospitalisation – training dataset

a) using unlimited CV severity score



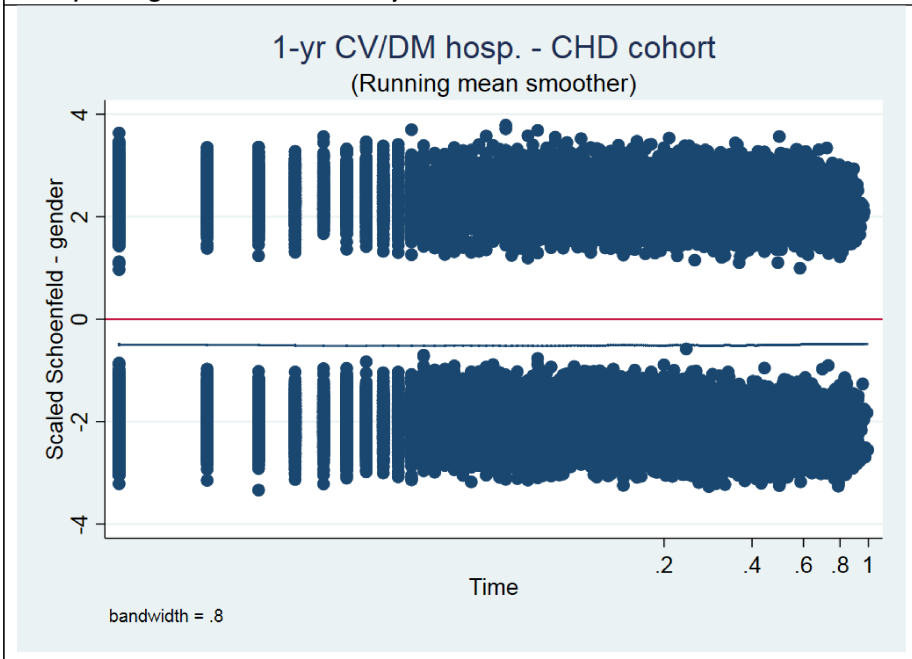
b) using socio-demographics variables only



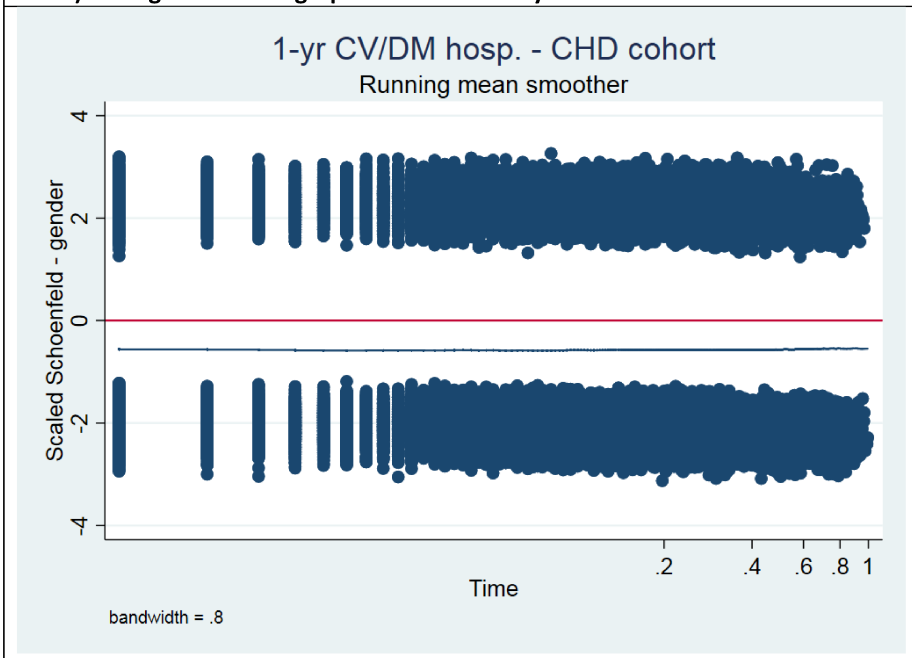
Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Figure S7 Schoenfeld residuals for testing proportional hazards of fitted survival models using unlimited CV severity score (by gender) for 1-yr CV/diabetes-related hospitalisation – training dataset

