A population-based study of 92 clinically recognised risk factors for heart failure

Banerjee, Amitava; Pasea, Laura; Chung, Sheng-chia; Direk, Kenan; Asselbergs, Folkert; Grobbee, Diederick E; Kotecha, Dipak; Anker, Stefan D; Dyszynski, Tomasz; Tyl, Benoît; Denaxas, Spiros; Lumbers, R Thomas; Hemingway, Harry

DOI:
10.1002/ejhf.2417

Banerjee Amitava (Orcid ID: 0000-0001-8741-3411)
Pasea Laura (Orcid ID: 0000-0001-9374-9691)
Asselbergs Folkert (Orcid ID: 0000-0002-1692-8669)
Kotecha Dipak (Orcid ID: 0000-0002-2570-9812)
Tyl Benoit (Orcid ID: 0000-0001-5297-8412)
Denaxas Spiros (Orcid ID: 0000-0001-9612-7791)
Lumbers R Thomas (Orcid ID: 0000-0002-9077-4741)

A population-based study of 92 clinically recognised risk factors for heart failure: co-occurrence, prognosis and preventive potential.

Amitava Banerjee[1,2,3], associate professor in clinical data science and honorary consultant cardiologist ami.banerjee@ucl.ac.uk*

Laura Pasea[1], post-doctoral research fellow l.pasea@ucl.ac.uk

Sheng-Chia Chung[1], senior research fellow s.chung@ucl.ac.uk

Kenan Direk[1, 4], research data manager k.direk@ucl.ac.uk

Folkert Asselbergs[1,2,5,6], professor of precision medicine and consultant cardiologist f.asselbergs@ucl.ac.uk

Diederick E Grobbee[7], professor of clinical epidemiology d.e.grobbee@umcutrecht.nl

Dipak Kotecha[6,8,9], professor of cardiology & cardiac imaging and honorary consultant cardiologist d.kotecha@bham.ac.uk

Stefan D Anker[10], professor of cardiology and cachexia research s.anker@cachexia.de

Tomasz Dyszynski[11], global safety leader tomasz.dyszynski@bayer.com

Benoît Tyl[12], director of opportunities and translational medicine benoit.tyl@servier.com

Spiros Denaxas[1,5], professor of biomedical informatics s.denaxas@ucl.ac.uk

R Thomas Lumbers[1,2,5], UKRI Rutherford Fellow and honorary consultant cardiologist t.lumbers@ucl.ac.uk

Harry Hemingway[1,5,13]. professor of clinical epidemiology and honorary consultant in public health h.hemingway@ucl.ac.uk

[1]*Institute of Health Informatics, University College London, 222 Euston Road, London, UK*
[2]*University College London Hospitals NHS Trust, 235 Euston Road, London, UK*
[3]*Barts Health NHS Trust, The Royal London Hospital, Whitechapel Rd, London, UK*
[4]*UCL Energy Institute, London, UK*
[5]*Health Data Research UK, London, UK*
[6]*Department of Cardiology, University Medical Center Utrecht, Utrecht, Netherlands*
[7]*Julius Center Research Program Cardiovascular Epidemiology, Utrecht University, Utrecht, Netherlands*
[8]*Institute of Cardiovascular Sciences, University of Birmingham, Birmingham, UK*
[9]*Health Data Research UK Midlands, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK*
[10] *Department of Cardiology, Charité Campus Virchow-Klinikum, Berlin, Germany*
[11]*Bayer AG, Medical Affairs & Pharmacovigilance, Pharmaceuticals TG Cardio, Thrombosis & Hemophilia Building M084, 112 13353 Berlin, Germany*
[12]*Center for Therapeutic Innovation, Cardiovascular and Metabolic Disease, Institut de Recherches Internationales Servier, 50 Rue Carnot, 92284, Suresnes Cedex, France*
[13]*National Institute for Health Research University College London Hospitals Biomedical Research Centre.*

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ejhf.2417

Abstract

Background: Primary prevention strategies for heart failure(HF) have had limited success, possibly due to a wide range of underlying risk factors(RFs). Systematic evaluations of the prognostic burden and preventive potential across this wide range of risk factors are lacking.

Objective: To estimate evidence, prevalence and co-occurrence for primary prevention and impact on prognosis of RFs for incident HF.

Methods: We systematically reviewed trials and observational evidence of primary HF prevention across 92 putative aetiologic RFs for HF identified from US and European clinical practice guidelines. We identified 170 885 individuals aged ≥30 years with incident HF from 1997-2017, using linked primary and secondary care UK electronic health records(EHR) and rule-based phenotypes(ICD-10, Read Version 2, OPCS-4 procedure and medication codes) for each of 92 RFs.

