

How Does Cardiovascular Disease First Present in Women and Men?

Incidence of 12 Cardiovascular Diseases in a Contemporary Cohort of 1 937 360 People

Julie George, PhD, FFPH*; Eleni Rapsomaniki, PhD*; Mar Pujades-Rodriguez, PhD;
Anoop Dinesh Shah, MSc, MRCP; Spiros Denaxas, PhD; Emily Herrett, PhD;
Liam Smeeth, PhD, FRGP; Adam Timmis, MD, FRCP; Harry Hemingway, FFPH, FRCP

Background—Given the recent declines in heart attack and stroke incidence, it is unclear how women and men differ in first lifetime presentations of cardiovascular diseases (CVDs). We compared the incidence of 12 cardiac, cerebrovascular, and peripheral vascular diseases in women and men at different ages.

Methods and Results—We studied 1 937 360 people, aged ≥ 30 years and free from diagnosed CVD at baseline (51% women), using linked electronic health records covering primary care, hospital admissions, acute coronary syndrome registry, and mortality (Cardiovascular Research Using Linked Bespoke Studies and Electronic Records [CALIBER] research platform). During 6 years median follow-up between 1997 and 2010, 114 859 people experienced an incident cardiovascular diagnosis, the majority (66%) of which were neither myocardial infarction nor ischemic stroke. Associations of male sex with initial diagnoses of CVD, however, varied from strong (age-adjusted hazard ratios, 3.6–5.0) for abdominal aortic aneurysm, myocardial infarction, and unheralded coronary death (particularly >60 years), through modest (hazard ratio, 1.5–2.0) for stable angina, ischemic stroke, peripheral arterial disease, heart failure, and cardiac arrest, to weak (hazard ratio <1.5) for transient ischemic attack, intracerebral hemorrhage, and unstable angina, and inverse (0.69) for subarachnoid hemorrhage (all $P < 0.001$).

Conclusions—The majority of initial presentations of CVD are neither myocardial infarction nor ischemic stroke, yet most primary prevention studies focus on these presentations. Sex has differing associations with different CVDs, with implications for risk prediction and management strategies.

Clinical Trial Registration—URL: <http://www.clinicaltrials.gov>. Unique identifier: NCT01164371.

(*Circulation*. 2015;132:1320-1328. DOI: 10.1161/CIRCULATIONAHA.114.013797.)

Key Words: aging ■ cardiovascular diseases ■ electronic health records ■ incidence
■ population ■ risk factors ■ sex

A central principle in cardiovascular disease (CVD) management is that the first lifetime diagnosis signals the failure of primary prevention and the need to initiate secondary prevention of recurrent or related CVD events. The decades-long emphasis given to prevention of myocardial infarction (MI) and stroke is reflected in remarkable declines – $\approx 33\%$ over the past decade – in their incidence in developed countries.¹ Incidence rates for chronic CVD presentations such as angina or heart failure, although less studied, do not appear to have similarly declined.^{1–3} Consequently, the spectrum of

initial presentations of CVD in contemporary practice is likely to have changed in comparison with the latter part of the last century. Cohort studies that report only fatal end points (final presentations),⁴ may have less relevance to informing the success of primary prevention than those which investigate initial presentations. Within studies that incorporate nonfatal events, acute MI and stroke have been more commonly investigated

Editorial see p 1303
Clinical Perspective on p 1328

Continuing medical education (CME) credit is available for this article. Go to <http://cme.ahajournals.org> to take the quiz

Received October 17, 2014; accepted July 13, 2015.

From Farr Institute of Health Informatics Research (London), University College London, United Kingdom (J.G., M.P.-R., A.D.S., S.D., H.H.); Worldwide Clinical Trials, Nottingham, United Kingdom (E.R.); Farr Institute of Health Informatics Research (London), London School of Hygiene & Tropical Medicine, United Kingdom (E.H., L.S.); and Farr Institute of Health Informatics Research (London) and Barts National Institute for Health Research Cardiovascular Biomedical Research Unit, Queen Mary University of London, United Kingdom (A.T.).

*Drs George and Rapsomaniki are joint first authors.

The online-only Data Supplement is available with this article at <http://circ.ahajournals.org/lookup/suppl/doi:10.1161/CIRCULATIONAHA.114.013797/-/DC1>.

Correspondence to Julie George, PhD, FFPH, Farr Institute of Health Informatics Research (London), UCL, 222 Euston Rd, London, United Kingdom, NW1 2DA. E-mail j.george@ucl.ac.uk

© 2015 The Authors. *Circulation* is published on behalf of the American Heart Association, Inc., by Wolters Kluwer. This is an open access article under the terms of the [Creative Commons Attribution](http://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution, and reproduction in any medium, provided that the original work is properly cited.

Circulation is available at <http://circ.ahajournals.org>

DOI: 10.1161/CIRCULATIONAHA.114.013797

than other chronic presentations.⁵⁻⁷ Large-scale contemporary studies that evaluate the first lifetime diagnosis in women and men across a wide range of acute and chronic CVDs including both fatal and nonfatal presentations can provide additional insight into the understanding of CVDs.

Fundamental unanswered questions about initial CVD presentation arise. First, what is the relative frequency of different CVDs as they affect women and men in contemporary practice? Second, is male sex an equally strong risk factor common to all CVDs, or does the association differ across a range of diseases?

The lack of large, contemporary, population-based cohorts with detailed clinical follow-up spanning hospital and ambulatory care has hindered the study of the initial presentation of a wide range of acute and chronic CVDs. It has been suggested that electronic health record (EHR) data might be meaningfully reused⁸ to create mega-cohorts for such research.⁹ We studied a contemporary, population-based cohort based on linked EHRs across primary, secondary, disease registry, and death records¹⁰⁻¹³ to address these 2 questions. We investigated a wide range of acquired symptomatic CVDs that are recognized to have differing pathogenic mechanisms.

Methods

Data Sources

Anonymized patients were selected from the Cardiovascular Research Using Llinked Bespoke Studies and Electronic Records (CALIBER) program, described¹⁴ and validated^{10-13,15} elsewhere. Patients were linked across 4 clinical data sources: the Clinical Practice Research Database (CPRD), the Myocardial Ischemia National Audit Project registry, Hospital Episodes Statistics, and the national death registry from the Office for National Statistics. CPRD provides primary care data on anthropometric measurements, laboratory tests, medical history, clinical diagnoses, prescriptions, medical procedures, and health behaviors, coded using the Read clinical coding scheme. Patients registered in practices submitting linkable data to CPRD, covering ≈4% of the English population, have been found to be representative of the English population in terms of age, sex, and ethnicity.^{16,17} Myocardial Ischemia National Audit Project is a national registry of patients admitted to the hospital with acute coronary syndromes. Hospital Episodes Statistics provides information on diagnoses and medical procedures related to all elective and emergency hospital admissions across all National Health Service hospitals in England.

Study Population

We studied 1937360 patients from 225 general practices across England registered between January 1997 and March 2010. We required that at study entry patients were aged ≥30 years, were free of diagnosed CVD, and had been followed up for at least 1 year. We used the entire medical history available on each patient to confirm they were free of diagnosed CVD. The look-back period ranged from 20 years to the minimum of 1 year, which previous research has indicated is a sufficient period to ensure accurate assessment of initial disease presentations.¹⁸ We used an open cohort design, so patients effectively entered the study when they met the inclusion criteria. Patients were censored on the earliest date from among: the date of first CVD presentation, date of death from other causes, date leaving the practice, or date of last practice data collection. (See Figure 1 in the online-only Data Supplement for study flow diagram.)

Risk Factors

The exposures of interest were sex and baseline age, analyzed as 10-year age groups between 30 and 80. A priori confounders were baseline age

as a continuous variable (in analyses estimating associations with sex), smoking status, body mass index, systolic blood pressure, total and high-density lipoprotein cholesterol, diabetes mellitus, socioeconomic status (based on area deprivation measure), use of statins, use of blood pressure medication, and, in women only, use of oral contraceptives or hormone replacement therapy. The baseline value for these confounders was taken as the most recent measurement as recorded during consultations in primary care (CPRD) up to 1 year before study entry. (Detailed definitions are in online-only Data Supplement Methods I.)

End Points

Primary end points were defined as the first recorded diagnosis of the 12 most common symptomatic manifestations of CVD, irrespective of underlying disease mechanism, arising from pathology in the head, heart, abdomen, or legs. The first diagnosis could occur in primary care, secondary care, or at death. We studied the following CVDs: stable angina, unstable angina, nonfatal MI, unheralded coronary death (UCD), heart failure, a composite of cardiac arrest, ventricular arrhythmia, and sudden cardiac death (SCD), transient ischemic attack, ischemic stroke, subarachnoid hemorrhage (SAH), intracerebral hemorrhage, abdominal aortic aneurysm (AAA), peripheral arterial disease (PAD), composite CVD, and other deaths. In secondary analysis, we examined associations in a subset of nonfatal MIs that were classified into ST-segment–elevation MI and non–ST-segment–elevation MI. Coronary heart disease (CHD) and stroke that were not otherwise specified (NOS) were also studied. We classified as fatal events where a death record exists for the same calendar date. (Overview of codes and data sources used to define cardiovascular end points available in online-only Data Supplement Methods II.)

Statistical Analysis

Hazard ratios (HRs) were estimated for the disease-specific Cox proportional-hazards models with length of follow-up as the timescale, stratified by practice, with women as the reference category, and included interactions between age (linear and quadratic term) and sex. Where we estimated the HR for baseline age, we additionally stratified by sex, to allow the baseline hazard to vary. The proportional hazard assumption was tested using Schoenfeld residuals, with no significant effects found.

In the main analyses, we estimated the association of each end point with age groups, the age-adjusted association with sex across all subjects, and by age group in a model with sex interactions. Assuming mutual independence between initial presentations, we assessed heterogeneity in the reported associations based on τ^2 , an estimate of the between-group variance of the log hazard ratio, and a way of summarizing the variability in effect sizes across all the end points in a single statistic.¹⁹

In secondary analysis, we examined whether associations with sex change after adjusting for smoking status, body mass index, diabetes mellitus, systolic blood pressure, total cholesterol, high-density lipoprotein cholesterol, and social deprivation, or additionally for baseline use of blood pressure–lowering medications (diazides, β -blockers, angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, or calcium channel blockers), statins, oral contraceptives and hormone replacement therapy. Missing covariate data were handled by multiple imputation. (Methods used for multiple imputation are described in online-only Data Supplement Methods III.) In sensitivity analyses we studied associations between sex and CVDs (1) ignoring primary care diagnoses and (2) restricting end points to fatal events.

In a post hoc analysis, we assessed the discrimination of age- and sex-adjusted models for each of the 12 end points by calculating the separate concordance index (C-index) for each.²⁰

Approval was granted by the Independent Scientific Advisory Committee of the Medicines and Healthcare Products Regulatory Agency and the Myocardial Ischemia National Audit Project Academic Group. We registered the protocol at clinicaltrials.gov (NCT01164371).

Results

Baseline characteristics of the cohort are shown in the Table. The cohort was young at baseline, as would be expected from a population free from CVD, and 90% were white. Both systolic and diastolic blood pressure increased with age, as did the proportion on blood pressure–lowering medication, with more women than men treated at all ages. More men than women were current or ex-smokers, the proportion of current smokers declining at >60 years of age. Rates of statin prescription were low, but were higher in men than in women at all ages.

