OPEN ACCESS

Original research

QRISK3 underestimates the risk of cardiovascular events in patients with COPD

Joseph Emil Amegadzie ,[1] Zhiwei Gao,[2] Jennifer K Quint ,[3] Richard Russell,[4,5] John R Hurst ,[6] Tae Yoon Lee,[1] Don D Sin ,[7] Wenjia Chen ,[8] Mona Bafadhel ,[5] Mohsen Sadatsafavi [1]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/thorax-2023-220615).

For numbered affiliations see end of article.

Correspondence to
Dr Mohsen Sadatsafavi, Faculty of Pharmaceutical Sciences, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada; msafavi@mail.ubc.ca

Received 21 June 2023
Accepted 26 October 2023

ABSTRACT

Background Patients with chronic obstructive pulmonary disease (COPD) are at increased risk of cardiovascular disease (CVD). The extent to which the excess CVD risk is captured by risk factors in QRISK, a widely used CVD risk scoring tool, is not well studied.

Methods We created an incidence cohort of diagnosed COPD patients from the United Kingdom (UK) Clinical Practice Research Datalink GOLD database (January 1998–July 2018). The outcome was a composite of fatal or non-fatal CVD events. Sex-specific age-standardised incidence ratios (SIR) were compared with values for the UK primary-care population. The observed 10-year CVD risk was derived using the Kaplan-Meier estimator and was compared with predicted 10-year risk from the QRISK3 tool.

Results 13 208 patients (mean age 64.9 years, 45% women) were included. CVD incidence was 3.53 events per 100 person-years. The SIR of CVD was 1.71 (95% CI 1.61 to 1.75) in women and 1.62 (95%CI 1.54–1.64) in men. SIR was particularly high among patients younger than 65 years (women=2.13 (95% CI 1.94 to 2.19); men=1.86 (95% CI 1.74 to 1.90)). On average, the observed 10-year risk was 52% higher than QRISK predicted score (33.5% vs 22.1%). The difference was higher in patients younger than 65 years (observed risk 82% higher than predicted).

Conclusion People living with COPD are at a significantly heightened risk of CVD over and beyond their predicted risk. This is particularly the case for younger people whose 10-year CVD risk can be >80% higher than predicted. Risk scoring tools must be validated and revised to provide accurate CVD predictions in patients with COPD.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ We know that the incidence of cardiovascular disease (CVD) is higher in patients with chronic obstructive pulmonary disease (COPD) than in the general population. But we do not know whether risk factors used in contemporary CVD risk scoring tools capture such an increased risk.

WHAT THIS STUDY ADDS

⇒ We show that a widely used CVD risk scoring tool, QRISK, substantially underestimates 10-year CVD risk in COPD. We show that such an underestimation of risk is especially pronounced in younger COPD patients.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Current CVD risk scoring tools must be validated and most likely updated to provide clinical utility for patients with COPD. COPD should be a distinct predictor in future CVD risk scoring tools.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a common chronic respiratory condition that affects over 300 million persons worldwide, is responsible for 3 million deaths per year and is the leading cause of hospitalisations globally.[1 2] In addition to respiratory symptoms, the course of COPD is characterised by the presence of various comorbidities, including cardiovascular disease (CVD). Several studies have demonstrated a higher burden of CVD in patients living with COPD compared with the general population of similar age/sex distribution.[3–5] CVD is a significant cause of morbidity and mortality in COPD, especially in those with milder disease.[6 7]

Many known risk factors are shared between COPD and CVD, such as ageing and smoking, which partly explain the coexistence of these two conditions. However, there are at least two other mechanisms by which CVD risk can be increased in COPD. First, there are likely unobserved common risk factors. Examples include shared genetic predispositions, oxidative stress, inflammation, common environmental triggers such as air pollution and viral respiratory tract infections and complex socioeconomic factors such as poverty and poor diet.[8] Yet an alternative pathway connecting the two conditions is the direct effect of COPD on CVD. For example, a recent systematic review showed that people living with COPD are highly susceptible to acute cardiovascular events during and shortly following severe exacerbations.[9]

To what extent the COPD–CVD association is mediated by unobserved risk factors or through direct influences is of clinical importance. Primary prevention of CVD is largely based on risk calculation, with most guidelines recommending preventive therapies (eg, with statins) for those with a 10-year risk ≥7.5%–10%.[10] Risk stratification is based on published risk scoring tools that are endorsed by expert panels, and many are incorporated into

Check for updates

To cite: Amegadzie JE, Gao Z, Quint JK, *et al*. *Thorax* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/thorax-2023-220615

electronic medical records and clinical decision support systems. An example is the QRISK scoring tool, which is widely used in the United Kingdom (UK) and other countries.[11] The latest version of the tool, QRISK3, has been developed based on data from 7.89 million patients from 981 general practices in the UK and has been validated among 2.67 million patients from 328 separate practices.[11] It includes up to 21 predictors relating to demographics, medical diagnoses, social deprivation, blood markers and health behaviour, making it one of the most comprehensive CVD risk assessment tools currently available. The influential UK's National Institute for Health and Care Excellence recently updated its guidelines on CVD risk assessment and reduction, strongly emphasising the use of QRISK3 for risk stratification.[12]

However, QRISK and other CVD risk scoring tools are developed based on data from general population. As such, they will generate valid predictions in patients living with COPD only if the entirety of increased CVD risk is captured by risk factors included in the tool. Any significant violation of this assumption will result in the underestimation of CVD risk, leading to missed opportunities to mitigate the risk of CVD in patients with COPD. To what extent commonly known predictors in contemporary CVD risk scoring tools explain the added CVD risk in COPD is currently not well understood.

To address this knowledge gap, in this study, we estimated CVD risk in a population-based primary care cohort of patients with COPD and compared it with predicted risk based on QRISK3 as an exemplary CVD risk scoring tool.[11] We hypothesised that due to the contribution of several unobserved risk factors and shared disease pathways, the CVD risk in patients living with COPD is higher than in the general population, and that QRISK3 will underestimate the true risk of CVD events in patients living with COPD.

METHODS

Study design and population

We conducted a retrospective cohort study of patients with COPD from the UK Clinical Practice Research Datalink (CPRD) GOLD primary care database between 1 January 1998 and 31 July 2018 (the study period). The CPRD consists of anonymous, longitudinal data from over 14 million individuals across ~700 general practices in the UK.[13] It includes data on medical symptoms and diagnoses, demographic information, specialist referrals, test results, prescriptions and lifestyle details such as smoking and alcohol consumption. As QRISK is developed based on the UK primary care data, it is expected to provide accurate predictions in the general sample from this data set. Indeed, previous versions of QRISK have been validated in CPRD, with the authors concluding that the performance of the tool was comparable with its performance in the original development sample.[14]

Patients who met the following criteria were included: they were permanently registered with a primary care practice, provided data for at least 1 year and had their data linked to UK Hospital Episode Statistics discharge data, the Office of National Statistics death data, and Index of Multiple Deprivations (a measure of socioeconomic status) within the study period. The data custodian, the independent Scientific Advisory Committee of the UK General Practice Research Datalink, approved our study protocol for anonymised research purposes, which required no human consent to participate.

Using a validated case definition,[15] we created an incidence cohort of COPD patients aged between 40 and 84 years. The incidence cohort approach was chosen because the time of COPD diagnosis is a natural vantage point for assessing CVD risk as recommended by guidelines.[16] The case definition for COPD (details in online supplemental table 1) has a positive predictive value of 86.5% in CPRD.[15] The upper age bound was placed to be aligned with the inclusion criteria of QRISK3 development sample, while the lower age bound was placed to reduce the risk of including patients with asthma. To be classified as an incident case, the initial COPD diagnosis must have been made after at least 1 year of data availability, with no COPD-related health resource utilisation during this period.