Results: Only 10/92 factors had high quality observational evidence for association with incident HF; 7 had effective RCT-based interventions for HF prevention(RCT-HF), and 6 for CVD prevention, but not HF(RCT-CVD), and the remainder had no RCT-based preventive interventions(RCT-0). We were able to map 91/92 risk factors to EHR using 5961 terms, and 88/91 factors were represented by at least one patient. In the 5 years prior to HF diagnosis, 44.3% had ≥4 RFs. By RCT evidence, the most common RCT-HF RFs were hypertension(48.5%), stable angina(34.9%), unstable angina(16.8%), myocardial infarction(15.8%), and diabetes(15.1%); RCT-CVD RFs were smoking(46.4%) and obesity(29.9%); and RCT-0 RFs were atrial arrhythmias(17.2%), cancer(16.5%),), heavy alcohol intake(14.9%). Mortality at 1 year varied across all 91 factors(lowest: pregnancy-related hormonal disorder 4.2%; highest: phaeochromocytoma 73.7%). Among new HF cases, 28.5% had no RCT-HF RFs and 38.6% had no RCT-CVD RFs. 15.6% had either no RF or only RCT-0 RFs.

Conclusion: 1 in 6 individuals with HF have no recorded RFs or RFs without trials. We provide a systematic map of primary preventive opportunities across a wide range of RFs for HF, demonstrating a high burden of co-occurrence and the need for trials tackling multiple RFs.

Introduction

Declines in incidence of heart failure(HF) have been slower than for ischaemic heart disease(IHD) and stroke[1, 2]. Primary prevention strategies exist for HF in individuals with hypertension, IHD and diabetes mellitus(DM)[3-5], but the European Society of Cardiology(ESC) identifies 89 discrete, frequently overlapping, risk factors(RFs), classified as 'diseased myocardium', 'abnormal loading conditions' and 'arrhythmias'(**Web Table 1**), partly explaining the limited success of HF primary prevention. A further three RFs are mentioned in American Heart Association(AHA) primary cardiovascular disease(CVD) prevention guidelines (smoking, reduced physical activity, PA, and reduced cardiorespiratory fitness)[6]. However, beyond suggesting broad diagnostic work-up, international HF guidelines neglect prevalence, co-occurrence, relative importance and prognosis by these 92 RFs[3].

In order to tackle the high and rising global burden of HF[1, 7-11], primary prevention strategies must prioritise evidence-based RF-specific interventions. The only cause-specific interventions for HF supported by randomised controlled trials(RCT) in primary CVD prevention guidelines are sodium/glucose cotransporter 2(SGLT2) inhibitors for DM, and blood pressure(BP)-lowering therapy for hypertension[6]. Canakinumab, an interleukin-1β inhibitor, may have a role in reducing HF events[12]. Other recommendations for HF prevention, such as increased PA[13], smoking cessation[14], or "ideal cardiovascular health"(smoking, cholesterol, BP, blood glucose, weight, diet and PA)[15-17] are not based on RCT evidence, which needs to be reviewed across the 92 RFs.

Effective, impactful prevention relies on knowledge of prevalence, co-occurrence and preventive potential across 92 RFs. However, studies to-date have assessed individual RFs[18], considering neither RFs comprehensively[9], nor basic HF sub-typing, e.g. with and without antecedent myocardial infarction(MI), hypertension and DM[19-26]. Despite proven validity of electronic health record(EHR) research in HF[27] for detection[28], prognosis[29], risk prediction[30] and burden of disease[1], "agnostic" approaches have not yet been used in national EHR across a wide range of RFs for incident HF, unlike genomics[31].

For each of 92 HF RFs reported in clinical guidelines, our objectives were to:

(i) Classify preventive potential by associated relative risk(RR) from observational studies, and effective interventions from RCTs (for HF: *RCT-HF*; CVD prevention, but unknown impact on HF: *RCT-CVD*; or no preventive intervention: *RCT-0*)

(ii) Develop reproducible coding and conduct a population-based, linked EHR study[32] to investigate: (a)prevalence and co-occurrence; (b)prognosis; and (c)preventable burden by effective treatments specific to HF and CVD prevention.

Methods:

Risk factors

We extracted RFs from guidelines: (i)ESC[8]: 89 RFs for HF (**Web Table 1**) and (ii)AHA[11]: 3 RFs for primary HF prevention: smoking, reduced PA and reduced cardiorespiratory fitness.