a) using unlimited CV severity score



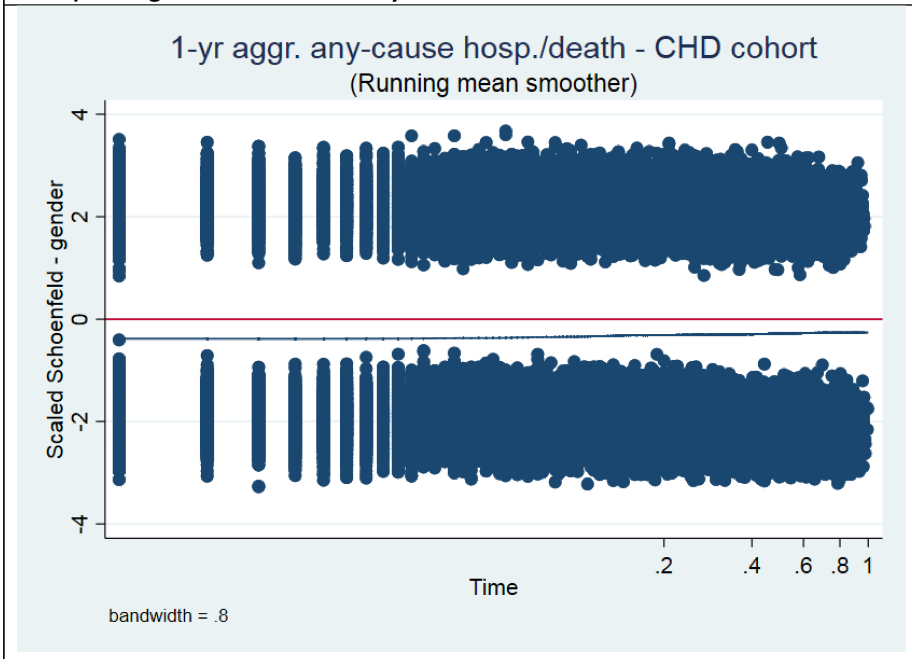
b) using socio-demographics variables only



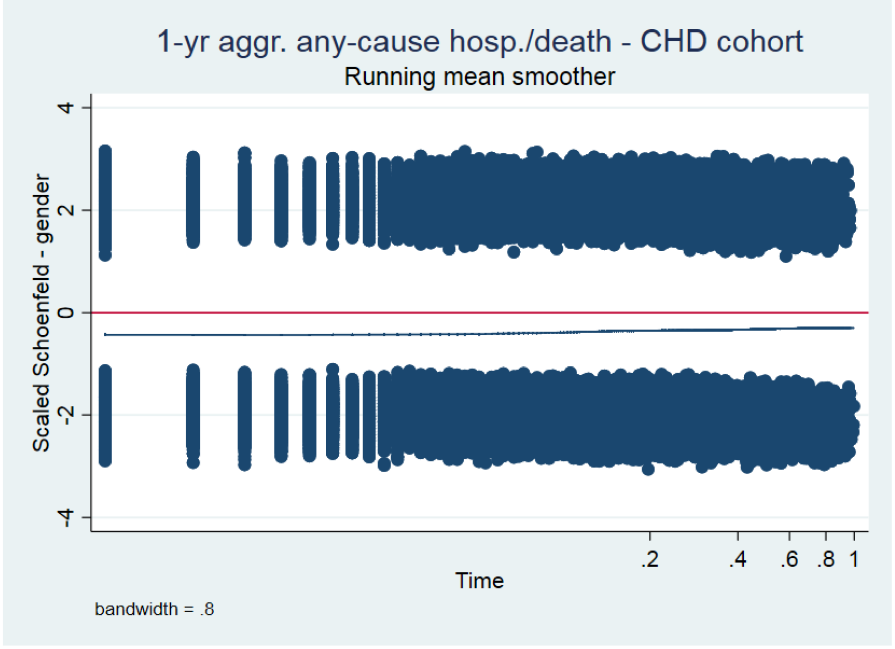
Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Figure S8 Schoenfeld residuals for testing proportional hazards of fitted survival models using unlimited CV severity score (by gender) for 1-yr aggregated any hospitalisation or death – training dataset