The *index date* was the date of COPD diagnosis. COPD patients were excluded if they had a missing Index of Multiple Deprivation, pre-existing CVD or were prescribed statins before the index date, in line with the intended use of QRISK (and the exclusion criteria of its development study). Follow-up was censored at the earliest recorded date of CVD, death from any cause, emigration outside a CPRD practice site or the end of the study period (31 July 2018), whichever occurred first.

Outcomes

CVD outcomes were defined following the development approach for the QRISK3 algorithm.[11] They included the first recorded episode of fatal or non-fatal coronary heart disease, ischaemic stroke, or transient ischaemic attack (as a composite endpoint). These events were determined from the GP database using Read codes (*see* online supplemental table 2) and from hospitalisation and mortality outcomes based on International Classification of Diseases, 10th revision codes (see online supplemental table 3).

Predictors

All predictors were ascertained according to the methodology used in the original QRISK3 development procedure (predictors are listed in table 1).[11] Accordingly, we obtained the most recent values of blood pressure and smoking status recorded before the index date and all comorbidities recorded at the baseline visit. We selected the closest value to cohort entry for laboratory test values such as total and high-density lipoprotein. The use of medications (corticosteroids and second-generation 'atypical' antipsychotics, which are predictors in QRISK3) at baseline was defined as at least two prescriptions, the most recent being not more than 28 days before the index date. Details on the evaluations of predictors and the measurement time windows are shown in online supplemental table 4.

Statistical analyses

We calculated the raw incidence rates per 100 person-years as the total number of CVD events divided by the total person-years of follow-up, multiplied by 100. To compare the CVD incidence between people with COPD and the QRSIK3 development cohort, the baseline sample of the latter cohort was used as the reference to calculate age-standardised (with 1-year bins) incidence ratios (SIR). SIRs were reported separately by sex and younger (≤65) versus older adults (65+ years of age). The chosen age cut-off was employed to facilitate meaningful comparisons of SIRs across age groups and is the cut-off frequently used to separate younger and older adults with COPD.[17 18]

We calculated 10-year observed risk using Kaplan-Meier analysis, in the whole cohort and in the subgroups defined by sex and age groups. We assessed the accuracy of QRISK3 by calculating the observed versus predicted 10-year CVD risk in the entire cohort and across subgroups. Aside from the Deprivation Index for which missingness was an exclusion criterion, we used

Table 1 Baseline characteristics of the included sample

Characteristics (QRISK3 predictors)	COPD patients (40–84 years) n=13 208	
	Women 5884 (44.5%)	Men 7324 (55.5%)
Age: mean (SD)	64.9 (10.4)	64.7 (10.2)
Index of deprivation: N (%)		
Least deprived	831 (14.1)	1105 (15.1)
Less deprived	1099 (18.7)	1443 (19.7)
Deprived	1204 (20.4)	1417 (19.3)
More deprived	1298 (22.1)	1624 (22.2)
Most deprived	1452 (24.7)	1735 (23.7)
Body mass index (kg/m^2): mean (SD)	25.7 (7.0)	25.7 (5.6)
% recorded	*96.5*	*95.6*
Total cholesterol recorded (mmol/L): mean (SD)	5.7 (1.5)	5.3 (1.2)
% recorded	*54.8*	*53.6*
HDL cholesterol recorded (mmol/L): mean (SD)	1.6 (1.0)	1.4 (1.0)
% recorded	*42.7*	*40.9*
Total cholesterol/HDL cholesterol ratio: mean (SD)	3.8 (1.2)	4.1 (1.3)
% recorded	*42.6*	*40.8*
Systolic blood pressure (mm Hg): mean (SD)	138.6 (18.2)	139.8 (17.3)
% recorded	*86.6*	*82.3*
Mean (SD) SBP variability (mm Hg)*	12.8 (7.8)	12.4 (8.7)
% recorded	*66.8*	*57.1*
Ethnic origin:		
White or not stated	5832 (94.9)	6973 (95.2)
Bangladeshi	†	†
Black African	†	†
Black Caribbean	†	17 (0.2)
Chinese	†	†
Indian	9 (0.2)	16 (0.2)
Other Asian	6 (0.1)	8 (0.1)
Pakistani	†	12 (0.2)
Other	279 (4.7)	290 (4.0)
Smoking status:		
Former smoker	1404 (10.6)	2035 (15.4)
Heavy smoker (20 cigs per day or over) ‡	1075 (8.1)	1459 (11.0)
Moderate smoker (10 to 19 cigs per day) ‡	1111 (8.4)	1021 (7.7)
Light smoker (less than 10 cigs per day) ‡	823 (6.2)	889 (6.7)
Non-smoker	586 (4.4)	559 (4.2)
% recorded	*85*	*81.4*
Medical characteristics: N (%)§		
Corticosteroid use	293 (5.0)	270 (3.7)
Atypical antipsychotic use	9 (0.2)	9 (0.1)
Family history of CHD	424 (7.2)	419 (5.7)
Chronic kidney disease	45 (0.8)	44 (0.6)
Rheumatoid arthritis	152 (2.6)	117 (1.6)
Atrial fibrillation	36 (0.6)	46 (0.6)
Migraine	443 (7.5)	202 (2.8)
Treated hypertension	1410 (24.0)	1427 (19.5)
SLE	18 (0.3)	†
Erectile dysfunction	NA	726 (9.9)
Severe mental illness	179 (3.0)	187 (2.6)
Type 1 diabetes	†	14 (0.2)
Type 2 diabetes	78 (1.3)	147 (2.0)
HIV/AIDS	0	0

*Based on a SD of ≥2 systolic blood pressure values.
†Cells with fewer than five recordings are not shown per the confidentiality policies of the Clinical Practice Research Data.
‡Smoking status: light smoker (<10 cigs), moderate smoke (10–19 cigs), heavy smoker (≥20 cigs).
§Assumed not present if not recorded.
AIDS, acquired immunodeficiency syndrome; CHD, coronary heart disease; COPD, chronic obstructive pulmonary disease; HDL, high density lipoprotein; HIV, human immunodeficiency virus; NA, not applicable; SBP, systolic blood pressure; SLE, systemic lupus erythematosus.

multiple imputations by chained equations to impute the value of other missing predictors. We created 10 independent data sets each with randomly imputed data, performed all the analyses separately within each data set and combined the results using Rubin's rule.[19]

Cohort creation and predictor and outcome extraction were conducted in SAS (V.9.4, SAS Institute, Cary, North Carolina). Statistical analyses related to missing value imputation, SIR calculation and Kaplan-Meier analyses were performed in R (V.4.2.2).[20] QRISK3 R package was used to calculate QRISK scores.[21] The statistical code is available from https://github.com/resplab/papercode/tree/main/COPD_CVD.

RESULTS

The original data set included 29 605 patients with obstructive airway disease, of whom 13 208 satisfied the case definition of COPD, had 1 year of data without COPD-related events before diagnosis, were free of a statin prescription on COPD diagnosis, had non-missing Deprivations Index and were without CVD at baseline. Figure 1 illustrates the cohort creation steps.

Table 1 summarises the characteristics of patients in the final data set. A slight majority (55%) were men; the average age at diagnosis was 64.9 years, with 48.7% younger than 65 years of age. The distribution of risk factors was generally similar between women and men (table 1).

Incidence

The incidence of CVD was 3.53 (95% CI 3.41 to 3.68) events per 100 person-years. This value was 74.6% higher in men than in women (4.01 (95% CI 3.81 to 4.21) vs 2.99 (95% CI 2.81 to 3.18)). In comparison, CVD incidence in the QRISK3 development population for individuals 40–84 years of age was 1.20 (95% CI 1.19 to 1.21) in total, 1.37 (95% CI 1.36 to 1.38) in men, and 1.02 (95% CI 1.01 to 1.03) in women.[11] The incidence rates of CVD in patients living with COPD were higher than those in the general population across all sex and age groups (figure 2; detailed results in online supplemental table 5). Overall, for both the COPD and the general population, CVD incidence increased with age, but the association between age and CVD incidence was sharper among patients with COPD.