Initial CVD Presentations

Over a 6-year median follow-up (interquartile range, 2–10), 114 859 initial CVD presentations were observed (52.3% in men), among which nonfatal MI, UCD, and ischemic or NOS stroke together accounted for 32.5%. The proportion of events varied by sex and age group (Figure 1; Table I in the online-only Data Supplement). The most frequent initial CVD presentation for men was nonfatal MI, which accounted for 27.9% of events in the 30 to 39 age group and more than double the proportion in women in the same age group (11.2%). This proportion declined in men as age increased, becoming similar to that in women in the >80 age group. In contrast, stable angina and unstable angina accounted for similar proportions of initial presentations in both men and women and declined with age. Although evident in younger age groups, heart failure and ischemic stroke as an initial presentation

started to increase in both sexes at age 60 to form the 2 most common initial presentations at age >80.

Associations With Age

The strength and shape of the association of CVDs with age varied from predominantly linear (in angina and nonfatal MI) to strongly quadratic (UCD, stroke, AAA), and from weak (SAH, unstable angina, and cardiac arrest/SCD) to very strong (heart failure and AAA). (See Figure II in the online-only Data Supplement.)

Associations With Sex

SAH was less common in men (HR men versus women, 0.69; 95% confidence interval [CI], 0.59–0.79); other CVDs were positively associated with male sex but with considerable heterogeneity ($\tau^2=0.196$; Figure 2). Specifically, the age-adjusted HR (all $P<0.001$) was <1.5 for transient ischemic attack, intracerebral hemorrhage, and unstable angina, 1.5 to 2.0 for stable angina, ischemic stroke, PAD, heart failure, and cardiac arrest/SCD, and 3.6 to 5.0 for AAA, MI, and UCD. The age-adjusted HR for men versus women was 4.14 (95% CI, 3.72–4.60) in ST-segment–elevation MI and 3.18 (95% CI, 2.86–3.52) in non–ST-segment–elevation MI. These associations changed little after adjustment for conventional CVD risk factors and baseline medications, with the exception of intracerebral hemorrhage, where the association reduced to null (Figure III in the online-only Data Supplement).

Table. Baseline Characteristics in Men and Women by 10-Year Age Groups

	30 to 39		40 to 49		50 to 59		60 to 69		70 to 79		>80	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
No. of patients	418 755	385 486	211 109	194 172	158 327	159 919	97 300	110 478	52 383	77 894	20 455	51 082
Age, y	33.3 (3.2)	33.3 (3.2)	44.5 (2.9)	44.6 (2.9)	54.4 (2.9)	54.5 (2.9)	64.4 (2.9)	64.5 (2.9)	74.2 (2.8)	74.5 (2.9)	85.2 (4.3)	86.2 (4.8)
White	85.7	86.6	89.1	89.2	94.2	93.5	95.1	94.7	96.0	96.6	97.4	98.1
Most deprived quintile	22.8	21.9	19.7	18.6	17.5	16.6	17.4	17.6	18.9	19.4	18.9	18.7
Number of GP visits in previous year	3.6 (4.9)	7.1 (6.8)	4.1 (5.5)	6.2 (6.7)	4.4 (6.0)	6.4 (6.8)	5.6 (6.9)	6.9 (7.3)	6.5 (7.7)	7.2 (7.9)	8.1 (9.3)	8.6 (9.3)
Body mass index, kg/m ²	26.2 (4.5)	25.5 (5.7)	27.1 (4.6)	26.4 (5.9)	27.5 (4.6)	27.1 (5.7)	27.2 (4.3)	27.3 (5.5)	26.3 (4.1)	26.5 (5.2)	24.9 (3.8)	24.4 (4.6)
Systolic BP, mm Hg	126 (14)	117 (13)	130 (15)	124 (16)	136 (17)	133 (18)	141 (18)	141 (18)	145 (19)	148 (20)	144 (21)	147 (22)
Diastolic BP, mm Hg	78 (10)	74 (9)	81 (10)	77 (10)	83 (10)	81 (10)	82 (10)	82 (9)	81 (10)	81 (10)	78 (11)	79 (11)
Total cholesterol, mmol/L	5.3 (1.1)	4.9 (1.0)	5.5 (1.1)	5.3 (1.0)	5.4 (1.1)	5.7 (1.1)	5.2 (1.1)	5.7 (1.2)	4.9 (1.1)	5.6 (1.2)	4.6 (1.1)	5.3 (1.2)
HDL, mmol/L	1.2 (0.3)	1.5 (0.4)	1.2 (0.4)	1.5 (0.4)	1.3 (0.4)	1.6 (0.5)	1.3 (0.4)	1.6 (0.4)	1.3 (0.4)	1.6 (0.5)	1.3 (0.4)	1.6 (0.4)
Ex-smoker	13.2	14.5	15.2	12.7	21.2	14.8	27.8	16.3	32.1	17.6	34.7	16.3
Current smoker	27.5	21.3	25.5	19.4	22.1	17.3	16.5	12.5	10.6	7.9	7.6	4.2
Diabetes mellitus	0.9	0.8	2.1	1.4	3.9	2.6	6.9	4.9	8.5	6.3	8.0	6.2
Statins	0.4	0.2	2.1	1.0	5.0	3.5	9.0	7.9	9.2	8.3	7.2	6.3
BP-lowering medications	4.2	7.6	8.8	14.1	17.9	25.8	29.8	35.4	38.5	43.7	39.0	45.6
Oral contraceptives	...	35.0	...	10.4	...	1.2	...	0.1
HRT	...	0.7	...	9.1	...	30.0	...	15.1	...	3.3	...	0.5

*Mean (standard deviation) for continuous variables and % for categorical variables. Continuous variables are the most recently recorded value in the year before study entry. BP indicates blood pressure; GP, general practice; HDL, high-density lipoprotein cholesterol; and HRT, hormone replacement therapy.

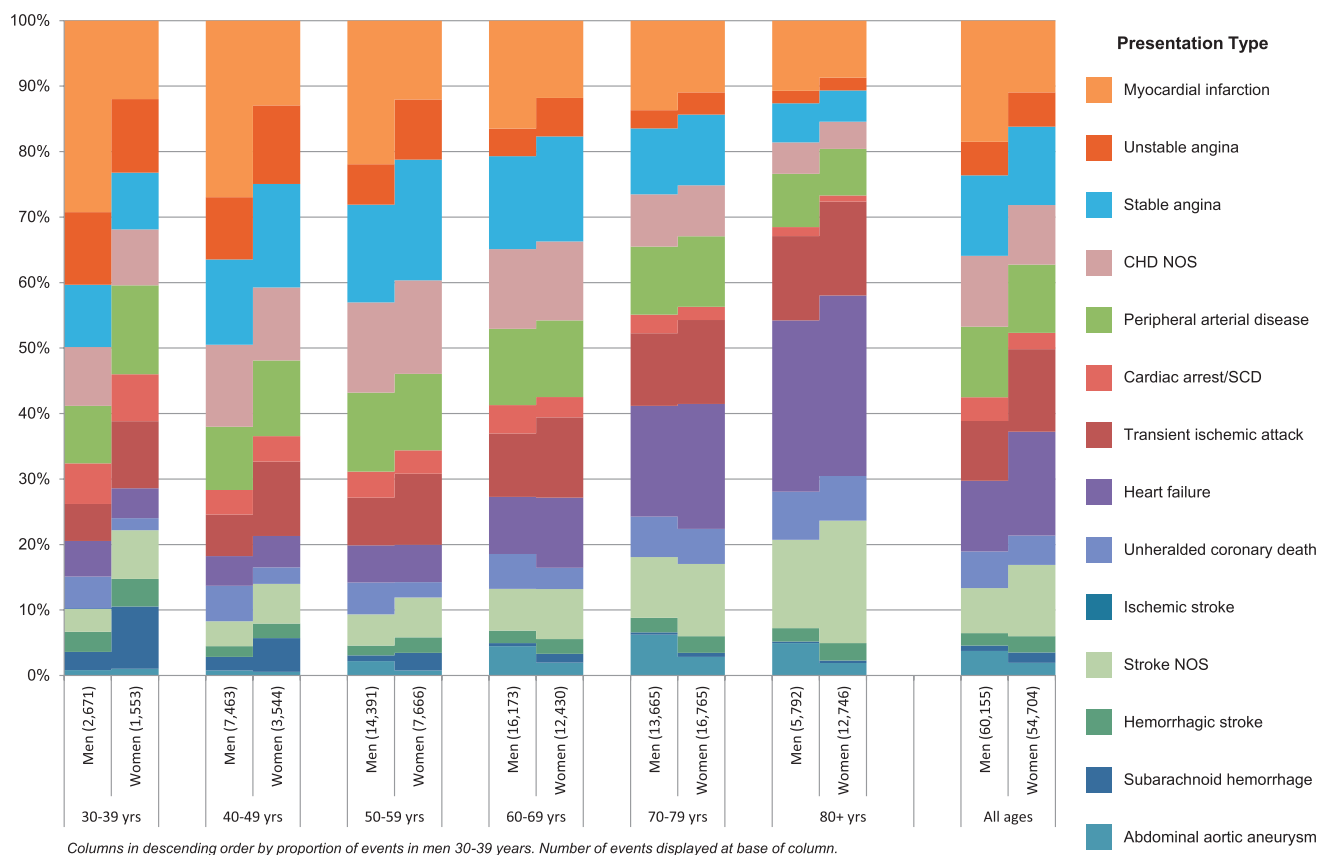


Figure 1. Age and sex distribution of 60 155 events in men and 54 704 in women representing the initial presentation of a wide range of CVDs. CHD indicates coronary heart disease; CVD, cardiovascular disease; NOS, not otherwise specified; and SCD, sudden cardiac death.

Associations between sex and initial CVD presentation were differentially modified by age (Figure 3). The largest differences in HRs for men versus women were observed in the younger (coronary end points) and middle (ischemic stroke, PAD, AAA) age groups. Most dramatically, men <60 years old had an >4-fold higher risk of MI or UCD than similarly aged women. In all cases, associations with male sex diminished with age.

Sensitivity Analyses

The pattern and magnitude of associations with sex were similar in multiply-adjusted analyses to analyses adjusted for age alone (see Figure III in the online-only Data Supplement). Stable angina and PAD were the only initial presentations where the association with male sex differed when the EHRs used were restricted to secondary care and mortality (Figure IV in the online-only Data Supplement).

Discrimination of Age- and Sex-Adjusted Models for Different CVDs

Using disease-specific age and sex coefficients in risk prediction models resulted in markedly different discrimination performance (Figure V in the online-only Data Supplement), with C-indices ranging from very low for SAH (0.57; 95% CI, 0.55–0.59) to relatively high for AAA (0.86; 95% CI, 0.85–0.88) in comparison with a conventional composite CVD model with C-index of 0.73 (95% CI, 0.72–0.73).