Evidence of preventive potential for 92 risk factors for heart failure

Following literature review of observational studies and RCTs, we investigated RFs by (i) *level of evidence*(GRADE A-D)[33] and *strength of association*(relative risk, RR) for incident HF; and (ii)*RF-specific interventions*: for primary prevention of HF(*RCT-HF*), CVD(*RCT-CVD*), or no interventions(*RCT-0*), noting relative risk reduction(RRR). GRADE levels of evidence were high(A: ≥2 high-quality cohort studies with consistent results or in special cases: one large, high-quality multi-centre trial), moderate(B:one high-quality cohort study and several cohort studies with some limitations), low(C:≥1 cohort studies with severe limitations) or very low(D: expert opinion, no direct research evidence, ≥1 studies with very severe limitations).

Electronic health record cohort and study population

We used primary care electronic health records (EHR) in Clinical Practice Research Datalink, (CPRD-GOLD), hospital admissions (Hospital Episodes Statistics, HES) and death registry (Office for National Statistics, ONS), with prospective recording and follow-up, linked by CPRD and NHS Digital using a unique national healthcare identifier[32]. MHRA (UK) Independent Scientific Advisory Committee [18_029R] approval was under Section 251 (NHS Social Care Act 2006). Eligible individuals were ≥30 years and free from HF at baseline. Patients with diagnosis of incident HF between 1 January 1997 and 1 January 2017, and ≥5 years of medical history available before HF diagnosis were included. Follow-up ceased at the date of death or on 1 January 2017. Incident HF was defined as the first coding of diagnosis after baseline (study entry) of fatal or non-fatal, hospitalised or non-hospitalised HF, identified in primary care (Read clinical terminology systems) and hospital inpatient admission (International Statistical Classification of Diseases-10th version; ICD-10) using a validated CALIBER phenotype[28, 32], involving ICD-10 I50, I110, I130, I132, I260 codes and Read code equivalents.

EHR phenotypes for 92 risk factors (14 groups) for heart failure

For each of the 92 RFs, phenotyping algorithms (code lists plus logic of how the codes are combined) are available at www.caliberresearch.org/portal (**Supplementary Code List**). Where available (n=66) we used existing EHR phenotyping algorithms. Hypertension was based on recorded values in primary care according to recent guidelines: ≥140 mmHg systolic BP (or ≥150 mmHg for people aged ≥60 years without DM and chronic kidney disease) and/or ≥90 mmHg diastolic BP(34). DM was defined at baseline (including type: 1, 2, or uncertain) by coded diagnoses recorded in CPRD or HES at or before study entry[35]. Heavy alcohol intake was defined by most recent record of alcohol consumption in the five years before study entry[36]. ESC guidelines list 5 different IHD subtypes, not directly available in EHR. Based on clinical judgment of 2 cardiologists(AB and TL), we used available EHR data("ESC" term) as follows: abnormal coronary microcirculation ("coronary artery aneurysm"), endothelial dysfunction("vasospastic angina"), unstable angina, UA,("myocardial stunning"), stable angina, SA,("epicardial coronary disease") and MI ("myocardial scar"). We developed 36 new phenotypes based on available data and by clinical judgment (AB and TL), using the CALIBER approach(**Table 1**)[32], a collaborative, iterative process involving multiple disciplines (e.g. clinicians, epidemiologists, computer scientists, public health researchers, statisticians), using Read codes (Version 2), ICD-10 coding, drugs and procedure (OPCS-4) codes. AB and TL independently agreed all EHR RF definitions and a third reviewer (HH) resolved cases of disagreement.

Follow-up

Participants who developed new-onset RFs during follow-up were analysed according to the baseline status of that RF. We considered RFs as ever (in the 5 years prior to first HF diagnosis), first ever (first RF recorded in the 5 years prior to HF diagnosis) or most recent (last RF recorded prior to or at HF diagnosis). RFs were curated as individual binary variables. Primary endpoint was 1-year all-cause mortality, defined by the record in either ONS or CPRD.

Analysis

For each of 92 RFs for incident HF, we calculated observed frequency for each RF ever in the 5 years prior to HF diagnosis. RFs were not mutually exclusive in the initial analysis, i.e. an individual patient could have multiple RFs. These analyses were repeated by first ever and most recent RFs. For the ten most prevalent RFs and the 14 RF groups(IHD; toxic damage; immune-mediated and inflammatory damage; infiltration; metabolic derangements; genetic abnormalities; hypertension; valve and myocardium structural defects; pericardial and endomyocardial pathologies; high output states; volume overload; tachyarrhythmias; bradyarrhythmias; primary prevention) "ever" in the 5 years prior to HF diagnosis, baseline characteristics were compared. The 92 "ever" RFs were analysed by age at HF diagnosis. The frequency of individuals was analysed by number of risk factors. We compared the

observed age-and-sex adjusted and case-mix-adjusted 1-year mortality by the 12 most prevalent RFs and the 14 RF groups for HF with Kaplan-Meier estimates and Cox proportional hazards models, adjusted for age and gender. The proportional hazard assumption and model fit was examined by Schoenfeld residuals and c-index. All analyses were performed with SAS (version 9.3) and R (version 3.4.3).