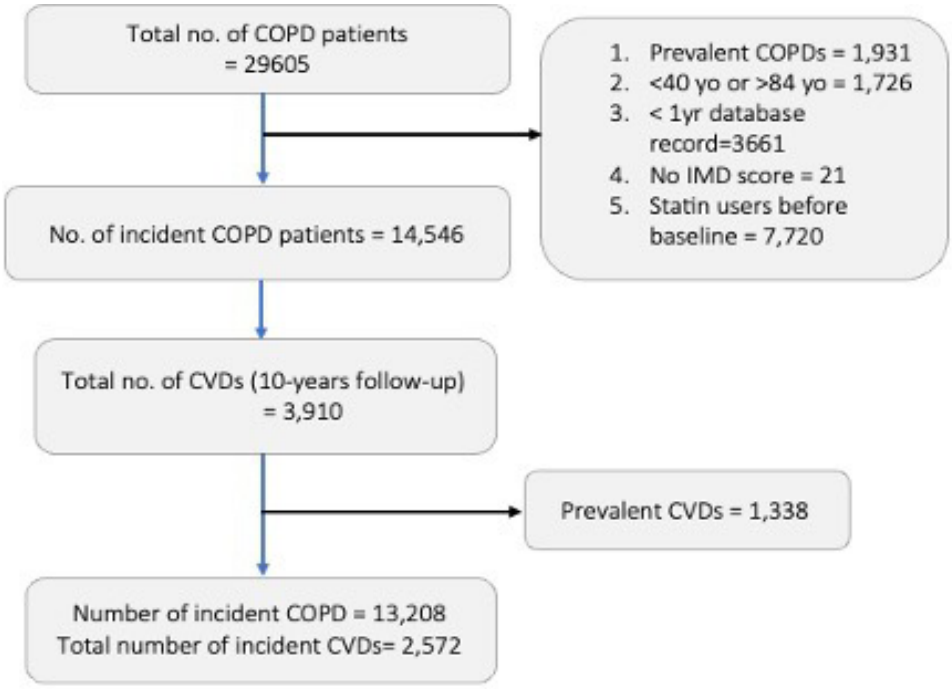

Figure 1 COPD patient selection and CVD outcome incidence. COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; IMD, Index of Multiple Deprivation.

CVD incidence was >3 times higher among COPD patients ≤65 years of age than their general population counterparts; this ratio was <1.5 among patients >65 years of age.

Standardised incidence ratios

When standardised to the QRISK3 development sample, the SIR was 1.62 (95% CI 1.54 to 1.64) among men and 1.71 (95% CI 1.61 to 1.75) among women. Again, there was a significant age effect in standardised CVD incidence between COPD and the general population. Among men, the SIR was 1.86 (95% CI 1.74 to 1.90) among those ≤65 years of age at diagnosis and 1.38 (95% CI 1.28 to 1.42) among older patients. Among women, the corresponding values were 2.13 (95% CI 1.94 to 2.19) and 1.46 (95% CI 1.33 to 1.50). Detailed results are provided in table 2.

Predicted versus observed 10-year CVD risk

The predicted and observed CVD risk, in total and by sex and age groups, are provided in table 3. The average predicted 10-year CVD risk was 22.1% (95% CI 21.8% to 22.4%) among patients with COPD. In comparison, the observed 10-year CVD risk was 33.5% (95% CI 32.3% to 34.7%). As such, the observed risk was 1.52 times higher than the predicted risk (p<0.001).

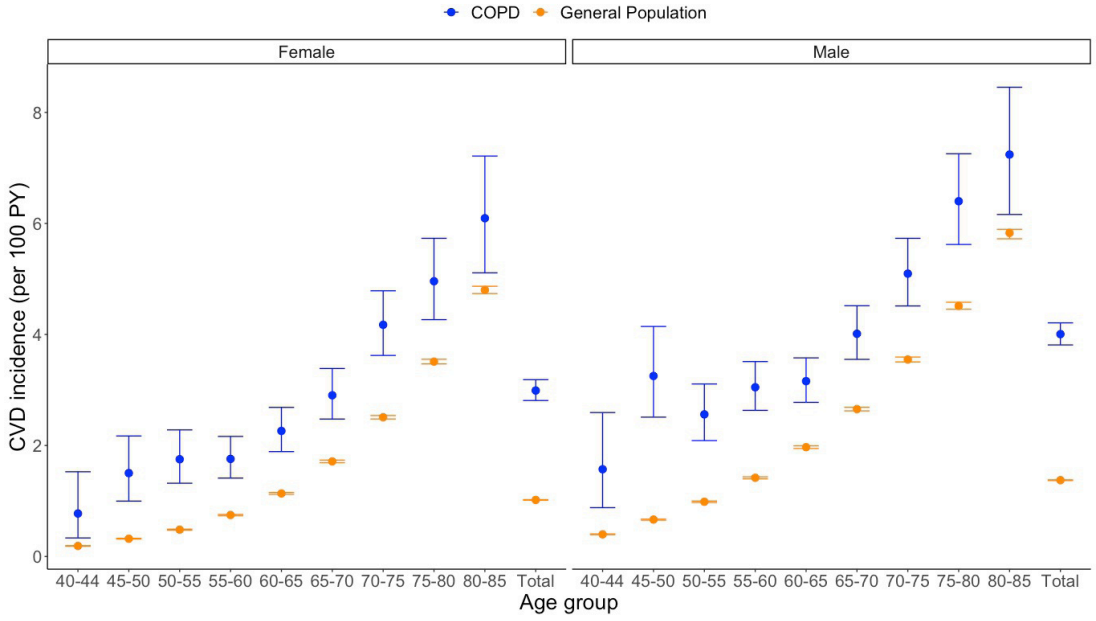

Figure 2 Incidence rates (95% CIs) by sex and age groups for COPD (blue) and the general population (orange). COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; PY, person-years.

Table 2 Standardised incidence ratio (SIR) between COPD and general primary care population

Group	Incidence in COPD cohort X 10000	Incidence in the QRISK3 development cohort X 10000 (Hippisley *et al*.[11])	Incidence ratio	SIR, 95%CI
All	354.2	118.9	2.98	1.65 (1.59 to 1.68)
Male	401.6	137.5	2.91	1.62 (1.54 to 1.64)
Female	299.3	101.9	2.94	1.71 (1.61 to 1.75)
Younger male (≤65 y/o)	320.9	107.2	2.99	1.86 (1.74 to 1.90)
Older male (65+y/o)	594.2	420.8	1.41	1.38 (1.28 to 1.42)
Younger female (≤65 y/o)	210.6	63.0	3.34	2.13 (1.94 to 2.19)
Older female (65+y/o)	483.4	329.8	1.47	1.46 (1.33 to 1.50

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; SIR, standardised incidence ratio.

The discrepancy between the predicted and observed risks was particularly pronounced in younger patients, with an observed risk that was 1.82 times higher than the predicted (p<0.001; observed: 25.7% (95% CI 24.2% to 27.2%), predicted: 14.1% (95% CI 13.9% to 14.3%)). In comparison, the observed risk was 1.38 times higher than predicted in older patients (p<0.001; observed: 42.6% (95% CI 40.7% to 44.5%), predicted: 30.8% (95% CI 30.5% to 31.3%)). Average predicted 10-year risk in women was 18.9% (95% CI 18.5% to 19.2%), while the observed risk was 29.5% (95% CI 27.8 to 31.3%), corresponding to an observed/predicted ratio of 1.56 (p<0.001). Average predicted 10-year risk in men was 24.7% (95% CI 24.4% to 25.1%) and the observed risk was 36.7% (95% CI 35.0% to 38.4%), giving rise to an observed/predicted ratio of 1.49 (p<0.001).