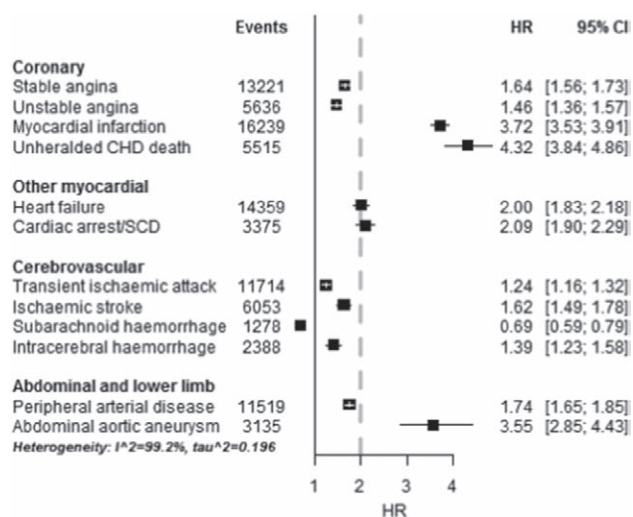
Discussion

Objectives Addressed, Summary of Main Findings

By linking EHRs from multiple sources we curated a cohort of nearly 2 million patients with >100 000 nonfatal and fatal CVD end points of 12 different types. We found that the majority of CVD first presentations are not MI or ischemic stroke but rather heart failure, angina, transient ischemic attack, and PAD. In our contemporary population-based cohort, we find that 51.3% of men and 41.2% of women experienced some form of CVD during their lifetime, with heart failure and stroke (primarily ischemic and NOS) becoming more common as the initial presentations in both men and women in later life. The variable associations of sex and age with different CVDs have important consequences for risk prediction.

Importance of Studying First Manifestations of CVD

We compared the relative frequency of 12 of the most common CVDs affecting atherothrombotic processes in the coronary, cerebral, and peripheral circulations, aneurysms in the cerebral and peripheral circulations, and disorder of myocardial function and cardiac arrhythmia. This family of diseases is clinically relevant, because having one is strongly associated with the subsequent development of another and should initiate a range of secondary preventive interventions.²¹



The vertical grey dotted line corresponds to the HR of the composite CVD endpoint. CHD NOS and Stroke NOS excluded from the main display because non-specific endpoints; their corresponding estimates are HR 2.03 (95% CI, 1.92-2.15; $n=10,895$) and 1.37 (95% CI, 1.26-1.49; $n=9,532$).

Figure 2. Hazard ratios of men in comparison with women for initial presentation of 12 different cardiovascular diseases among a population of 1.93 million adults. CHD indicates coronary heart disease; CI, confidence interval; CVD, cardiovascular disease; HR, hazard ratio; NOS, not otherwise specified; and SCD, sudden cardiac death.

Despite the insights to be gained from considering the first presentation among these diseases together, this first-life-time-presentation approach has rarely been reported in the literature²² and has tended to exclude major diseases such as heart failure, been restricted to small cohorts, or reported in men only.²³⁻²⁷

Innovative Role of Large-Scale Health Record Linkages

Through the use of linked EHRs, we were able to capture diseases first presenting in primary care and were not confined to hospitalized cases. Our cohort is population based, with >99% of the English population estimated to be registered with a family physician,²⁸ unlike other recent large cohort studies, such as UK Biobank, with response rates <10%.²⁹ The size of cohort—nearly 2 million people—possible with this EHR platform allows us to include serious but less commonly diagnosed events such as AAA and SAH, and to have sufficient events in women to study sex differences reliably. We were able to follow up actual events in a clinically meaningful 5-year time frame, similar to the time horizon of randomized, controlled trials. More broadly, we demonstrate the potential of linked EHR cohorts to complement bespoke, investigator-led cohorts. The UK Biobank,³⁰ the Research Program for Genes Environment and Health in Kaiser,³¹ and precision medicine initiatives³² all place a major emphasis on specific disease types and follow-up through health records. Such large EHR cohorts further hold out the promise of lower cost to research funders for data collection, the intrinsic clinical relevance of real-world data, the opportunities to study diseases with higher specificity given the cohort sizes possible, and the prospect for

researchers and clinicians to work across the boundaries that currently impede the translation of new discoveries into public health benefit.³³

We expect the pattern of the age and sex associations we found with the CVD outcomes to apply to the broader UK population and other European populations free from symptomatic CVD. Our patients were drawn from >200 practices representative of the English population. Indeed, a recent article investigating similar questions in a smaller investigator-led Dutch cohort found broadly similar associations, albeit with fewer end points.²² Different cohorts, especially those with more people from differing ethnic groups or differing baseline risk profiles, may well present different associations.

Validity of Risk Factor and Disease Measurements in EHRs

Although a principal strength of this study is the ability to resolve a wide range of CVDs in a large-scale cohort, the principal limitation is the possibility of errors in the individual EHR data sources.^{34,35} However, evidence for the validity of our risk factor and disease end points comes from several sources. First, in this population, using identical phenotypic definitions for these same 12 diseases, we have replicated anticipated risk factor – disease associations with systolic and diastolic blood pressure,¹¹ type 2 diabetes mellitus,¹⁵ smoking,¹⁰ and socio-economic deprivation.¹² These findings support the prospective prognostic validity of both the risk factor and the disease measurements. Second, a recent systematic review of studies validating diagnoses in CPRD found a median positive predictive value of 88% across a wide range of diagnoses,⁷ whereas a separate systematic review found the accuracy of discharge coding in Hospital Episodes Statistics to be 83%.³⁵ Third, the associations we found when considering events from all data sources (Figure V in the online-only Data Supplement) were consistent with those when excluding nonfatal cases or those from primary care. The doctors and coders responsible, and the information on which these diagnoses are based, differ for each data source (primary care, hospital, and death); it was reassuring that the associations were broadly similar. Finally, we¹³ and others³⁶ have demonstrated the validity of using linked data for end point follow-up.

Male Sex as a Risk Factor for Different CVDs

We demonstrate that male sex does not have a common underlying association on the incidence of different CVDs. Rather, the strength of this association is highly variable, ranging from protective for SAH; minor for transient ischemic attack, intracerebral hemorrhage, and unstable angina; moderate for stable angina, ischemic stroke, PAD, heart failure, and cardiac arrest/SCD; and strong for AAA, MI, and UCD. Additionally, we found that these associations change with age, with sex differences in proportion of initial presentation of MI and coronary death reducing with age, and with heart failure and stroke (ischemic and NOS) emerging as the most common initial presentations in both sexes. These findings suggest that stratifying patients into low-, intermediate-, and high-risk groups based on their total and disease-specific risks,³⁷ accompanied by the establishment of new cost-effective treatment thresholds,³⁸ could improve risk management, particularly for diseases such as heart failure and

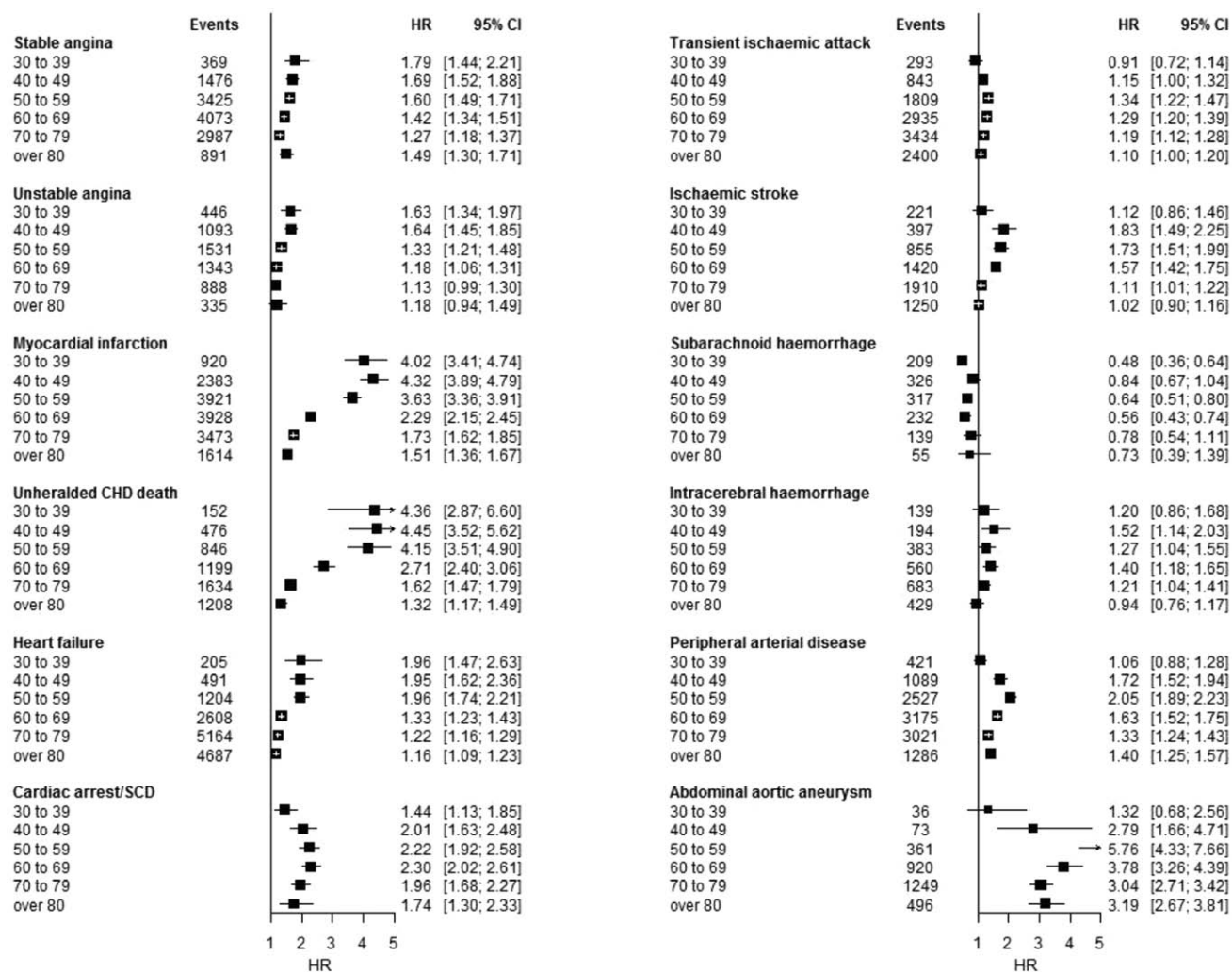


Figure 3. Hazard ratios for men in comparison with women for initial presentation of 12 cardiovascular diseases by baseline age group among a population of 1.93 million adults. CHD indicates coronary heart disease; CI, confidence interval; HR, hazard ratio; NOS, not otherwise specified; and SCD, sudden cardiac death.

stroke that affect high proportions of women but are underman- aged based on current clinical risk assessment.³⁹

Clinical Implications and Risk Prediction

Current risk algorithms in common use focus on CHD⁴⁰ and CVD,⁴¹ as does the new American College of Cardiology/ American Heart Association Guideline on the Assessment of Cardiovascular Risk,⁴² yet we show that chronic disease, such as heart failure and PAD, account for a substantial proportion of initial CVD presentations in contemporary practice. These diseases are associated with marked increased risk of subse- quent events and death, yet have been excluded from many risk prediction algorithms. Given the recent decline in the incidence of acute events of MI and stroke, our findings raise the question of whether risk algorithms should take account of the current burden of CVDs and, in efforts to personalize cardiovascular risk, whether there is a need for algo- rithms tailored to account for specific diseases. For clinical use the latter would only have a role if decisions on preven- tion strategies were altered by using a more specific than a more generic risk prediction tool. Our post hoc analysis of the

discrimination performance of risk prediction models using disease-specific age and sex coefficients supports the impor- tance of having more tailored risk algorithms.