Results

Review of observational evidence and RCTs

Level of evidence was A for 10/92 RFs (B: n=24 and C: n=58). Associations with incident HF were *very strong*(RR >3.5; n=4: MI, HCM, pregnancy(pre-eclampsia), and atrial arrhythmias(atrial fibrillation)); *strong*(RR=2.5-3.5; n=5: hypertension, smoking, reduced cardiorespiratory fitness, connective tissue diseases and sinus node dysfunction); *moderate*(RR=1.5-2.5; n=15: SA, DM, reduced PA, Conn's syndrome, phaeochromocytoma, obesity, acquired valve disease, arteriovenous fistula, severe anaemia, thyrotoxicosis, renal failure and conduction disorders); and *weak*(RR<1.5; n=4: UA, alcohol, metabolic syndrome and parathyroid disorders). The remaining 64/92 RFs (including thyroid disease:9.1%, iron deficiency:6.1% and cytostatic drugs: 4.1%) lacked available evidence for strength of association with incident HF(**Table 1**). Only 7/92 RFs were *RCT-HF:* UA, SA, MI, hypertension, cytostatic drugs, DM and renal failure. Six RFs (smoking, reduced PA, obesity, aortic valve disorders, reduced cardiorespiratory fitness and amyloidosis) were *RCT-CVD*.

Study population, prevalence and co-occurrence of RFs

Using 5961 controlled clinical terminology terms, we developed phenotypes for 91/92 RFs(no codes available for cardiorespiratory fitness), including 170885 individuals with incident HF(**Web Figure 1, Web Table 2**). Mean age at HF diagnosis was 73.7(s.d.14.3) years.

Hypertension(48.5%), smoking(46.4%), SA(34.9%), obesity(29.9%), atrial arrhythmias(17.2%), UA(16.8%), cancer(16.5%), MI(15.8%), DM(15.1%), alcohol(14.9%), severe anaemia(14.3%) and thyroid disorders(9.1%) were commonest. Prevalence was <1% for 63/91 RFs and zero for 3 RFs(endomyocardial fibrosis, immunomodulating drugs and Chagas disease)(**Figure 1**, **Table 2**). 8.0% of those with incident HF had 0/91 RFs. IHD, atrial arrhythmias, hypertension, obesity, DM and cancer had >15% prevalence, among 12 commonest RFs. Bradyarrhythmias, toxic damage, genetic abnormalities and IHD were more common in males than females, unlike high output states and immune-mediated/inflammatory which were more common in females(**Web Table 3)**.

When RFs were analysed by age at HF diagnosis, individuals with atrial arrhythmias were oldest (mean age 80.1, s.d. 10 yrs) and with none of the 91 RFs were youngest (mean age

67.1, s.d. 17.1 yrs). Analysing "first ever" RFs in the 5 years preceding HF diagnosis, the commonest were hypertension, smoking, SA, obesity, other cause (no history of any of the 91 RFs), heavy alcohol intake, cancer, DM, severe anaemia, atrial arrhythmias and MI. Analysing "most recent" RFs, the commonest were smoking, hypertension, other cause, SA, atrial arrhythmias, obesity, UA, MI, cancer, severe anaemia and heavy alcohol intake(**Web Figures 2 and 3**). Among the 4 commonest RFs overall, for hypertension, SA and obesity, prevalence of CVD and RFs was higher in "first ever" than "last ever" classification, whereas for atrial arrhythmias, the opposite trend was true(**Web Table 4**).