DISCUSSION

We used a representative sample of UK English primary care patients with COPD to evaluate the incidence and 10 year risk of CVD at the time of COPD diagnosis and determined to what extent QRISK3 can accurately predict CVD risk in people living with COPD. We showed that for both men and women, the incidence of CVD in people with COPD is substantially higher than in the general population, particularly among younger people with COPD. As an example, the annual incidence of CVD among 55-year-old individuals with COPD diagnosis was similar to the annual incidence among 70-year-olds in the general population. This gap narrowed by age, with the annual incidence of CVD among 70-year-old individuals on COPD diagnosis close to the annual incidence among 75-year-old members of the general population.

Critically, we showed that the QRISK3 score captures only part of the excess risk of CVD in patients with COPD. According to QRISK3, the 10-year CVD risk was, on average, 22.1% at the time of COPD diagnosis. However, at 33.5%, the observed 10-year risk was >50% higher than the predicted. Again, the discrepancy between the predicted and observed risk was particularly pronounced in younger individuals, with actual risk being >80% higher than predicted in adults 65 years of age or younger at the time of COPD diagnosis. One potential mechanism is that owing to genetic variations and/or environmental exposures, many individuals who develop COPD at a younger age have reduced peak lung growth. Reduced lung size has been causally linked with coronary artery disease and CVDs, though the exact mechanism(s) by which this occurs remain obscure.[22] The QRISK scores do not consider genetic or many salient early life environmental factors, which may have outsized roles in the pathogenesis of COPD and CVD in younger persons.

Our findings of elevated CVD risk in patients with COPD are consistent with existing studies. A recent population-based study from the province of Ontario, Canada, documented a two-fold higher sex-standardised and age-standardised risk of major CVD events in COPD compared with the general population.[23] After controlling for a large set of classical CVD risk factors, COPD was still associated with CVD, with an adjusted HR of 1.25 (95% CI 1.23 to 1.27). Another retrospective Canadian study found that the risk of hospitalisation due to various cardiovascular causes was more than twofold higher in the COPD group than in the comparison group.[24] A meta-analysis of observational data has shown that adults with COPD are at a higher risk of developing various CVD events and cardiovascular risk factors such as hypertension and diabetes.[5] Furthermore, a US-based study found that individuals with COPD had a significantly higher prevalence of CVD, with COPD increasing the odds of developing CVD by 170%.[25] However, the utility of existing global CVD risk tools in predicting CVD events in patients with COPD has not been well studied. As such, our assessment of QRISK provides novel insight into the discrepancy between the predicted and observed CVD risk in COPD, with important clinical and research implications.

These results have important clinical implications. CVD risk assessment is often neglected in COPD patients.[26] However, our results show that even when CVD risk is assessed, it

Table 3 Predicted (by QRISK3) versus observed 10-year CVD risk and observed/predicted ratio, in total and by sex and age groups

Group	Predicted risk (95% CI)	Observed risk (95%CI)	Observed/predicted ratio*
All	22.1% (21.8% to 22.4%)	33.5% (32.3% to 34.7%)	1.52
Male	24.7% (24.4% to 25.1%)	36.7% (35.0% to 38.4%)	1.49
Female	18.9% (18.5% to 19.2%)	29.5% (27.8 to 31.3%)	1.56
Younger (≤65 y/o)	14.1% (13.9% to 14.3%)	25.7% (24.2% to 27.2%)	1.82
Older (65+y/o)	30.8% (30.5% to 31.3%)	42.6% (40.7% to 44.5%)	1.38

*All p values comparing predicted vs observed risk were significant (<0.001).
y/o, years old.

underestimates the true risk, leading to potential missed opportunities for CVD risk modification through lifestyle interventions and therapeutics. This underestimation is particularly the case among individuals diagnosed with COPD earlier in their lives, who presumably stand to benefit the most from primary CVD prevention. Underestimated risks also hinder informed shared decision-making. Indeed, the magnitude of risk can be a motivator for behaviour change.[27]

Our results also have important implications for research. QRISK is not the only commonly used CVD risk scoring tool, and the validity of similar tools in COPD patients need to be rigorously examined. Other major scoring tools such as Framingham[28] and ASCVD (AtheroSclerotic CardioVascular Disease)[29] risk scores have fewer predictors and thus are potentially more prone to underestimating risk in COPD. More broadly, given the high prevalence of COPD and the high incidence of CVD, the inclusion of COPD as a predictor in CVD risk scoring tools should be seriously considered.

QRISK includes predictors related to other conditions such as rheumatoid arthritis, chronic kidney disease and migraine. Lack of consideration of COPD in contemporary CVD risk scoring tools probably reflects the general lack of awareness about the prevalence of airway diseases and an old but persistent notion that COPD is a disease of smokers and the inclusion of smoking as a risk factor is sufficient. We do not criticise the creators of CVD risk scoring tools, as it is ultimately the responsibility of the respiratory community to raise awareness around the burden and multifaceted pathophysiology of airway diseases. We also note that the choice of predictors in common CVD risk scoring tools is based on their discovered associations with CVD in the general population. Other predictors (eg, coronary artery calcium score[30]) might prove valuable in patients with COPD. However, given the significant underestimation of risk, a more immediate step while awaiting the next generation of CVD risk scoring tools is to update existing tools to generate calibrated CVD risk for patients with COPD. This has been undertaken for short-term CVD outcomes. For example, Rothnie *et al* reported that Global Registry of Acute Coronary Events underestimates 6-month mortality after acute coronary events in patients with COPD.[31] They found that correcting the score by a multiplicative factor of 1.3 significantly improves its calibration in COPD.

The strengths of our study include access to a large, representative primary care cohort of people with COPD and a comprehensive assessment of CVD risk factors and outcomes, which was faithful to the definitions used to develop and validate QRISK. The relatively large sample size enabled us to investigate CVD incidence and predicted and observed risks within important patient subgroups with high statistical power, resulting in unequivocal interpretations about the performance of QRISK. However, our study also has limitations. This was not a dedicated validation study of QRISK. Such studies are required to follow specific methodology to evaluate the discrimination, calibration and clinical utility of a scoring tool, and if needed, propose modified versions of the tool that correct for miscalibration.[32 33] The assessment of predicted and observed risk was not performed in several subgroups (eg, by ethnicity, socioeconomic status) as well as in prevalent COPD cases and those with previous CVD events. These can be performed as part of a future dedicated validation study. Similar to the QRISK development algorithm, we used multiple imputation to include individuals with missing predictor values, but the pattern of missingness might be different between individuals with COPD and the general population. Several predictors (eg, lab tests outside regular check-ups) are often ascertained based on clinical suspicion, and as such, the prevalence of abnormalities (associated with higher CVD risk) in the imputed predictor values might be higher than the true, unobserved values. However, any bias due to such overestimation is in the opposite direction of our results (as it will result in overestimating the QRISK score), and thus is unlikely to threaten the validity of our findings. We did not have a comparison group from the general CPRD population. As such, to draw comparisons, we relied on the data provided in the QRISK3 original study for the general population.[11] While the QRISK development population and our study population are similar (both from electronic medical records of primary care UK practices), we cannot rule out differences due to different sampling methods in the two populations. However, this has little implications for the most important finding of our results: QRISK3 underestimates CVD risk in this representative primary care COPD sample.

In conclusion, increased CVD risk in COPD patients cannot simply be explained by shared risk factors such as age and smoking. CVD risk scoring tools developed for the general population capture a fraction of the added risk. This is particularly the case for younger patients whose actual CVD risk can be >80% higher than predicted values. Such discrepancy can result in missed opportunities for the prevention of CVD through risk factor modification, behavioural change and therapeutics. Current CVD risk scoring tools need to be validated and recalibrated to provide clinical utility for patients with COPD, and COPD should be considered as a distinct predictor in future CVD risk scoring tools.