A more nuanced application of age and sex in the clinical setting that takes account of their heterogeneous associations with different CVDs is provided by the following example: A 69-year-old woman with untreated hypertension has a 20% 10-year general risk of CVD, fulfilling guideline criteria for primary prevention. With heart failure her most likely initial CVD presentation within that 10-year time frame (see Figure 1), a tailored blood pressure-lowering regime that excludes calcium antagonists would optimize CVD preven- tion because these drugs are relatively less effective at reduc- ing risk of heart failure.⁴³ At earlier ages, where CHD is the more common initial presentation, the choice of blood pres- sure-lowering medication is likely to make little difference to outcomes. This is just 1 example of the way in which under- standing of the heterogeneity of risks associated with specific end points could lead to more personalized risk modification.

We also provide further evidence of the need to protect women against CVD with the same vigor as for men. The

current strategy of evaluating and treating short-term risk of total CVD has the consequence that almost all men aged >70 should be on treatment, irrespective of their CVD risk factors. However, a wider group of people with high risk of specific CVDs could be targeted and treated earlier by increasing the sensitivity (by extending the time horizon to lifetime, as suggested by the Joint British Societies latest recommendations⁴⁴) and specificity (by using more specific diagnoses) of risk predictions. Given that the majority of initial CVD presentations in our cohort were nonfatal (84% in men and 80% in women), such opportunities for earlier intervention via refinement of prediction tools should not be missed.

Furthermore, our findings have potentially important consequences for the accuracy of models used to predict CVD risk in clinical practice. We found large differences in the associations of different CVDs with age (from very weak with SAH to very strong with UCD, heart failure, stroke, and AAA) and male sex (from negative with SAH to very strong with AAA, nonfatal MI, and UCD). So far, most efforts to improve the prediction of CVD have focused on refining current models with new predictors. Although there are several models for specific CVDs (eg, heart failure,⁴⁵ stroke⁴⁶), current guidelines recommend assessment of total CVD risk to simplify clinical decision making.²¹ Here we show that this one-size-fits-all approach reduces the ability to discriminate between individuals with high and low risk of specific CVDs.

Implications for Research

Our findings suggest that future research on the primary prevention of CVDs should take account of current patterns of disease presentation and redress the imbalance of previous literature that has focused extensively on heart attack and stroke. Our findings have implications for the design and interpretation of observational studies, randomized trials, and meta-analyses investigating the primary prevention of CVDs. Because the fundamental risk factors of age and sex have such heterogeneous associations with different CVDs, and most studies are only sufficiently powered to examine CVD aggregates, it is important to account for the relative proportion of each disease in the composite end point in meta-analysis. Despite an extensive literature on the underlying biological and behavioral pathways by which sex may influence aggregates of CVD and CHD, there is a lack of mechanism studies that investigate why sex has such heterogeneous associations on different CVDs.

Limitations

Our study has important limitations. First, we were not able to resolve some disease subtypes, eg, systolic versus diastolic heart failure or ruptured versus nonruptured cases of AAA. We did find that the association of MI with male sex was more marked for ST-segment–elevation MI than non–ST-segment–elevation MI, suggesting an even greater degree of heterogeneity may be unmasked by investigating more specific diagnoses. Second, we did not evaluate common CVDs that are commonly asymptomatic such as atrial fibrillation. Third, EHRs contain limited covariates for explaining the heterogeneity in sex differences that we report. Fourth, there were 2 less well-specified diagnoses (CHD NOS and stroke

NOS) which we were unable to resolve further, but which we included to ensure all potential initial presentations were taken into account. Stroke NOS is likely to be largely ischemic stroke, based on proportion of strokes that are ischemic¹ and the behavior of this end point in modeling, indicating that we may have overestimated the association of ischemic stroke with male sex. We believe CHD NOS is a mixture of stable and unstable angina given the associations in this article and others, but are unable to substantiate this.

Conclusion

In an era of modern primary prevention, CVDs commonly first present with heart failure, transient ischemic attack, stable angina, and PAD – diseases that have seldom been the focus of primary prevention studies. Predicting CVD risk should take account a wide range of CVDs, and the different association each has with age and sex, as well.

Sources of Funding

This study was supported by the National Institute for Health Research (RP-PG-0407-10314), Wellcome Trust (WT 086091/Z/08/Z), and the Farr Institute of Health Informatics Research, funded by The Medical Research Council (K006584/1), in partnership with Arthritis Research UK, the British Heart Foundation, Cancer Research UK, the Economic and Social Research Council, the Engineering and Physical Sciences Research Council, the National Institute of Health Research, the National Institute for Social Care and Health Research (Welsh Assembly Government), the Chief Scientist Office (Scottish Government Health Directorates), and the Wellcome Trust. Dr George was funded by a National Institute for Health Research Doctoral Fellowship (DRF-2009-02-50). Dr Shah is supported by a Clinical Research Training Fellowship from the Wellcome Trust (0938/30/Z/10/Z). Dr Smeeth is supported by a Wellcome Trust Senior Research Fellowship in Clinical Science. Dr Timmis acknowledges support of St Bartholomew's and the London Cardiovascular Biomedical Research Unit, funded by the National Institute for Health Research. This article presents independent research funded in part by the NIHR. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The authors declare that the funding sources had no role in the conduct, analysis, interpretation and writing of this manuscript.

Disclosures

None.

References

- Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Borden WB, Bravata DM, Dai S, Ford ES, Fox CS, Franco S, Fullerton HJ, Gillespie C, Hailpern SM, Heit JA, Howard VJ, Huffman MD, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH, Lisabeth LD, Magid D, Marcus GM, Marelli A, Matchar DB, McGuire DK, Mohler ER, Moy CS, Mussolino ME, Nichol G, Paynter NP, Schreiner PJ, Sorlie PD, Stein J, Turan TN, Virani SS, Wong ND, Woo D, Turner MB; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2013 update: a report from the American Heart Association. *Circulation*. 2013;127:e6–e245. doi: 10.1161/CIR.0b013e31828124ad.
- Murphy NF, MacIntyre K, Capewell S, Stewart S, Pell J, Chalmers J, Redpath A, Frame S, Boyd J, McMurray JJ. Hospital discharge rates for suspected acute coronary syndromes between 1990 and 2000: population based analysis. *BMJ*. 2004;328:1413–1414. doi: 10.1136/bmj.38111.650741.F7.
- Beaglehole R, Ebrahim S, Reddy S, Voûte J, Leeder S; Chronic Disease Action Group. Prevention of chronic diseases: a call to action. *Lancet*. 2007;370:2152–2157. doi: 10.1016/S0140-6736(07)61700-0.
- Lewington S, Clarke R, Qizilbash N, Peto R, Collins R; Prospective Studies Collaboration. Age-specific relevance of usual blood pressure