a) using unlimited CV severity score



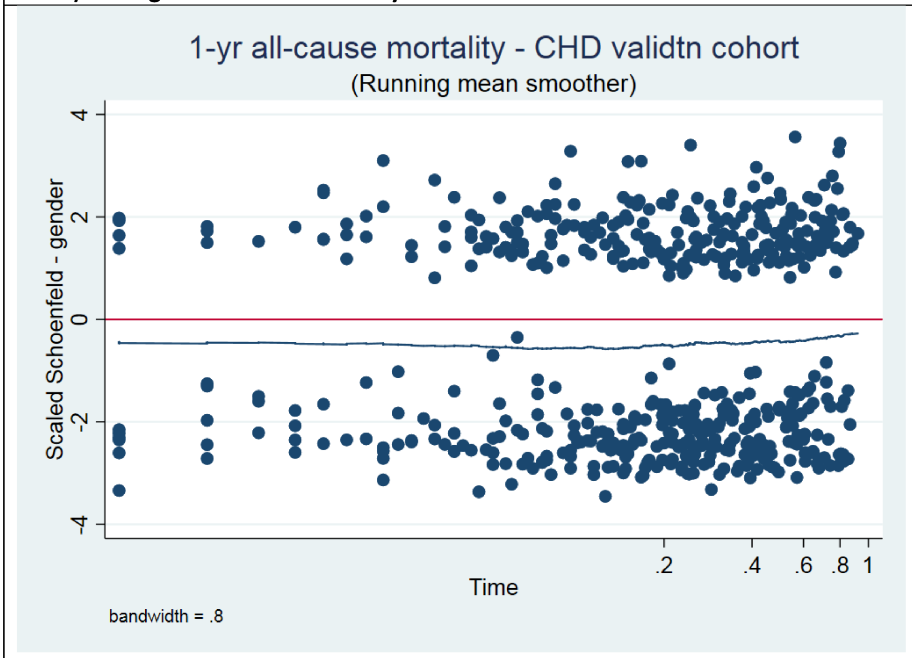
b) using socio-demographics variables only



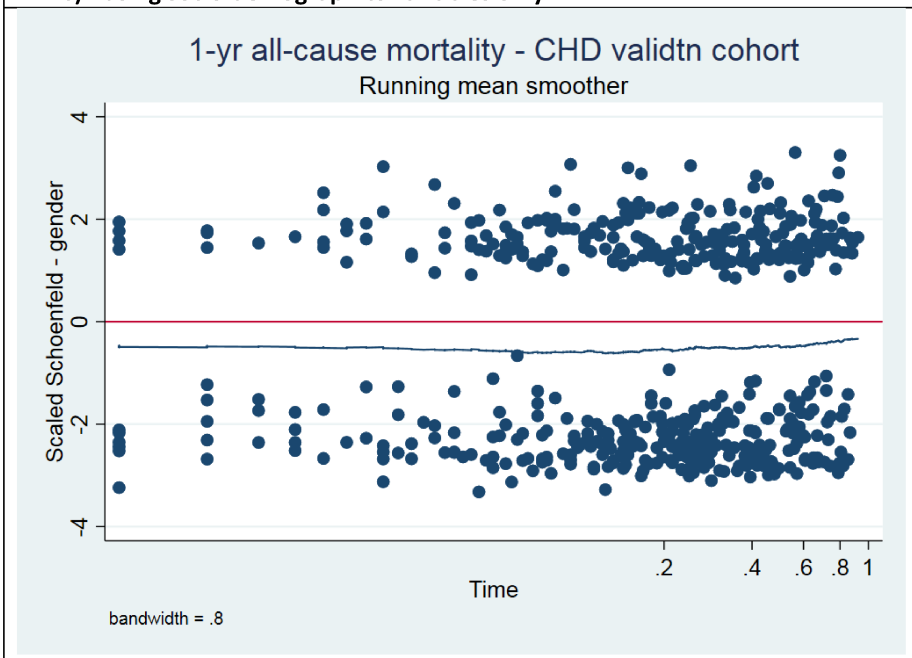
Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Figure S9 Schoenfeld residuals for testing proportional hazards of fitted survival models using unlimited CV severity score (by gender) for 1-yr all-cause mortality – validation dataset