Author affiliations
[1]Respiratory Evaluation Sciences Program, Faculty of Pharmaceutical Sciences, The University of British Columbia, Vancouver, British Columbia, Canada
[2]Division of Community Health and Humanities, Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada
[3]School of Public Health, National Heart and Lung Institute, Imperial College London, London, UK
[4]Respiratory Medicine Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK
[5]King's Centre of Lung Health, Peter Gorer Dept of Immunobiology, Faculty of Life Sciences and Medicine, King's College London, London, UK
[6]Academic Unit of Respiratory Medicine, University College London Medical School, London, UK
[7]Centre for Heart Lung Innovation, Faculty of Medicine, The University of British Columbia, Vancouver, British Columbia, Canada
[8]Saw Swee Hock School of Public Health, National University of Singapore, Singapore

Twitter Mohsen Sadatsafavi @mo_safavi

Contributors MS and MB conceived the study question. JEA, MS and TYL developed the analytic plan, performed the literature review and conducted the analyses. ZG facilitated data access. MB and JKQ revised the initial study protocol. RR, JRH and DDS revised and finalised the protocol. MS, JKQ, TL, WC and MB supervised the study progress and provided regular feedback. JEA and MS contributed to the study design and TL performed part of the statistical analyses. JA wrote the first draft of the manuscript. All authors critically revised the manuscript and approved the final copy. MS is the guarantor of this work.

Funding This study was funded by The Canadian Institutes of Health Research (Team Grant PHT 178432). JEA was supported by Legacy for Airway Health and a scholarship from MITACS Canada.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Ethical approval was granted by the Health Research Ethics Board at Memorial University of Newfoundland (HREB #2017.024), and the Independent Scientific Advisory Committee overseeing CPRD approved our study protocol (Protocol 18_005RA3PAA).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. This study is partly based on data from the Clinical Practice Research Datalink (CPRD-GOLD) obtained under licence from the UK Medicines and

Healthcare Products Regulatory Agency. However, the interpretation and conclusions contained in this study are those of the authors alone. Data are available upon request from the data custodian, the CPRD.

ORCID iDs
Joseph Emil Amegadzie http://orcid.org/0000-0002-0459-7241
Jennifer K Quint http://orcid.org/0000-0003-0149-4869
John R Hurst http://orcid.org/0000-0002-7246-6040
Don D Sin http://orcid.org/0000-0002-0756-6643
Wenjia Chen http://orcid.org/0000-0001-8201-7145
Mona Bafadhel http://orcid.org/0000-0002-9993-2478
Mohsen Sadatsafavi http://orcid.org/0000-0002-0419-7862

REFERENCES

1 López-Campos JL, Tan W, Soriano JB. Global burden of COPD. *Respirology* 2016;21:14–23.
2 Sheikh K, Coxson HO, Parraga G. This is what COPD looks like. *Respirology* 2016;21:224–36.
3 Rasmussen T, Køber L, Pedersen JH, *et al*. Relationship between chronic obstructive pulmonary disease and subclinical coronary artery disease in long-term smokers. *Eur Heart J Cardiovasc Imaging* 2013;14:1159–66.
4 Cazzola M, Calzetta L, Bettoncelli G, *et al*. Cardiovascular disease in asthma and COPD: a population-based retrospective cross-sectional study. *Respiratory Medicine* 2012;106:249–56.
5 Chen W, Thomas J, Sadatsafavi M, *et al*. Risk of cardiovascular comorbidity in patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Lancet Respir Med* 2015;3:631–9.
6 Balasubramanian VP, Varkey B. Chronic obstructive pulmonary disease: effects beyond the lungs. *Curr Opin Pulm Med* 2006;12:106–12.
7 Berry DA. A case for Bayesianism in clinical trials. *Stat Med* 1993;12:1377–93;
8 Arnett DK, Blumenthal RS, Albert MA, *et al*. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: executive summary: a report of the American college of cardiology/American Heart Association task force on clinical practice guidelines. *J Am Coll Cardiol* 2019;74:1376–414.
9 Chen Y-WR, Leung JM, Sin DD, *et al*. A systematic review of diagnostic biomarkers of COPD exacerbation. *PLoS ONE* 2016;11:e0158843.
10 Lloyd-Jones DM, Morris PB, Ballantyne CM, *et al*. 2017 focused update of the 2016 ACC expert consensus decision pathway on the role of non-statin therapies for LDL-cholesterol lowering in the management of atherosclerotic cardiovascular disease risk: a report of the American college of cardiology task force on expert consensus decision pathways. *J Am Coll Cardiol* 2017;70:1785–822.
11 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.
12 Samarasekera EJ, Clark CE, Kaur S, *et al*. Cardiovascular disease risk assessment and reduction: summary of updated NICE guidance. *BMJ* 2023;381:1028.
13 Herrett E, Gallagher AM, Bhaskaran K, *et al*. Data resource profile: clinical practice research datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.
14 Hippisley-Cox J, Coupland C, Brindle P. The performance of seven Qprediction risk scores in an independent external sample of patients from general practice: a validation study. *BMJ Open* 2014;4:e005809.
15 Quint JK, Müllerova H, DiSantostefano RL, *et al*. Validation of chronic obstructive pulmonary disease recording in the clinical practice research datalink (CPRD-GOLD). *BMJ Open* 2014;4:e005540.
16 Bhatt SP, Wells JM, Kinney GL, *et al*. Clinician estimation of cardiovascular risk in patients with advanced COPD: the COPD and cardiovascular disease education study (CCED). *Chest* 2014;146:348–58.
17 Vaz Fragoso CA, Gill TM, McAvay G, *et al*. Respiratory impairment and mortality in older persons: a novel spirometric approach. *J Investig Med* 2011;59:1089–95.
18 Fragoso CAV. Epidemiology of chronic obstructive pulmonary disease (COPD) in aging populations. *COPD: Journal of Chronic Obstructive Pulmonary Disease* 2016;13:125–9.
19 Marshall A, Altman DG, Holder RL, *et al*. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* 2009;9:57.
20 The R Foundation. The R project for statistical computing, version 4.2.2. n.d. Available: https://www.r-project.org/
21 Li Y. Package 'Qrisk3'. n.d. Available: https://cran.r-project.org/web/packages/QRISK3/QRISK3.pdf
22 Higbee DH, Granell R, Sanderson E, *et al*. Lung function and cardiovascular disease: a two-sample Mendelian randomisation study. *Eur Respir J* 2021;58:2003196.
23 Maclagan LC, Croxford R, Chu A, *et al*. Quantifying COPD as a risk factor for cardiac disease in a primary prevention cohort. *Eur Respir J* 2023;62:2202364.
24 Curkendall SM, DeLuise C, Jones JK, *et al*. Cardiovascular disease in patients with chronic obstructive pulmonary disease, Saskatchewan Canada cardiovascular disease in COPD patients. *Ann Epidemiol* 2006;16:63–70.
25 Finkelstein J, Cha E, Scharf SM. Chronic obstructive pulmonary disease as an independent risk factor for cardiovascular morbidity. *Int J Chron Obstruct Pulmon Dis* 2009;4:337–49.
26 Vivodtzev I, Maltais F. Cardiovascular risk in COPD: searching for a culprit. *Chest* 2020;157:753–4.
27 Naik G, Ahmed H, Edwards AGK. Communicating risk to patients and the public. *Br J Gen Pract* 2012;62:213–6.
28 DAWBER TR, MEADORS GF, MOORE FE Jr. Epidemiological approaches to heart disease: the Framingham study. *Am J Public Health Nations Health* 1951;41:279–81.
29 Sampson M, Wolska A, Amar M, *et al*. Estimated atherosclerotic cardiovascular disease risk score: an automated decision aid for statin therapy. *Clinical Chemistry* 2022;68:1302–10.
30 Bhatt SP, Kazerooni EA, Newell JD, *et al*. Visual estimate of coronary artery calcium predicts cardiovascular disease in COPD. *Chest* 2018;154:579–87.
31 Rothnie KJ, Smeeth L, Pearce N, *et al*. Predicting mortality after acute coronary syndromes in people with chronic obstructive pulmonary disease. *Heart* 2016;102:1442–8.
32 Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014;35:1925–31.
33 Sadatsafavi M, Adibi A, Puhan M, *et al*. Moving beyond AUC: decision curve analysis for quantifying net benefit of risk prediction models. *Eur Respir J* 2021;58:2101186.