- to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet*. 2002;360:1903–1913. doi: 10.1016/S0140-6736(02)11911-8
5. Anand SS, Islam S, Rosengren A, Franzosi MG, Steyn K, Yusufali AH, Keltai M, Diaz R, Rangarajan S, Yusuf S; INTERHEART Investigators. Risk factors for myocardial infarction in women and men: insights from the INTERHEART study. *Eur Heart J*. 2008;29:932–940. doi: 10.1093/eurheartj/ehn018.
 6. Hyvärinen M, Qiao Q, Tuomilehto J, Söderberg S, Eliasson M, Stehouwer CD. The difference between acute coronary heart disease and ischaemic stroke risk with regard to gender and age in Finnish and Swedish populations. *Int J Stroke*. 2010;5:152–156. doi: 10.1111/j.1747-4949.2010.00423.x.
 7. Kappert K, Böhm M, Schmieder R, Schumacher H, Teo K, Yusuf S, Sleight P, Unger T; ONTARGET/TRANSCEND Investigators. Impact of sex on cardiovascular outcome in patients at high cardiovascular risk: analysis of the Telmisartan Randomized Assessment Study in ACE-Intolerant Subjects With Cardiovascular Disease (TRANSCEND) and the Ongoing Telmisartan Alone and in Combination With Ramipril Global End Point Trial (ONTARGET). *Circulation*. 2012;126:934–941. doi: 10.1161/CIRCULATIONAHA.111.086660.
 8. Blumenthal D, Tavenner M. The “meaningful use” regulation for electronic health records. *N Engl J Med*. 2010;363:501–504. doi: 10.1056/NEJMp1006114.
 9. Gaziano JM. The evolution of population science: advent of the mega cohort. *JAMA*. 2010;304:2288–2289. doi: 10.1001/jama.2010.1691.
 10. Pujades-Rodriguez M, George J, Shah AD, Rapsomaniki E, Denaxas S, West R, Smeeth L, Timmis A, Hemingway H. Heterogeneous associations between smoking and a wide range of initial presentations of cardiovascular disease in 1937360 people in England: lifetime risks and implications for risk prediction. *Int J Epidemiol*. 2015;44:129–141. doi: 10.1093/ije/dyu218.
 11. Rapsomaniki E, Timmis A, George J, Pujades-Rodriguez M, Shah AD, Denaxas S, White IR, Caulfield MJ, Deanfield JE, Smeeth L, Williams B, Hingorani A, Hemingway H. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1.25 million people. *Lancet*. 2014;383:1899–1911. doi: 10.1016/S0140-6736(14)60685-1.
 12. Pujades-Rodriguez M, Timmis A, Stogiannis D, Rapsomaniki E, Denaxas S, Shah A, Feder G, Kivimaki M, Hemingway H. Socioeconomic deprivation and the incidence of 12 cardiovascular diseases in 1.9 million women and men: implications for risk prediction and prevention. *PLoS One*. 2014;9:e104671. doi: 10.1371/journal.pone.0104671.
 13. Herrett E, Shah AD, Boggon R, Denaxas S, Smeeth L, van Staa T, Timmis A, Hemingway H. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ*. 2013;346:f2350. doi: 10.1136/bmj.f2350.
 14. Denaxas SC, George J, Herrett E, Shah AD, Kalra D, Hingorani AD, Kivimaki M, Timmis AD, Smeeth L, Hemingway H. Data resource profile: cardiovascular disease research using linked bespoke studies and electronic health records (CALIBER). *Int J Epidemiol*. 2012;41:1625–1638. doi: 10.1093/ije/dys188.
 15. Shah AD, Langenberg C, Rapsomaniki E, Denaxas S, Pujades-Rodriguez M, Gale CP, Deanfield J, Smeeth L, Timmis A, Hemingway H. Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1.9 million people. *Lancet Diabetes Endocrinol*. 2015;3:105–113. doi: 10.1016/S2213-8587(14)70219-0.
 16. Gallagher AM, Puri S, Staa TV. Linkage of the General Practice Research Database (GPRD) with other data sources. *Pharmacoepidemiol Drug Saf*. 2011;S230–S364.
 17. Mathur R, Bhaskaran K, Chaturvedi N, Leon DA, vanStaa T, Grundy E, Smeeth L. Completeness and usability of ethnicity data in UK-based primary care and hospital databases. *J Public Health (Oxf)*. 2014;36:684–692. doi: 10.1093/pubmed/fdt116.
 18. Lewis JD, Bilker WB, Weinstein RB, Strom BL. The relationship between time since registration and measured incidence rates in the General Practice Research Database. *Pharmacoepidemiol Drug Saf*. 2005;14:443–451. doi: 10.1002/pds.1115.
 19. Thompson SG, Pyke SD, Hardy RJ. The design and analysis of paired cluster randomized trials: an application of meta-analysis techniques. *Stat Med*. 1997;16:2063–2079. doi: 10.1002/(SICI)1097-0258(19970930)
 20. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361–387. doi: 10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4.
 21. Department of Health. *Cardiovascular Disease Outcomes Strategy: Improving outcomes for people at risk or with cardiovascular disease*. 2013. <https://www.gov.uk/government/publications/improving-cardiovascular-disease-outcomes-strategy>. Accessed September 19, 2014.
 22. Leening MJ, Ferket BS, Steyerberg EW, Kavousi M, Deckers JW, Nieboer D, Heeringa J, Portegies ML, Hofman A, Ikram MA, Hunink MG, Franco OH, Stricker BH, Witteman JC, Roos-Hesselink JW. Sex differences in lifetime risk and first manifestation of cardiovascular disease: prospective population based cohort study. *BMJ*. 2014;349:g5992. doi: 10.1136/bmj.g5992.
 23. Rothwell PM, Coull AJ, Silver LE, Fairhead JF, Giles MF, Lovelock CE, Redgrave JN, Bull LM, Welch SJ, Cuthbertson FC, Binney LE, Gutnikov SA, Anslow P, Banning AP, Mant D, Mehta Z; Oxford Vascular Study. Population-based study of event rate, incidence, case fatality, and mortality for all acute vascular events in all arterial territories (Oxford Vascular Study). *Lancet*. 2005;366:1773–1783. doi: 10.1016/S0140-6736(05)67702-1.
 24. Baena-Diez JM, Vidal-Solsona M, Byram AO, González-Casafont I, Ledesma-Ulloa G, Martí-Sans N. The epidemiology of cardiovascular disease in primary care: the Zona Franca Cohort study in Barcelona, Spain. *Rev Esp Cardiol*. 2010;63:1261–1269.
 25. Canoui-Poitrine F, Luc G, Juhan-Vague I, Morange PE, Arveiler D, Ferrieres J, Amouyel P, Bingham A, Montaye M, Ruidavets JB, Haas B, Evans A, Ducimetiere P, Empana JP; PRIME Study Group. Respective contribution of conventional risk factors and antihypertensive treatment to stable angina pectoris and acute coronary syndrome as the first presentation of coronary heart disease: the PRIME Study. *Eur J Cardiovasc Prev Rehabil*. 2009;16:550–555. doi: 10.1097/HJR.0b013e32832c88d1.
 26. Glynn RJ, Rosner B. Comparison of risk factors for the competing risks of coronary heart disease, stroke, and venous thromboembolism. *Am J Epidemiol*. 2005;162:975–982. doi: 10.1093/aje/kwi309.
 27. Ducimetière P, Ruidavets JB, Montaye M, Haas B, Yarnell J; PRIME Study Group. Five-year incidence of angina pectoris and other forms of coronary heart disease in healthy men aged 50–59 in France and Northern Ireland: the Prospective Epidemiological Study of Myocardial Infarction (PRIME) Study. *Int J Epidemiol*. 2001;30:1057–1062. doi: 10.1093/ije/30.5.1057
 28. Social Exclusion Task Force. *Improving the Way We Meet the Primary Health Care Needs of the Socially Excluded*. London: Cabinet Office; 2010.
 29. Manolio TA, Weis BK, Cowie CC, Hoover RN, Hudson K, Kramer BS, Berg C, Collins R, Ewart W, Gaziano JM, Hirschfeld S, Marcus PM, Masys D, McCarty CA, McLaughlin J, Patel AV, Peakman T, Pedersen NL, Schaefer C, Scott JA, Sprosen T, Walport M, Collins FS. New models for large prospective studies: is there a better way? *Am J Epidemiol*. 2012;175:859–866. doi: 10.1093/aje/kwr453.
 30. Collins R. What makes UK Biobank special? *Lancet*. 2012;379:1173–1174. doi: 10.1016/S0140-6736(12)60404-8.
 31. Risch N, Schaefer C. Kaiser Permanente Research Program on Genes, Environment and Health. 2011. www.rpgeh.kaiser.org. Accessed September 19, 2014.
 32. National Research Council (US) Committee on A Framework for Developing a New Taxonomy of Disease. *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease*. Washington, DC: National Academies Press; 2013.
 33. Lauer MS. Time for a creative transformation of epidemiology in the United States. *JAMA*. 2012;308:1804–1805. doi: 10.1001/jama.2012.14838.
 34. Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol*. 2010;69:4–14. doi: 10.1111/j.1365-2125.2009.03537.x.
 35. Burns EM, Rigby E, Mamidanna R, Bottle A, Aylin P, Ziprin P, Faiz OD. Systematic review of discharge coding accuracy. *J Public Health (Oxf)*. 2012;34:138–148. doi: 10.1093/pubmed/fdr054.
 36. Payne RA, Abel GA, Simpson CR. A retrospective cohort study assessing patient characteristics and the incidence of cardiovascular disease using linked routine primary and secondary care data. *BMJ Open*. 2012;2:e000723. doi: 10.1136/bmjopen-2011-000723.
 37. Marma AK, Berry JD, Ning H, Persell SD, Lloyd-Jones DM. Distribution of 10-year and lifetime predicted risks for cardiovascular disease in US adults: findings from the National Health and Nutrition Examination Survey 2003 to 2006. *Circ Cardiovasc Qual Outcomes*. 2010;3:8–14. doi: 10.1161/CIRCOUTCOMES.109.869727.
 38. Hippisley-Cox J, Coupland C, Robson J, Brindle P. Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of

- cardiovascular disease: cohort study using QResearch database. *BMJ*. 2010;341:c6624. doi: 10.1136/bmj.c6624.
39. Mosca L, Benjamin EJ, Berra K, Bezanson JL, Dolor RJ, Lloyd-Jones DM, Newby LK, Piña IL, Roger VL, Shaw LJ, Zhao D, Beckie TM, Bushnell C, D'Armiento J, Kris-Etherton PM, Fang J, Ganiats TG, Gomes AS, Gracia CR, Haan CK, Jackson EA, Judelson DR, Kelepouris E, Lavie CJ, Moore A, Nussmeier NA, Ofili E, Oparil S, Ouyang P, Pinn VW, Sherif K, Smith SC Jr, Sopko G, Chandra-Stroboş N, Urbina EM, Vaccarino V, Wenger NK. Effectiveness-based guidelines for the prevention of cardiovascular disease in women—2011 update: a guideline from the American Heart Association. *Circulation*. 2011;123:1243–1262. doi: 10.1161/CIR.0b013e31820faaf8.
 40. Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults. Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA*. 2001;285:2486–2497. doi:10.1001/jama.285.19.2486.
 41. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ*. 2007;335:136. doi: 10.1136/bmj.39261.471806.55.
 42. Goff DC Jr, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB Sr, Gibbons R, Greenland P, Lackland DT, Levy D, O'Donnell CJ, Robinson JG, Schwartz JS, Shero ST, Smith SC Jr, Sorlie P, Stone NJ, Wilson PWF. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;129(suppl 2):S49–S73.
 43. Turnbull F; Blood Pressure Lowering Treatment Trialists' Collaboration. Effects of different blood-pressure-lowering regimens on major cardiovascular events: results of prospectively-designed overviews of randomised trials. *Lancet*. 2003;362:1527–1535. doi: 10.1016/S0140-6736(03)14739-3
 44. JBS3 Board. Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3). *Heart*. 2014;100(suppl):ii1–ii67.
 45. Aaronson KD, Cowger J. Heart failure prognostic models: why bother? *Circ Heart Fail*. 2012;5:6–9. doi:10.1161/CIRCHEARTFAILURE.111.965848.
 46. Hippisley-Cox J, Coupland C, Brindle P. Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: a prospective open cohort study. *BMJ*. 2013;346:f2573. doi: 10.1136/bmj.f2573.

CLINICAL PERSPECTIVES

The first lifetime presentation of cardiovascular disease in men and women in the 21st century is not currently well understood, with contemporary studies of sufficient size and clinical resolution to distinguish the most common cardiovascular diseases (CVDs) lacking. Traditional cohort studies have, to date, commonly focused on incident heart attack and stroke, but it is well recognized that both have been rapidly declining in incidence. Understanding how CVDs first present is important for developing primary prevention strategies that protect against specific phenotypes and against the wider cascade of other CVDs that often follow. Electronic health records based on usual clinical practice in unselected, contemporary populations provide an important opportunity to assess how CVD first presents in women and men across a wide range of 12 different diseases affecting the head, heart, abdominal, and peripheral circulations. In a study of 1.9 million adults, 114 859 people experienced an incident cardiovascular diagnosis, the majority (66%) of which were neither myocardial infarction nor ischemic stroke. Sex has differing associations with different CVDs, with implications for risk prediction and management strategies. Chronic disease, such as heart failure and peripheral arterial disease, account for a substantial proportion of initial lifetime CVD presentations, yet are been excluded from many risk prediction algorithms. Given the recent decline in the incidence of acute events, our findings emphasize the relevance of risk algorithms that take account of the current burden of CVDs.

Go to <http://cme.ahajournals.org> to take the CME quiz for this article.

SUPPLEMENTAL MATERIAL

Content

Supplemental methods

eText 1 Definition of risk factors

eText 2 Overview of codes and data sources used to define each cardiovascular endpoints

eText 3 Multiple imputation

Supplemental tables

eTable 1 Number and proportion of total events for wide range of cardiovascular diseases by gender and age group

Supplemental figures

eFigure 1 Study flow diagram

eFigure2 Hazard ratios of the association of baseline age with initial presentations of twelve different cardiovascular diseases, with age 40-49 years as reference

eFigure 3 Hazard ratios for men compared to women for initial presentations of twelve different cardiovascular diseases, adjusted for age, selected cardiovascular risk factors, and medications

eFigure 4 Age-adjusted hazard ratios for men compared to women for initial presentation of twelve cardiovascular diseases by different data sources

eFigure 5 The C-index of risk predictions based on CVD phenotype-specific models, adjusted for age & sex only

Supplemental References

Supplemental methods

eText1 Definition of risk factors

All risk factors were based on data recorded in CPRD during primary care consultations in the year prior to the index date, unless otherwise specified. Where multiple blood test results were recorded, the values from the test closest in time to the index date was used.

Medications were deemed to be regular medications if patients had at least two repeat prescriptions, covering a two-month supply, in the year prior to the index date.

Patients' age was measured in years as at the index date. Socioeconomic status was based on an area measure of deprivation, the Index of Multiple Deprivation,¹ linked to their postcode (which was removed prior to receipt of the data). The most recent smoking record prior to the index date was used to classify patients as never, ex- or current smokers.

Patients identified as current smokers with no smoking record within the three years before study entry were reclassified as having missing smoking data. Never smokers who had a previous record of smoking were reclassified as ex-smokers. Body mass index (kg/m²) was calculated using weight measurement closest in time to the index date. Patients were defined as diabetic if they ever had a diagnosis of diabetes or were receiving regular prescriptions for either insulin or metformin. Baseline SBP was based on readings taken during routine primary care consultations; where more than one measurement was taken on the same day, the average measurement was used. Both total cholesterol and high density lipoprotein were defined as the level in routine blood tests, in mmol/L.