a) using unlimited CV severity score



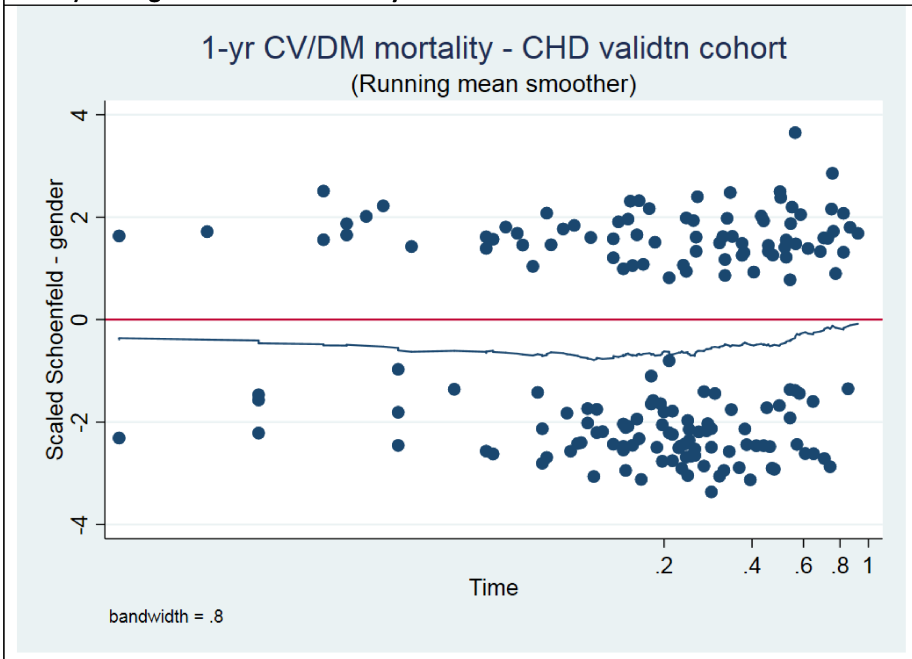
b) using socio-demographics variables only



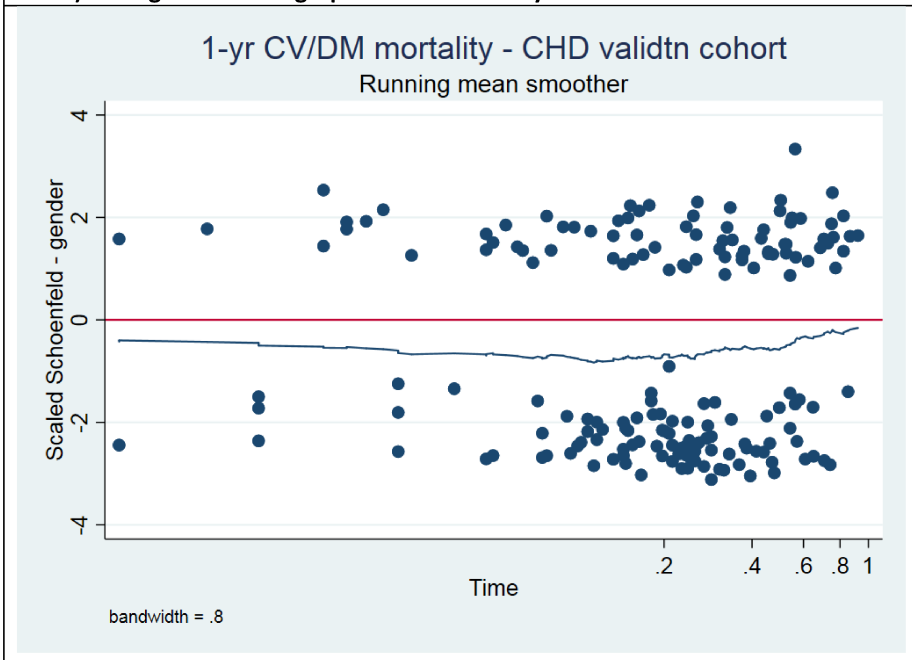
Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Figure S10 Schoenfeld residuals for testing proportional hazards of fitted survival models using unlimited CV severity score (by gender) for1-yr 1-year CV/diabetes-related mortality – validation dataset