QRISK3 underestimates the risk of cardiovascular events in patients with COPD

Supplementary Table 1: READ codes to identify COPD patients in the CPRD

Read Code	Description/Read Term
8H2R.00	Admit COPD emergency
9NgP.11	On COPD (chr obstruc pulmonary disease) supportv cre pathway
66Yi.00	Multiple COPD emergency hospital admissions
66YL.11	COPD follow-up
661M300	COPD self-management plan agreed
66Ye.00	Emergency COPD admission since last appointment
9kf0.11	COPD patient unsuitable for pulmonary rehabilitation
66Yd.00	COPD accident and emergency attendance since last visit
9Oi1.00	Chronic obstructive pulmonary disease monitoring 2nd letter
9Oi2.00	Chronic obstructive pulmonary disease monitoring 3rd letter
H36..00	Mild chronic obstructive pulmonary disease
66Yg.00	Chronic obstructive pulmonary disease disturbs sleep
66YD.00	Chronic obstructive pulmonary disease monitoring due
H37..00	Moderate chronic obstructive pulmonary disease
H38..00	Severe chronic obstructive pulmonary disease
9Oi3.00	Chronic obstructive pulmonary disease monitoring verb invite
66YT.00	Chronic obstructive pulmonary disease monitoring by doctor
9Oi4.00	Chronic obstructive pulmonary disease monitor phone invite
9NgP.00	On chronic obstructive pulmonary disease supprtv cre pathway
H3y..11	Other specified chronic obstructive pulmonary disease
8CMV.00	Has chronic obstructive pulmonary disease care plan
H320.00	Chronic bullous emphysema
H32z.00	Emphysema NOS
H32.00	Emphysema
1J71.00	Suspected chronic obstructive pulmonary disease
H312100	Emphysematous bronchitis
H322.00	Centrilobular Emphysema
679V.00	Health education - chronic obstructive pulmonary disease
H39..00	Very severe chronic obstructive pulmonary disease
66YS.00	Chronic obstructive pulmonary disease monitoring by nurse
H320z00	Chronic bullous emphysema NOS
H312z00	obstructive Chronic bronchitis NOS
66YL.00	Chronic obstructive pulmonary disease follow-up
66Yh.00	Chronic obstructive pulmonary disease does not disturb sleep
9Oi0.00	Chronic obstructive pulmonary disease monitoring 1st letter
8CMW500	Chronic obstructive pulmonary disease care pathway
8BMW.00	Issue of chronic obstructive pulmonary disease rescue pack
66YB200	Telehealth chronic obstructive pulmonary disease monitoring
66YB.00	Chronic obstructive pulmonary disease monitoring
H3...00	Chronic obstructive pulmonary disease
66YM.00	Chronic obstructive pulmonary disease annual review
Hyu3100	[X]Other specified chronic obstructive pulmonary disease
66YB100	Chronic obstructive pulmonary disease 6 monthly review

Amegadzie JE, *et al. Thorax* 2023;0:1–7. doi: 10.1136/thorax-2023-220615

66YB100	Chronic obstructive pulmonary disease 3 monthly review
H3z..11	Chronic obstructive airways disease NOS
H3...11	Chronic obstructive airways disease
H3y..00	Other specified chronic obstructive airways disease
H3A..00	End stage chronic obstructive airways disease
663K.00	Airways obstructn irreversible
H312200	Acute exacerbation of chronic obstructive airways disease
H312200	Chronic obstructive airways disease

Amegadzie JE, *et al. Thorax* 2023;0:1–7. doi: 10.1136/thorax-2023-220615

Supplementary Table 2: Read codes used to identify patients with cardiovascular disease from GP records.[1]

Group Name	Read Term	Read Description
Coronary Heart Disease	G3	Ischaemic heart disease
Coronary Heart Disease	G3-1	Arteriosclerotic heart disease
Coronary Heart Disease	G3-2	Atherosclerotic heart disease
Coronary Heart Disease	G3-3	IHD - Ischaemic heart disease
Coronary Heart Disease	G30	Acute myocardial infarction
Coronary Heart Disease	G30-1	Attack - heart
Coronary Heart Disease	G30-2	Coronary thrombosis
Coronary Heart Disease	G30-3	Cardiac rupture following myocardial infarction (MI)
Coronary Heart Disease	G30-4	Heart attack
Coronary Heart Disease	G30-5	MI - acute myocardial infarction
Coronary Heart Disease	G30-6	Thrombosis - coronary
Coronary Heart Disease	G30-7	Silent myocardial infarction
Coronary Heart Disease	G30-98	Coronary thrombosis
Coronary Heart Disease	G30-99	Myocardial Infarction
Coronary Heart Disease	G300	Acute anterolateral infarction
Coronary Heart Disease	G301	Other specified anterior myocardial infarction
Coronary Heart Disease	G3010	Acute anteroapical infarction
Coronary Heart Disease	G3011	Acute anteroseptal infarction
Coronary Heart Disease	G301z	Anterior myocardial infarction NOS
Coronary Heart Disease	G302	Acute inferolateral infarction
Coronary Heart Disease	G303	Acute inferoposterior infarction
Coronary Heart Disease	G304	Posterior myocardial infarction NOS
Coronary Heart Disease	G305	Lateral myocardial infarction NOS
Coronary Heart Disease	G306	True posterior myocardial infarction
Coronary Heart Disease	G307	Acute subendocardial infarction
Coronary Heart Disease	G3070	Acute non-Q wave infarction
Coronary Heart Disease	G3071	Acute non-ST segment elevation myocardial infarction
Coronary Heart Disease	G308	Inferior myocardial infarction NOS
Coronary Heart Disease	G309	Acute Q-wave infarct
Coronary Heart Disease	G30A	Mural thrombosis
Coronary Heart Disease	G30B	Acute posterolateral myocardial infarction
Coronary Heart Disease	G30X	Acute transmural myocardial infarction of unspecif site
Coronary Heart Disease	G30X0	Acute ST segment elevation myocardial infarction
Coronary Heart Disease	G30y	Other acute myocardial infarction
Coronary Heart Disease	G30y0	Acute atrial infarction
Coronary Heart Disease	G30y1	Acute papillary muscle infarction
Coronary Heart Disease	G30y2	Acute septal infarction
Coronary Heart Disease	G30yz	Other acute myocardial infarction NOS
Coronary Heart Disease	G30z	Acute myocardial infarction NOS
Coronary Heart Disease	G31	Other acute and subacute ischaemic heart disease
Coronary Heart Disease	G31-99	Acute/subacute IHD NOS
Coronary Heart Disease	G310	Postmyocardial infarction syndrome
Coronary Heart Disease	G310-1	Dressler's syndrome
Coronary Heart Disease	G311	Preinfarction syndrome
Coronary Heart Disease	G311-1	Crescendo angina