Blood-pressure-lowering medications included in our definition are thiazide diuretics, beta-blockers, angiotensin converting enzyme-inhibitors, angiotensin receptor blockers, or calcium-channel blockers. Hormonal therapy (in women only) included combined oral contraceptives, progestogen-only oral contraceptives or hormone replacement therapy. Statins included atorvastatin, fluvastatin, pravastatin, rosuvastatin, or simvastatin.

Additionally the following co-variables were used for the multiple imputation:

- White cell count , haemoglobin, creatinine, alanine transferase from blood tests requested during routine clinical care;
- Regular prescriptions (repeat prescriptions covering period of at least two months) for aspirin; and
- Medical history of depression, cancer, renal disease, liver disease or chronic obstructive pulmonary disease.

eText 2 Overview of codes and data sources used to define each cardiovascular endpoints. (Further details given in the CALIBER data portal, www.caliberresearch.org/portal)

Endpoint	CPRD – Read codes	MINAP – specific disease registry	HES – OPCS 4 hospital procedures	HES – ICD 10 hospital diagnoses†	ONS – ICD 10 causes of death‡
Stable angina	<p>G33..00: Stable Angina.</p> <p>G33z.00: Angina pectoris NOS + 25 other codes for diagnosis of stable angina pectoris.</p> <p>30 codes for evidence of coronary artery disease at angiography (CT,MR, invasive or not specified).</p> <p>151 Read codes for evidence of myocardial ischaemia (Resting ECG, exercise ECG, stress echo, radioisotope scan).</p> <p>Two or more successive prescriptions for anti-anginals.</p>	nu	<p>K40-K46: Coronary artery bypass graft.</p> <p>K49,K50 and K75: Percutaneous coronary intervention, not within 30 days of an acute coronary syndrome.</p>	I20: Stable angina pectoris excluding unstable angina (I20.0).	nu
Unstable angina	<p>G311.13/G311100: Unstable angina.</p> <p>G233200: Angina at rest.</p> <p>G311400: Worsening angina + 13 other codes.</p>	<p>Discharge diagnosis of unstable angina, no raised ST elevation.</p> <p>No raised troponin levels.</p>	nu	<p>I20.0: Unstable or worsening angina.</p> <p>I24: Acute ischaemic heart disease.</p> <p>I24.0: Coronary thrombosis not resulting in MI.</p>	nu

Endpoint	CPRD – Read codes	MINAP – specific disease registry	HES – OPCS 4 hospital procedures	HES – ICD 10 hospital diagnoses†	ONS – ICD 10 causes of death‡
				I24.8: Other forms of ischaemic heart disease. I24.9: Acute ischaemic heart disease, unspecified.	
Coronary heart disease not otherwise specified	G3...00: Ischaemic heart disease + 90 other codes including CHD NOS, chronic ischaemic heart disease, silent myocardial infarction.	nu	nu	CHD NOS, chronic ischaemic heart disease, silent MI (I25) excluding I25.2, old MI.	nu
Acute Myocardial Infarction (MI)	G30X000: Acute ST segment elevation myocardial infarction. G307100: Acute non-ST segment elevation myocardial infarction. G30..14: Heart attack. G30..15: MI Acute myocardial infarction + 60 other codes as Acute MI not otherwise specified.	MI with or without ST elevation based on initial electrocardiogram findings, raised troponins and clinical diagnosis.	nu	I21: Acute myocardial infarction. I23: Current complications of acute MI.	nu

Endpoint	CPRD – Read codes	MINAP – specific disease registry	HES – OPCS 4 hospital procedures	HES – ICD 10 hospital diagnoses†	ONS – ICD 10 causes of death‡
Unheralded coronary death	Any CVD excluded.	Any CVD excluded.	Any CVD excluded.	Any CVD excluded.	I20: Angina Pectoris. I21: Acute MI. I22: Subsequent MI. I23: Certain current complications following acute MI. I24: Other acute ischaemic heart diseases. I25: Chronic ischaemic heart disease.
Heart failure	G58..00: Heart Failure + 92 other Read codes for heart failure diagnosis.	nu	nu	I50: Heart failure. I11.0: Hypertensive heart disease with (congestive) heart failure. I13.0: Hypertensive heart and renal disease with (congestive) heart failure. I13.2: Hypertensive heart and renal disease with both (congestive) heart failure and renal disease.	I50 Heart failure. I11.0 Hypertensive heart disease with (congestive) heart failure. I13.0: Hypertensive heart and renal disease with (congestive) heart failure. I13.2: Hypertensive heart and renal disease with both (congestive) heart failure and renal disease

Endpoint	CPRD – Read codes	MINAP – specific disease registry	HES – OPCS 4 hospital procedures	HES – ICD 10 hospital diagnoses†	ONS – ICD 10 causes of death‡
Ventricular arrhythmias, cardiac arrest and sudden cardiac death	G574.00: Ventricular fibrillation and flutter. G757.00: Cardiac arrest + 35 other Read codes for ventricular fibrillation, asystole, cardiac arrest, cardiac resuscitation, electro-mechanical dissociation. G575100: Sudden cardiac death.	nu	X50: Implanted cardiac defibrillation device. K59: Implantation, revision and renewal of cardiac defibrillator.	I46: Cardiac arrest. I47.0: Re-entry ventricular arrhythmia. I47.2: Ventricular tachycardia.	I46: Cardiac arrest. I47.0: Re-entry ventricular arrhythmia. I47.2: Ventricular tachycardia.
Transient ischaemic attack	Fyu5500: [X]Other transient cerebral ischaemic attacks + related symptoms + 5 other Read codes.	nu	nu	G458: Other transient cerebral ischaemic attacks and related syndromes. G459: Transient cerebral ischaemic attack, unspecified.	nu
Ischaemic stroke	G64..11: CVA – cerebral artery occlusion, G64..13 Stroke due to cerebral arterial occlusion. G6W..00: Cerebral infarction due to unspecified occlusion/stenosis of precerebral arteries. G6X..00: Cerebral	nu	Stroke NOS with carotid endarterectomy or stenting within 90 days (OPCS codes L294, L295, L311, L314; Read codes 7A20300 + 4 others).	I63: Cerebral infarction.	I63: Cerebral infarction.

Endpoint	CPRD – Read codes	MINAP – specific disease registry	HES – OPCS 4 hospital procedures	HES – ICD 10 hospital diagnoses†	ONS – ICD 10 causes of death‡
	infarction due to unspecified occlusion/stenosis of cerebral arteries plus 8 other codes.				
Subarachnoid haemorrhage	G601.00: Subarachnoid haemorrhage from carotid siphon and bifurcation. G602.00: Subarachnoid haemorrhage from middle cerebral artery. G60X.00: Subarachnoid haemorrhage from intracranial artery, unspecified.	nu	nu	I60: Subarachnoid haemorrhage.	I60: Subarachnoid haemorrhage.
Intracerebral haemorrhage	Gyu6F00: [x] Intracerebral haemorrhage in hemisphere, unspecified + 16 other codes.	nu	nu	I61: Intracerebral haemorrhage.	I61: Intracerebral haemorrhage.
Stroke not otherwise specified	G66..11: Cerebrovascular accident unspecified + 14 other Read codes.	nu	U54.3: Delivery of rehabilitation for stroke.	I64: Stroke not specified as haemorrhage or infarction. G463-G467: Stroke syndromes.	I64: Stroke not specified as haemorrhage or infarction. I672: Cerebral atherosclerosis. I679: Cerebrovascular disease, unspecified.

Endpoint	CPRD – Read codes	MINAP – specific disease registry	HES – OPCS 4 hospital procedures	HES – ICD 10 hospital diagnoses†	ONS – ICD 10 causes of death‡
Peripheral arterial disease	63 codes for lower limb peripheral arterial disease diagnosis (including diabetic PAD, gangrene, arterial thrombosis of the leg and intermittent claudication). Evidence of atherosclerosis of iliac and lower limb arteries based on angiography or Dopplers.	nu	L50-L54: Bypass, reconstruction and other open operations on iliac artery. L58-L60, L62: Bypass, reconstruction, transluminal operations or other open operations of femoral artery. L65: Revision of reconstruction of artery.	I70.2: atherosclerosis of arteries of extremities. I73.9: Peripheral vascular disease intermittent claudication E10.05,E11-E14: Peripheral complications of diabetes including gangrene, insulin dependent diabetes mellitus, non-insulin-dependent diabetes mellitus, malnutrition-related diabetes mellitus, other specified diabetes mellitus, unspecified diabetes mellitus.	I70.2: Atherosclerosis of arteries of extremities. I73.9: Peripheral vascular disease intermittent claudication. Peripheral complications of diabetes including gangrene 0.5 suffix of E10: Insulin dependent diabetes mellitus, E11: Non-insulin-dependent diabetes mellitus, E12: Malnutrition-related diabetes mellitus, E13: Other specified diabetes mellitus, E14: Unspecified diabetes mellitus
Abdominal aortic aneurysm	G714.00: AAA without mention of rupture + 12 more codes for AAA diagnosis. 42 codes for AAA procedures.	nu	L16: Extra anatomic bypass of aorta. L18-L23: Replacement of aneurysmal segment of aorta, bypass of segment of aorta, plastic repair of aorta. L25-L28: Transluminal or endovascular insertion of stent on aneurysmal segment of	I71.3: Ruptured AAA. 171.4: AAA without rupture. I71.5: Ruptured thoraco-abdominal aortic aneurysm. I71.6: Thoracoabdominal aortic aneurysm without mention of rupture. I71.8: Aortic aneurysm of	I71.3: Ruptured AAA. I71.4: AAA without rupture. I71.5: Ruptured thoraco-abdominal aortic aneurysm. I71.6: Thoracoabdominal aortic aneurysm without mention of rupture.

Endpoint	CPRD – Read codes	MINAP – specific disease registry	HES – OPCS 4 hospital procedures	HES – ICD 10 hospital diagnoses†	ONS – ICD 10 causes of death‡
			aorta.	unspecified site, ruptured. I71.9: Aortic aneurysm of unspecified site, without mention of rupture.	I71.8: Aortic aneurysm of unspecified site, ruptured. I71.9: Aortic aneurysm of unspecified site, without mention of rupture.

Note: AAA, aortic abdominal aneurysm; CVD, cardiovascular disease; MI, myocardial infarction; NOS, not otherwise specified; nu = not used in definition; OPCS, Office of Population Censuses and Surveys Classification of Interventions and Procedures. †Primary cause of admission. ‡Underlying cause of death.

eText 3 Multiple imputation

Multiple imputation² was implemented using the *mice* algorithm in the statistical package R.

Imputation models were estimated separately for men and women and included:

a) all the baseline covariates used in the main analysis (age, quadratic age, index of multiple deprivation, smoking, body mass index, diabetes, systolic blood pressure, total cholesterol, and HDL cholesterol);

b) prior (between 1 and 4 years before study entry) and post (between 0 and 1 year after study entry) averages of continuous main analysis covariates and other measurements not in the main analysis (white cell count, haemoglobin, creatinine, alanine transferase);

c) baseline medications (statins, blood pressure medications, aspirin, and oral contraceptives and hormone replacement therapy (in women only));

d) coexisting medical conditions (history of depression, cancer, renal disease, liver disease and chronic obstructive pulmonary disease);

e) the Nelson-Aalen hazard and the event status for each endpoint analysed in the data³.