a) using unlimited CV severity score



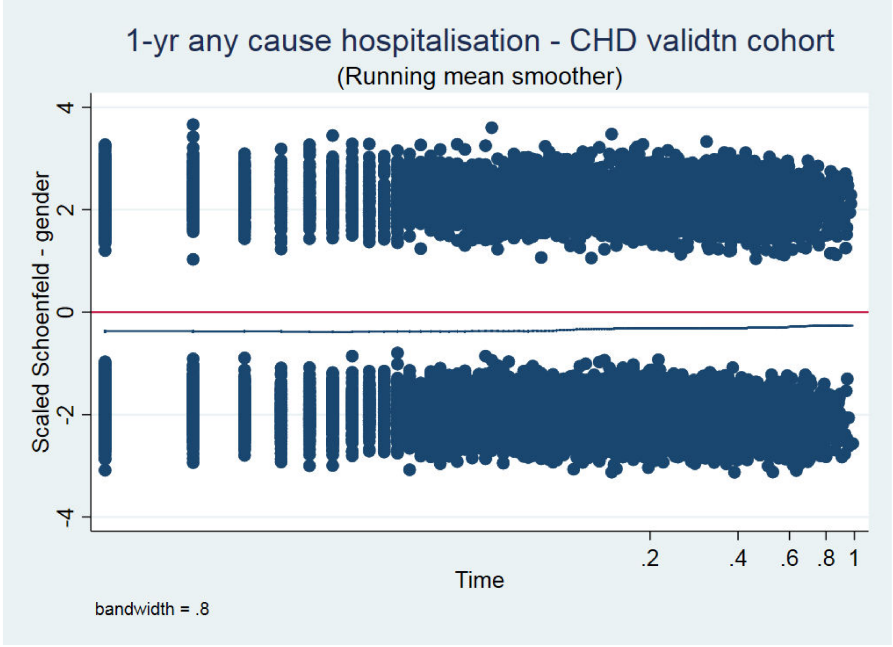
b) using socio-demographics variables only



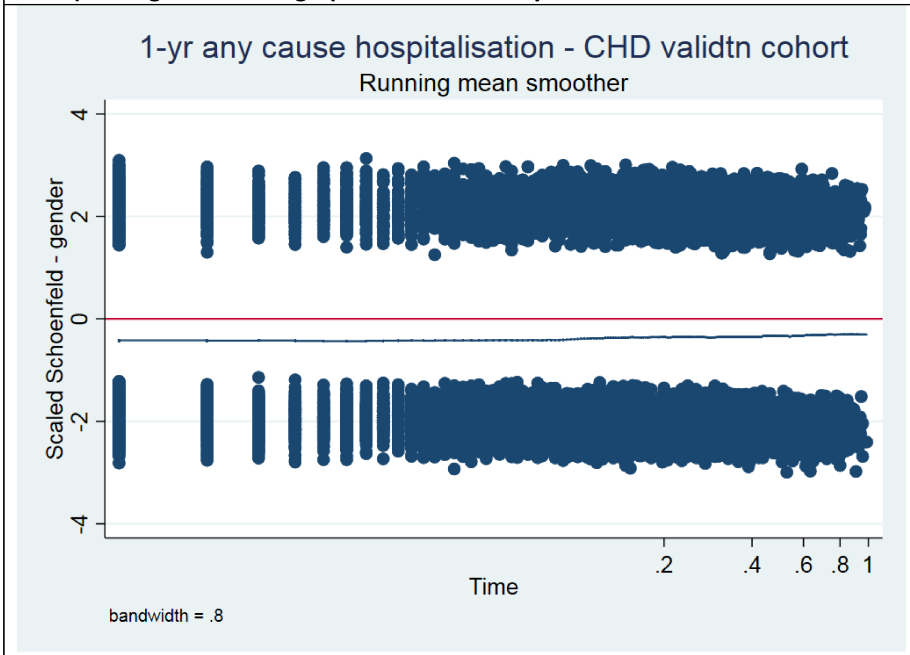
Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Figure S11 Schoenfeld residuals for testing proportional hazards of fitted survival models using unlimited CV severity score (by gender) for 1-yr any cause hospitalisation – validation dataset