Coronary Heart Disease	G311-2	Impending infarction
Coronary Heart Disease	G311-3	Unstable angina
Coronary Heart Disease	G311-4	Angina at rest
Coronary Heart Disease	G3110	Myocardial infarction aborted
Coronary Heart Disease	G3110-1	MI - myocardial infarction aborted
Coronary Heart Disease	G3111	Unstable angina
Coronary Heart Disease	G3112	Angina at rest
Coronary Heart Disease	G3113	Refractory angina
Coronary Heart Disease	G3114	Worsening angina
Coronary Heart Disease	G3115	Acute coronary syndrome
Coronary Heart Disease	G311z	Preinfarction syndrome NOS
Coronary Heart Disease	G312	Coronary thrombosis not resulting in myocardial infarction
Coronary Heart Disease	G31y	Other acute and subacute ischaemic heart disease
Coronary Heart Disease	G31y0	Acute coronary insufficiency
Coronary Heart Disease	G31y0-99	Acute coronary syndrome
Coronary Heart Disease	G31y1	Microinfarction of heart
Coronary Heart Disease	G31y2	Subendocardial ischaemia
Coronary Heart Disease	G31y3	Transient myocardial ischaemia
Coronary Heart Disease	G31yz	Other acute and subacute ischaemic heart disease NOS
Coronary Heart Disease	G32	Old myocardial infarction
Coronary Heart Disease	G32-1	Healed myocardial infarction
Coronary Heart Disease	G32-2	Personal history of myocardial infarction
Coronary Heart Disease	G33	Angina pectoris
Coronary Heart Disease	G330	Angina decubitus
Coronary Heart Disease	G3300	Nocturnal angina
Coronary Heart Disease	G330z	Angina decubitus NOS
Coronary Heart Disease	G331	Prinzmetal's angina
Coronary Heart Disease	G331-1	Variant angina pectoris
Coronary Heart Disease	G332	Coronary artery spasm
Coronary Heart Disease	G33z	Angina pectoris NOS
Coronary Heart Disease	G33z0	Status anginosus
Coronary Heart Disease	G33z1	Stenocardia
Coronary Heart Disease	G33z2	Syncope anginosa
Coronary Heart Disease	G33z3	Angina on effort
Coronary Heart Disease	G33z4	Ischaemic chest pain
Coronary Heart Disease	G33z5	Post infarct angina
Coronary Heart Disease	G33z6	New onset angina
Coronary Heart Disease	G33z7	Stable angina
Coronary Heart Disease	G33zz	Angina pectoris NOS
Coronary Heart Disease	G34	Other chronic ischaemic heart disease
Coronary Heart Disease	G34-99	Chr. ischaemic heart dis. NOS
Coronary Heart Disease	G340	Coronary atherosclerosis
Coronary Heart Disease	G340-1	Triple vessel disease of the heart
Coronary Heart Disease	G340-2	Coronary artery disease
Coronary Heart Disease	G3400	Single coronary vessel disease
Coronary Heart Disease	G3401	Double coronary vessel disease
Coronary Heart Disease	G342	Atherosclerotic cardiovascular disease
Coronary Heart Disease	G343	Ischaemic cardiomyopathy
Coronary Heart Disease	G344	Silent myocardial ischaemia
Coronary Heart Disease	G34y	Other specified chronic ischaemic heart disease

Coronary Heart Disease	G34y0	Chronic coronary insufficiency
Coronary Heart Disease	G34y1	Chronic myocardial ischaemia
Coronary Heart Disease	G34yz	Other specified chronic ischaemic heart disease NOS
Coronary Heart Disease	G34z	Other chronic ischaemic heart disease NOS
Coronary Heart Disease	G34z0	Asymptomatic coronary heart disease
Coronary Heart Disease	G35	Subsequent myocardial infarction
Coronary Heart Disease	G350	Subsequent myocardial infarction of anterior wall
Coronary Heart Disease	G351	Subsequent myocardial infarction of inferior wall
Coronary Heart Disease	G353	Subsequent myocardial infarction of other sites
Coronary Heart Disease	G35X	Subsequent myocardial infarction of unspecified site
Coronary Heart Disease	G36	Certain current complication follow acute myocardial infarct
Coronary Heart Disease	G360	Haemopericardium/current comp folow acut myocard infarct
Coronary Heart Disease	G361	Atrial septal defect/curr comp folow acut myocardal infarct
Coronary Heart Disease	G362	Ventric septal defect/curr comp fol acut myocardal infarctn
Coronary Heart Disease	G363	Ruptur cardiac wall w'out haemopericard/cur comp fol ac MI
Coronary Heart Disease	G364	Ruptur chordae tendinae/curr comp fol acute myocard infarct
Coronary Heart Disease	G365	Rupture papillary muscle/curr comp fol acute myocard infarct
Coronary Heart Disease	G366	Thrombosis atrium,auric append&vent/curr comp foll acute MI
Coronary Heart Disease	G38	Postoperative myocardial infarction
Coronary Heart Disease	G380	Postoperative transmural myocardial infarction anterior wall
Coronary Heart Disease	G381	Postoperative transmural myocardial infarction inferior wall
Coronary Heart Disease	G382	Postoperative transmural myocardial infarction other sites
Coronary Heart Disease	G383	Postoperative transmural myocardial infarction unspec site
Coronary Heart Disease	G384	Postoperative subendocardial myocardial infarction
Coronary Heart Disease	G38z	Postoperative myocardial infarction, unspecified
Coronary Heart Disease	G3y	Other specified ischaemic heart disease
Coronary Heart Disease	G3z	Ischaemic heart disease NOS
Coronary Heart Disease	G501	Post infarction pericarditis
Coronary Heart Disease	Gyu34	[X]Acute transmural myocardial infarction of unspecif site
Stroke or TIA	F4236	Amaurosis fugax
Stroke or TIA	Fyu55	[X]Other transnt cerebral ischaemic attacks+related syndroms
Stroke or TIA	G63y0	Cerebral infarct due to thrombosis of precerebral arteries
Stroke or TIA	G63y1	Cerebral infarction due to embolism of precerebral arteries
Stroke or TIA	G64	Cerebral arterial occlusion
Stroke or TIA	G64-1	CVA - cerebral artery occlusion
Stroke or TIA	G64-2	Infarction - cerebral
Stroke or TIA	G64-3	Stroke due to cerebral arterial occlusion
Stroke or TIA	G640	Cerebral thrombosis
Stroke or TIA	G6400	Cerebral infarction due to thrombosis of cerebral arteries
Stroke or TIA	G641	Cerebral embolism
Stroke or TIA	G641-1	Cerebral embolus
Stroke or TIA	G6410	Cerebral infarction due to embolism of cerebral arteries
Stroke or TIA	G64z	Cerebral infarction NOS
Stroke or TIA	G64z-1	Brainstem infarction NOS
Stroke or TIA	G64z-2	Cerebellar infarction
Stroke or TIA	G64z-99	Cerebral A. occlusion NOS
Stroke or TIA	G64z0	Brainstem infarction
Stroke or TIA	G64z1	Wallenberg syndrome
Stroke or TIA	G64z1-1	Lateral medullary syndrome
Stroke or TIA	G64z2	Left sided cerebral infarction

Stroke or TIA	G64z3	Right sided cerebral infarction
Stroke or TIA	G64z4	Infarction of basal ganglia
Stroke or TIA	G65	Transient cerebral ischaemia
Stroke or TIA	G65-1	Drop attack
Stroke or TIA	G65-2	Transient ischaemic attack
Stroke or TIA	G65-3	Vertebro-basilar insufficiency
Stroke or TIA	G65-99	Transient Ischaemic Attacks
Stroke or TIA	G650	Basilar artery syndrome
Stroke or TIA	G650-1	Insufficiency - basilar artery
Stroke or TIA	G652	Subclavian steal syndrome
Stroke or TIA	G653	Carotid artery syndrome hemispheric
Stroke or TIA	G654	Multiple and bilateral precerebral artery syndromes
Stroke or TIA	G656	Vertebrobasilar insufficiency
Stroke or TIA	G65y	Other transient cerebral ischaemia
Stroke or TIA	G65z	Transient cerebral ischaemia NOS
Stroke or TIA	G65z-99	Transient Ischaemic Attacks
Stroke or TIA	G65z0	Impending cerebral ischaemia
Stroke or TIA	G65z1	Intermittent cerebral ischaemia
Stroke or TIA	G65zz	Transient cerebral ischaemia NOS
Stroke or TIA	G66	Stroke and cerebrovascular accident unspecified
Stroke or TIA	G66-1	CVA unspecified
Stroke or TIA	G66-2	Stroke unspecified
Stroke or TIA	G66-3	CVA - Cerebrovascular accident unspecified
Stroke or TIA	G66-98	Stroke/CVA - undefined
Stroke or TIA	G66-99	Stroke
Stroke or TIA	G667	Left sided CVA
Stroke or TIA	G668	Right sided CVA
Stroke or TIA	G6760	Cereb infarct due cerebral venous thrombosis, nonpyogenic
Stroke or TIA	G6W	Cereb infarct due unsp occlus/stenos precerebr arteries
Stroke or TIA	G6X	Cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs
Stroke or TIA	Gyu63	[X]Cerebrl infarctn due/unspcf occlusn or sten/cerebrl artrs
Stroke or TIA	Gyu64	[X]Other cerebral infarction
Stroke or TIA	Gyu65	[X]Occlusion and stenosis of other precerebral arteries
Stroke or TIA	Gyu66	[X]Occlusion and stenosis of other cerebral arteries
Stroke or TIA	ZV12D	[V]Personal history of transient ischaemic attack