Non-normally distributed variables were log-transformed for imputation and exponentiated back to their original scale for analysis. Five multiply imputed datasets were generated, and Cox models fitted to each dataset. Coefficients were combined using Rubin's rules.

We checked whether the imputations were plausible by comparing plots of the distribution of observed and imputed values of all variables.

Supplemental Tables

eTable 1. Age and sex distribution of 60,155 events in men and 54,704 in women representing the first lifetime presentations of range of CVDs

Presentation	30-39 years		40-49 years		50-59 years		60-69 years		70-79 years		80+years		All ages	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Myocardial infarction	746 (27.9)	174 (11.2)	1,942 (26.0)	441 (12.4)	3,038 (21.1)	883 (11.5)	2,536 (15.7)	1,392 (11.2)	1,765 (12.9)	1,708 (10.2)	584 (10.1)	1,030 (8.1)	10,611 (17.6)	5,628 (10.3)
Unstable angina	512 (19.2)	287 (18.5)	1,589 (21.3)	788 (22.2)	2,767 (19.2)	1,723 (22.5)	2,523 (15.6)	2,117 (17.0)	1,391 (10.2)	1,746 (10.4)	365 (6.3)	723 (5.7)	9,147 (15.2)	7,384 (13.5)
Stable angina	243 (9.1)	126 (8.1)	938 (12.6)	538 (15.2)	2,071 (14.4)	1,354 (17.7)	2,182 (13.5)	1,891 (15.2)	1,301 (9.5)	1,686 (10.1)	325 (5.6)	566 (4.4)	7,060 (11.7)	6,161 (11.3)
Coronary heart disease not otherwise specified	229 (8.6)	124 (8.0)	904 (12.1)	380 (10.7)	1,911 (13.3)	1,048 (13.7)	1,873 (11.6)	1,424 (11.5)	1,033 (7.6)	1,216 (7.3)	261 (4.5)	492 (3.9)	6,211 (10.3)	4,684 (8.6)
Peripheral arterial disease	224 (8.4)	197 (12.7)	695 (9.3)	394 (11.1)	1,671 (11.6)	856 (11.2)	1,795 (11.1)	1,380 (11.1)	1,339 (9.8)	1,682 (10.0)	444 (7.7)	842 (6.6)	6,168 (10.3)	5,351 (9.8)
Ischemic stroke	209 (7.8)	209 (13.5)	531 (7.1)	345 (9.7)	1,190 (8.3)	772 (10.1)	1,766 (10.9)	1,534 (12.3)	1,951 (14.3)	2,880 (17.2)	1,079 (18.6)	3,119 (24.5)	6,726 (11.2)	8,859 (16.2)
Cardiac arrest/ Sudden cardiac death	159 (6.0)	104 (6.7)	271 (3.6)	132 (3.7)	550 (3.8)	262 (3.4)	667 (4.1)	366 (2.9)	366 (2.7)	312 (1.9)	74 (1.3)	112 (0.9)	2,087 (3.5)	1,288 (2.4)
Transient ischemic attack	144 (5.4)	149 (9.6)	456 (6.1)	387 (10.9)	1,012 (7.0)	797 (10.4)	1,488 (9.2)	1,447 (11.6)	1,432 (10.5)	2,002 (11.9)	702 (12.1)	1,698 (13.3)	5,234 (8.7)	6,480 (11.8)
Heart failure	139 (5.2)	066 (4.2)	328 (4.4)	163 (4.6)	784 (5.4)	420 (5.5)	1,345 (8.3)	1,263 (10.2)	2,185 (16.0)	2,979 (17.8)	1,425 (24.6)	3,262 (25.6)	6,206 (10.3)	8,153 (14.9)
Unheralded coronary death	125 (4.7)	27 (1.7)	390 (5.2)	86 (2.4)	674 (4.7)	172 (2.2)	816 (5.0)	383 (3.1)	799 (5.8)	835 (5.0)	400 (6.9)	808 (6.3)	3,204 (5.3)	2,311 (4.2)
Stroke not otherwise specified	89 (3.3)	108 (7.0)	273 (3.7)	206 (5.8)	660 (4.6)	447 (5.8)	982 (6.1)	898 (7.2)	1,198 (8.8)	1,723 (10.3)	735 (12.7)	2,213 (17.4)	3,937 (6.5)	5,595 (10.2)
Hemorrhagic stroke	78 (2.9)	61 (3.9)	118 (1.6)	76 (2.1)	209 (1.5)	174 (2.3)	294 (1.8)	266 (2.1)	287 (2.1)	396 (2.4)	112 (1.9)	317 (2.5)	1,098 (1.8)	1,290 (2.4)
Subarachnoid hemorrhage	71 (2.7)	138 (8.9)	151 (2.0)	175 (4.9)	120 (0.8)	197 (2.6)	72 (0.4)	160 (1.3)	44 (0.3)	95 (0.6)	12 (0.2)	43 (0.3)	470 (0.8)	808 (1.5)
Abdominal aortic aneurysm	21 (0.8)	15 (1.0)	54 (0.7)	19 (0.5)	305 (2.1)	56 (0.7)	689 (4.3)	231 (1.9)	805 (5.9)	444 (2.6)	270 (4.7)	226 (1.8)	2,144 (3.6)	991 (1.8)

NB Columns in descending order of proportion by ordering in men 30-39 years.

Supplemental Figures

eFigure 1 Study flow diagram

eFigure 2 Hazard ratios of the association of baseline age with initial presentations of twelve different cardiovascular diseases, with age 40-49 years as reference

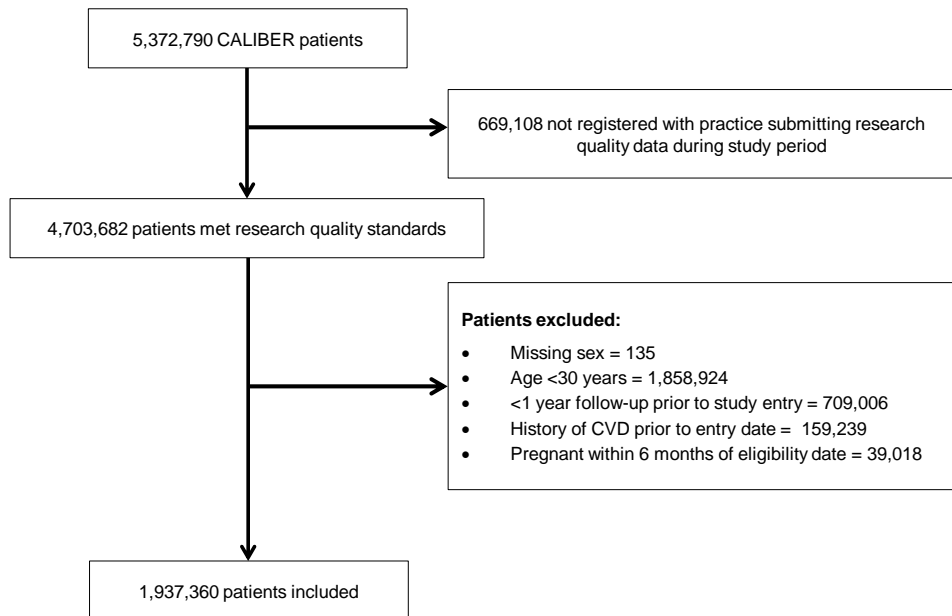
eFigure 3 Hazard ratios for men compared to women for initial presentations of twelve different cardiovascular diseases, adjusted for age, selected cardiovascular risk factors, and medications

eFigure 4 Age-adjusted hazard ratios for men compared to women for initial presentation of twelve cardiovascular diseases by different data sources

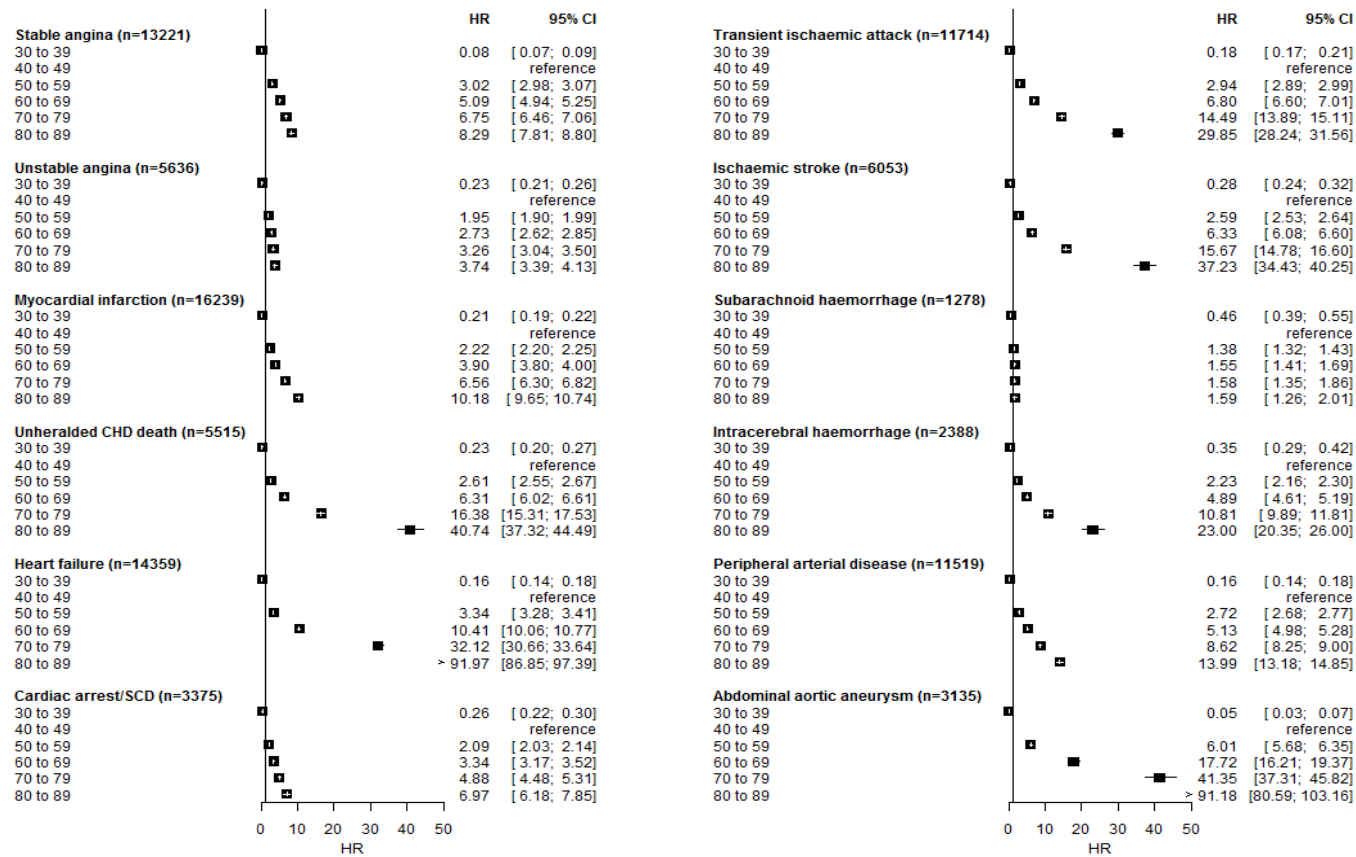
eFigure 5 The C-index of risk predictions based on CVD phenotype-specific models, adjusted for age & sex only

Supplemental References

1. Noble M, McLennan D, Wilkinson K, Whitworth A, Exley S, Barnes H, Dibben C, 2007, *The English Indices of Deprivation*. Wetherby: Department of Communities and Local Government. 2007.
2. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley and Sons. 1987.
3. White IR, Royston P. Imputing missing covariate values for the Cox model. *Stat Med*. 2009;28:1982–1998.