a) using unlimited CV severity score



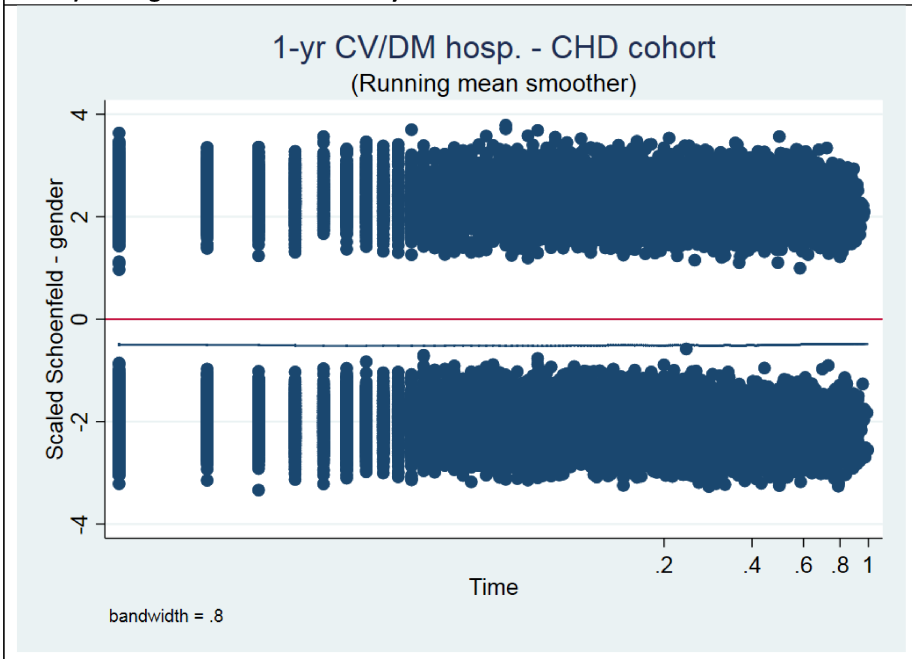
b) using socio-demographics variables only



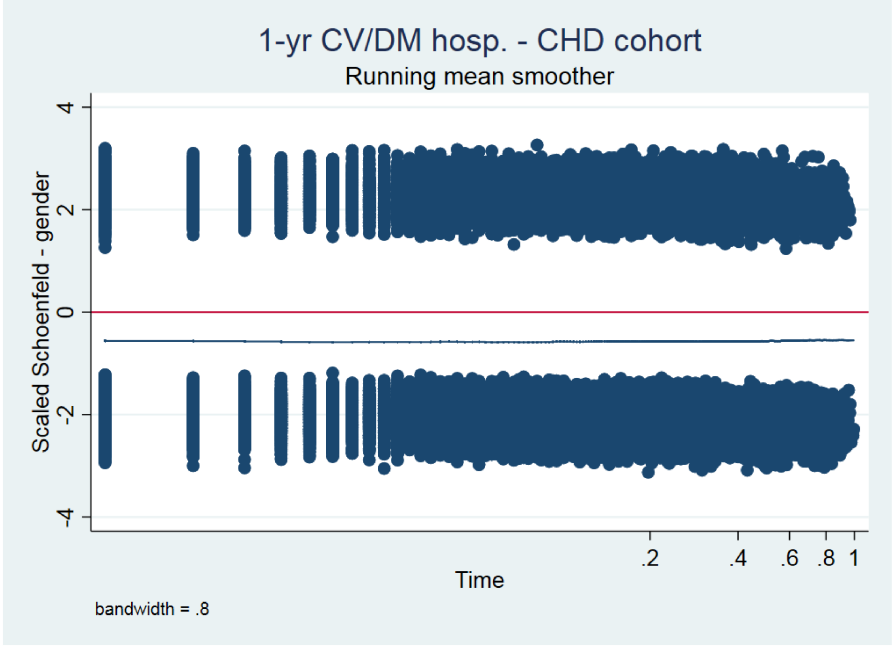
Supplementary data to *'Assessing the severity of CVD in 213,088 patients with CHD'- Zghebi et al. 2020*