Supplementary Table 3: ICD-10 codes used for hospitalisation outcomes, adopted from the QRISK tool.[1]

ICD-10	Description
G45	Transient ischaemic attack and related syndromes
I20	Angina pectoris
I21	Acute myocardial infarction
I22	Subsequent myocardial infarction
I23	Complications after myocardial infarction
I24	Other acute ischaemic heart disease
I25	Chronic ischaemic heart disease
I63	Cerebral infarction
I64	Stroke not specified as haemorrhage or infarction

Supplementary Table 4: Definition and Evaluation of Predictors

Predictors	**Definition of Variables**	**Time-window consideration**	**Missing data** (percentages)
Age	Difference between the birth year in the general practice record and the date of cohort entry	Determined at the date of diagnosis of COPD	None
Ethnicity	Available and recorded general practice record as White, Black_Caribbean, Black_African, Black_Other, Indian, Pakistani, Bangladeshi, Other_Asian, Chinese, Mixed, Other (where there is no majority)	Latest information recorded before entry to the cohort	Assumed white if missing
Cholesterol	Has two categories: Total and HDL cholesterol ratio. This predictor will be assessed as the ratio of serum versus HDL	Closest value to cohort entry and restricting values after the baseline date to those before the CVD event or before any statin prescription	Total cholesterol = 45.9 HDL cholesterol = 58.3 Total and HDL cholesterol ratio = 58.4
Smoking status	This predictor will be categorized into three categories "current-, ex- and no-smoker." We will also look at the number of packs smoked each year. However, we anticipate significant missingness when assessing pack-years smoked	Latest information recorded before entry to the cohort	17.0
Gender	Male or female as defined by general practice records	Latest information recorded before entry to the cohort	None
Family history of CHD	Family history(diagnosis) of cardiovascular disease in first-degree relatives less than 60 years of age	Recorded at the baseline visit	Assumed not present if not recorded
Diabetes	Diagnosis of diabetes recorded in the general practice data	Recorded at the baseline visit	Assumed not present if not recorded
Rheumatoid Arthritis	Diagnosis of rheumatoid arthritis, Felty's syndrome, Caplan's syndrome, adult-onset Still's disease or inflammatory polyarthropathy in the general practice data	Recorded at the baseline visit	Assumed not present if not recorded
Atrial Fibrillation	Diagnosis of Atrial fibrillation (including atrial fibrillation, atrial flutter, and paroxysmal atrial fibrillation) in the general practice data	Recorded at the baseline visit	Assumed not present if not recorded
Chronic Kidney Disease	Chronic kidney disease (stage 4 or 5) and major chronic renal disease (including nephrotic syndrome, chronic glomerulonephritis, chronic pyelonephritis, renal dialysis, and renal transplant) in the general practice data	Recorded at the baseline visit	Assumed not present if not recorded
Blood Pressure	Only systolic blood pressure will be considered per the QRISK study. This will be defined as any systolic blood pressure recorded within five years of cohort entry. For two or more records, the standard deviation will be computed	Recorded in the five years before study entry	Systolic blood pressure = 15.8 ≥2 systolic blood pressure = 38.6

Migraine	Diagnosis of migraine (including classic migraine, atypical migraine, abdominal migraine, cluster headaches, basilar migraine, hemiplegic migraine, and migraine with or without aura) in the general practice data	Recorded at the baseline visit	Assumed not present if not recorded
Socio-Economic Status (Index Of Multiple Deprivations)	Characterized into five categories as least deprived, less deprived, deprived, more deprived or most deprived	Recorded before the baseline visit (CPRD: 2015 updated recorded)	None
Body Mass Index	The ratio of weight in kilograms versus height in meters squared	Most recent recorded at the baseline visit	4.2
Treated Hypertension	Considered as a diagnosis of hypertension and treatment with at least one antihypertensive drug	Recorded at the baseline visit	Assumed not present if not recorded
Corticosteroid Use	At least two prescriptions of oral or parenteral prednisolone, betamethasone, cortisone, depo-medrone, dexamethasone, deflazacort, efcortesol, hydrocortisone, methylprednisolone, or triamcinolone with the most recent one being not more than 28 days prior to the cohort enrollment date	At least two prescriptions with the most recent one not more than 28 days before the date of cohort entry	Assumed not present if not recorded
Systemic Lupus Erythematosus (SLE)	Diagnosis of SLE (including disseminated lupus erythematosus or Libman-Sacks disease)	Recorded at the baseline visit	Assumed not present if not recorded
Second-Generation “Atypical” Antipsychotic	The use of amisulpride, aripiprazole, clozapine, lurasidone, olanzapine, paliperidone, quetiapine, risperidone, sertindole, or zotepine with the most recent one being not more than 28 days prior to the cohort enrollment date	At least two prescriptions with the most recent one not more than 28 days before the date of cohort entry	Assumed not present if not recorded
Severe Mental Illness	The diagnosis of severe mental illness, including psychosis, schizophrenia, or bipolar affective disease in the general practice data	Recorded at the baseline visit	Assumed not present if not recorded
HIV or AIDS	The diagnosis of HIV/AIDS in the general practice data	Recorded at the baseline visit	Assumed not present if not recorded
Erectile Dysfunction	The diagnosis of erectile dysfunction or treatment for erectile dysfunction, including alprostadil, phosphodiesterase type 5 inhibitors, papaverine, or phentolamine	Recorded at the baseline visit	Assumed not present if not recorded

Supplementary Material – Table 5: Incidence by age group and total

Age group	Gender	No. of patients	Total FU time	No. of CVD	Incidence CVD X 10000	Incidence CVD X 10000 (Hippisley et al)
40-44	M	195	954.1	15	157.2183	39.9
40-44	F	165	1033.4	8	77.41826	19
45-50	M	369	1998.8	65	325.2	66.5
45-50	F	324	1864.0	28	150.2183	32
50-55	M	658	3983.9	102	256.028	98.6
50-55	F	516	3139.6	55	175.1831	48.3
55-60	M	1048	6297.5	192	304.8812	141.8
55-60	F	812	5063.4	89	175.7701	74.7
60-65	M	1361	7819.8	247	315.8632	196.9
60-65	F	988	5747.3	130	226.193	113.6
65-70	M	1216	6829.0	274	401.2321	265.5
65-70	F	967	5612.6	163	290.4203	171.3
70-75	M	1061	5436.4	277	509.533	354.8
70-75	F	879	4911.8	205	417.3641	250.8
75-80	M	839	3812.7	244	639.9671	451.6
75-80	F	740	3690.3	183	495.8978	351.1
80-85	M	577	2209.9	160	724.0178	582.9
80-85	F	493	2215.4	135	609.3821	480.2
TOTAL	M	7324	39342.1	1576	400.6	137.5
TOTAL	F	5884	33277.6	996	299.3	101.9

Reference

1. 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(Journal Article):j2099-j2099. doi:10.1136/bmj.j2099