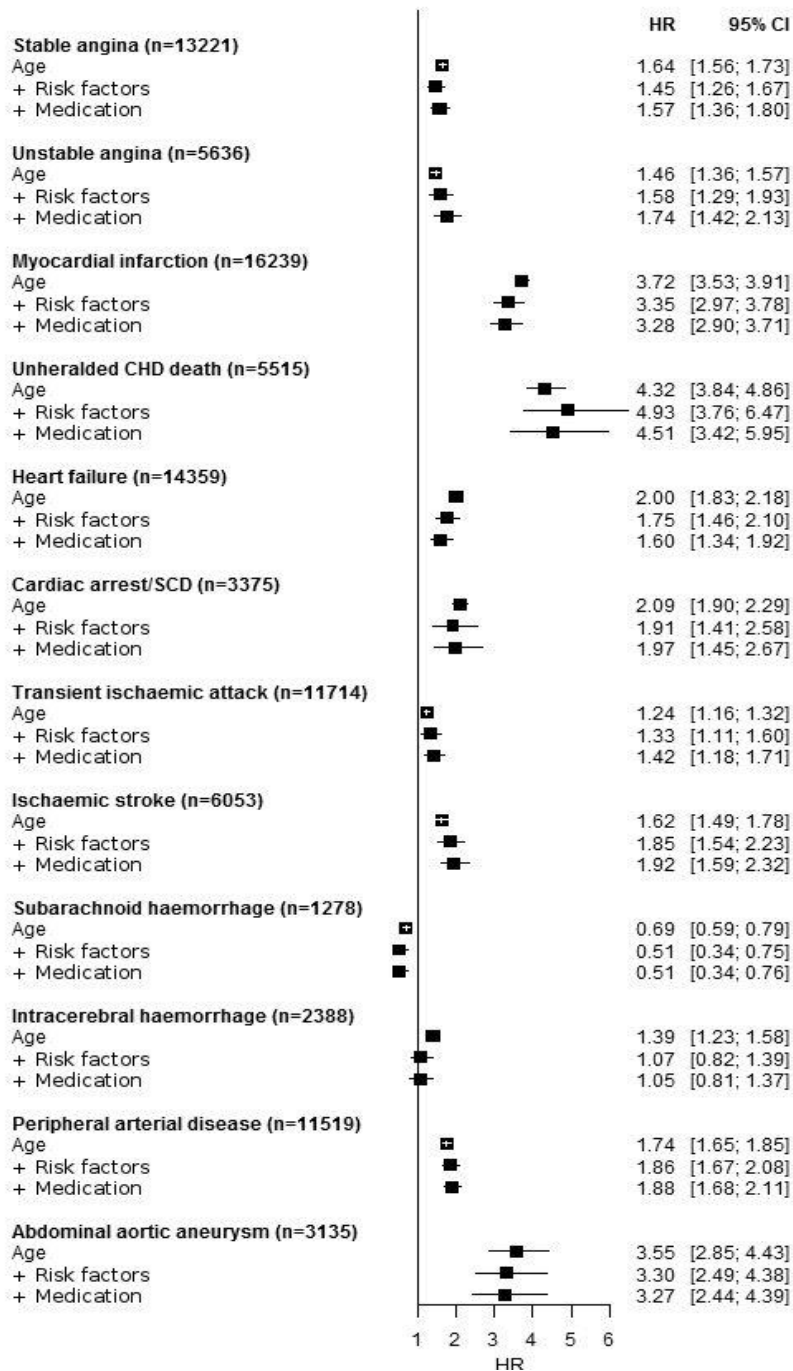
eFigure 1



HRs comparing age groups to reference age group (40-49 years), adjusted for sex and stratified by primary care practice. CHD indicates coronary heart disease; SCD, sudden coronary death.

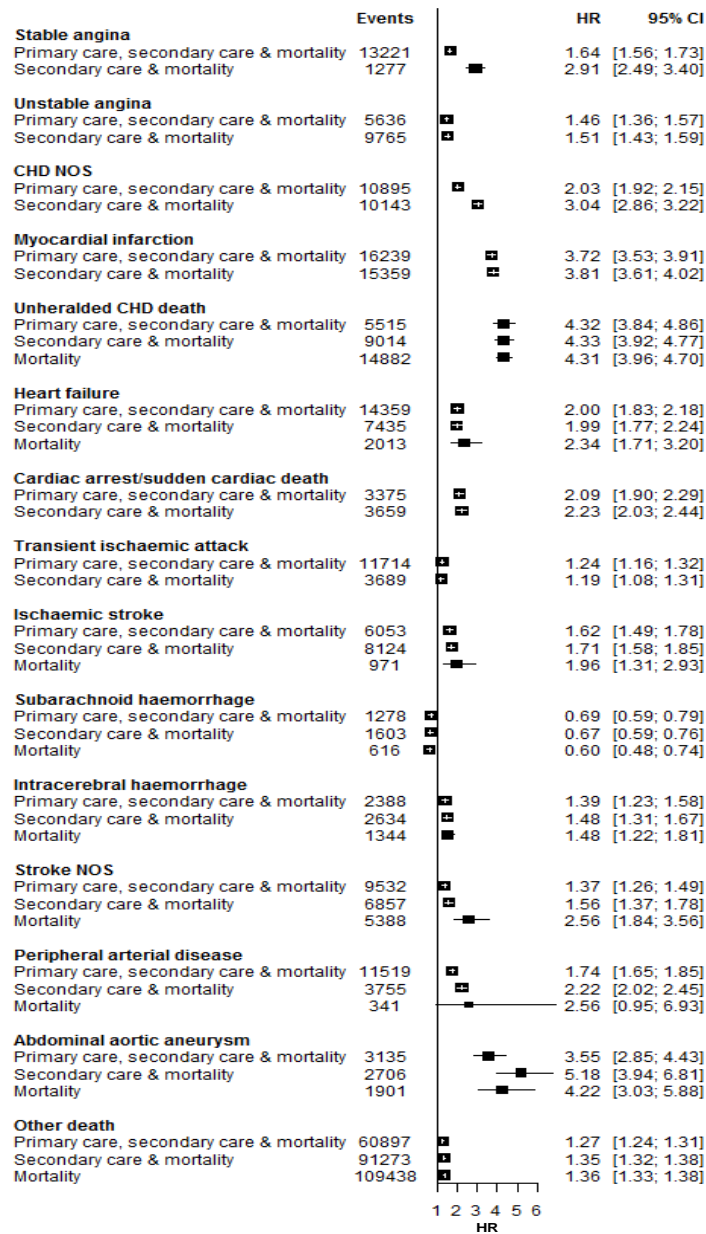
eFigure 2

eFigure 3



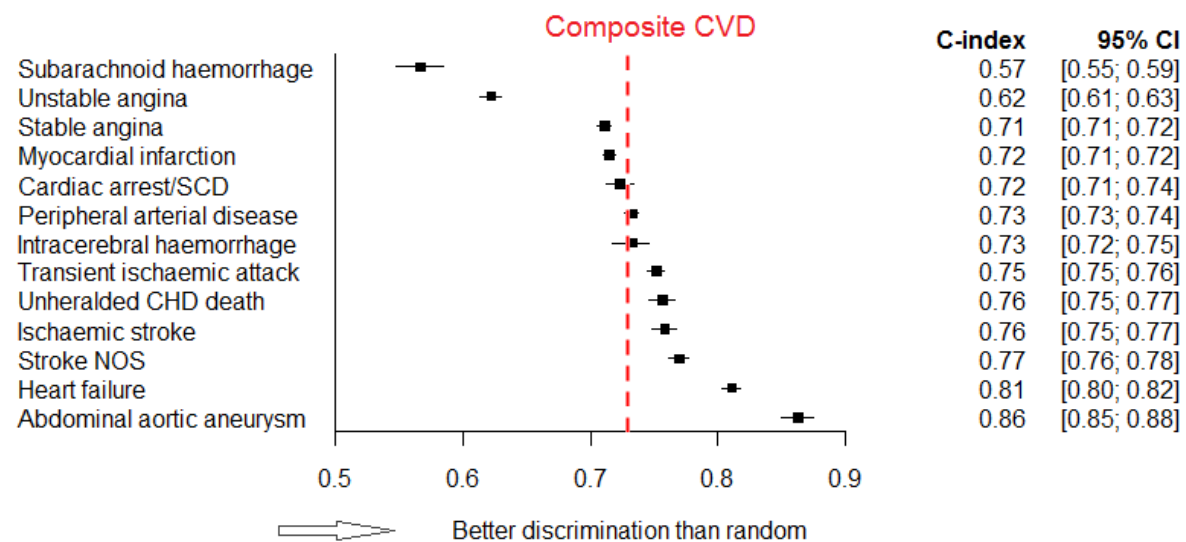
HRs adjusted for age, risk factors and medications, with missing data handled using multiple imputation. Adjustment for risk factors included smoking status, body mass index, systolic blood pressure, total and high-density lipoprotein cholesterol, diabetes mellitus and social economic status. Adjustment for medications included statins, blood pressure medications (includes thiazides, beta-blockers, ARB/ACE inhibitors) and, in women only, oral contraceptives or hormone replacement therapy. CHD indicates coronary heart disease; SCD, sudden cardiac death.

eFigure 4



HRs comparing men to women by source of endpoint data, adjusted for age and stratified by primary care practice. CHD indicates coronary heart disease; NOS, not otherwise specified.

eFigure 5



Data restricted to baseline age 40 to 74 (N= 1,004,190). Composite CVD (red dashed line), C-index 0.73 (95% CI, 0.72, 0.73). SCD indicates sudden cardiac death; CHD, coronary heart disease, NOS, not otherwise specified

How Does Cardiovascular Disease First Present in Women and Men?: Incidence of 12 Cardiovascular Diseases in a Contemporary Cohort of 1 937 360 People

Julie George, Eleni Rapsomaniki, Mar Pujades-Rodriguez, Anoop Dinesh Shah, Spiros Denaxas, Emily Herrett, Liam Smeeth, Adam Timmis and Harry Hemingway

Circulation. 2015;132:1320-1328; originally published online September 1, 2015;
doi: 10.1161/CIRCULATIONAHA.114.013797

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2015 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the
World Wide Web at:

<http://circ.ahajournals.org/content/132/14/1320>

Free via Open Access

Data Supplement (unedited) at:

<http://circ.ahajournals.org/content/suppl/2015/09/01/CIRCULATIONAHA.114.013797.DC1.html>

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Circulation* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the [Permissions and Rights Question and Answer](#) document.

Reprints: Information about reprints can be found online at:
<http://www.lww.com/reprints>

Subscriptions: Information about subscribing to *Circulation* is online at:
<http://circ.ahajournals.org/subscriptions/>