Figure S12 Schoenfeld residuals for testing proportional hazards of fitted survival models using unlimited CV severity score (by gender) for 1-yr CV/diabetes-related hospitalisation – validation dataset

a) using unlimited CV severity score

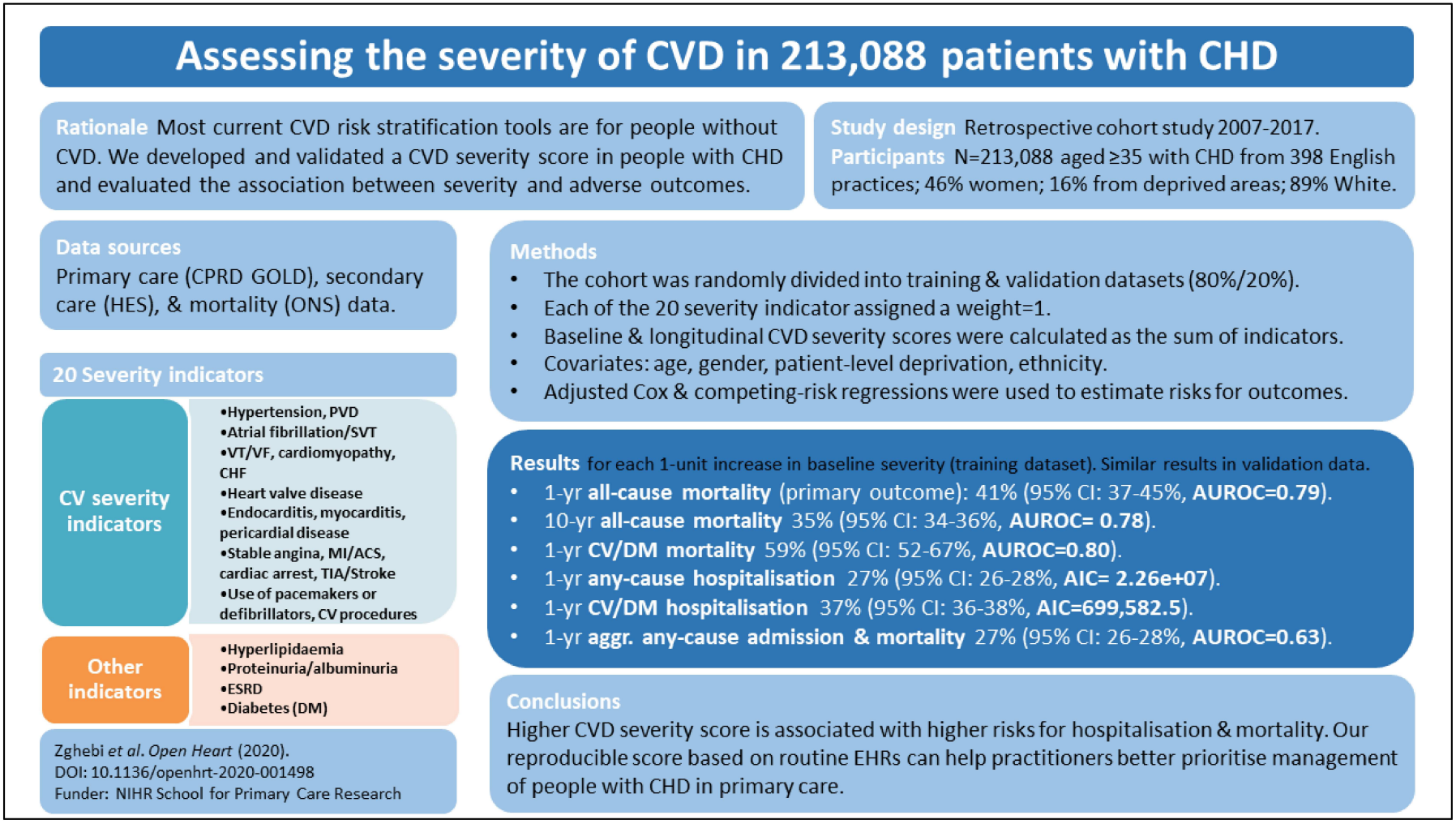


b) using socio-demographics variables only



Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- *Zghebi et al. 2020*

Figure S13 Summary figure illustrating the study methodology and main findings



Supplementary data to '*Assessing the severity of CVD in 213,088 patients with CHD*'- *Zghebi et al. 2020*

Figure S14 Visual abstract of the study