8.0%, 14.3%, 17.2%, 16.2% and 44.3% of individuals with HF had 0, 1, 2, 3 and ≥4 RFs, respectively. Prevalence of ≥4 RFs increased with age at HF onset(1.2%, 3.0%, 5.8%, 12.9% and 20.5% for <50, 50-59, 60-69, 70-79, and ≥80 years)(**Web Figure 4**). Hypertension, SA and obesity were most commonly associated with other RFs. Almost all(n=85) RFs were co-morbid with hypertension. For those with a RF, probability of hypertension was 53.3%(average over 85 RFs). Commonest combinations of 2, 3, 4 and 5 RFs were hypertension and smoking; hypertension, obesity and smoking; hypertension, SA, MI and smoking; and hypertension, smoking, SA, UA, and MI. For the 12 most prevalent RFs, the proportion with 0 and ≥4 RFs in addition to the named RF was 6.8% and 43.4% for hypertension, 6.5% and 46.9% for smoking. 3.6% and 57.1% for SA, 3.9% and 52.1% for obesity, 4.7% and 53.9% for atrial arrhythmias, 0.7% and 72.0% for UA, 4.5% and 54.0% for cancer, 1.0% and 65.1% for MI, 1.7% and 66.7% for DM, and 3.8% and 55.7% for heavy alcohol intake, 4.7% and 56.8% for severe anaemia and 3.4% and 57.3% for thyroid disorders. For the same RFs, in those *without* the named RF, the proportion of individuals with 0 and ≥4 RFs was 15.5% and 28.9% for hypertension, 14.3% and 27.6% for smoking, 12.3% and 28.7% for SA, 11.4% and 33.7% for obesity, 9.7% and 38.8% for atrial arrhythmias, 9.6% and 35.9% for UA, 9.6% and 39.1% for cancer, 9.5% and 37.5% for MI, 9.4% and 37.5% for DM, and 9.4% and 39.4% for heavy alcohol intake, 9.3% and 39.7% for severe anaemia and 8.8% and 41.4% for thyroid disorders.

Prognosis

One-year mortality was 16.7%, increasing with number of RFs(8.5%, 10.2%, 12.8%, 16.2% and 23.1% for 0, 1, 2, 3 and ≥4 RFs respectively). For individual RFs, 1- and 5-year mortality were highest for phaeochromocytoma(73.7% and 79.0%) and lowest for pregnancy-related hormonal disorder(7.6% and 15.4%)(**Figure 2**). Among the commonest RFs, cancer(55.0%), atrial arrhythmias(53.1%) and severe anaemia(52.3%) had worst 5-year prognosis(**Figure 3**).

Preventable burden

Among hypertensive individuals, only 51.7% were on angiotensin converting enzyme inhibitors(ACEI) and 53.7% on calcium antagonists. Among those with SA, 73.5% and 63.1% were on antiplatelets and statins respectively**(Table 1)**. Individuals with 0/91 RFs were

younger and less likely to be on medications at HF diagnosis. Of the commonest RFs, 5/12 were RCT-HF. Of those with ≥1 RF, most had ≥1 RCT-HF or RCT-CVD(**Web Figure 5, Table 1**). Of all new HF cases, 28.5% had no RCT-HF RFs and 38.6% had no RCT-CVD RFs. 15.6% had either no risk factor, or a risk factor without evidence of preventive potential. Individuals >80 years with 1 or 2 RFs in the 5 years prior to HF diagnosis were less likely to have ≥1 treatable RF than individuals aged <65 or 65-75 years(**Figure 4**).

Discussion

We provide the first systematic map of primary prevention opportunities across a wide range of RFs for heart failure, with four main findings. First, we show poor quality evidence for RCT-supported interventions to prevent HF across 92 RFs. Second, we rank order the prevalence of RFs recorded prior to the first diagnosis of HF (and therefore amenable to primary preventive efforts), of which hypertension, smoking, obesity, atrial arrhythmias, myocardial infarction, DM and heavy alcohol intake are noteworthy. Third, one- and 5-year mortality for HF was highly variable, depending on specific causes (e.g. ischaemic vs non-ischaemic) and the number of co-occurring RFs. Fourth, the majority of individuals with HF (84.4%), had at least one RF amenable to preventive treatment in the five years preceding diagnosis.

Trials to support preventive interventions are lacking (i.e. of 92 RFs for HF, only 7 were directly supported by RCT data). Moreover, the level of observational evidence (by GRADE criteria) is poor (i.e. of 92 RFs, levels A=10, B=24, C=58), and 64/92 RFs had no available data for strength of association with incident HF. Lack of evidence limits coordinated approaches to HF prevention at individual and population levels, across research, guidelines and practice.

We provide reusable EHR definitions of each of the HF RFs (https://www.caliberresearch.org/portal). Definitions and coding have varied across different study designs(e.g. trial, cohort, EHR, registry) and settings(e.g. community, primary care, hospital), and may not be representative of the population, hampering the transferability and interoperability of definitions. Standardisation of these definitions may form the basis of new classifications and subphenotypes, “discovered” by machine learning and other methods. A small number of RFs (n=12) may explain 81% of “first” or 65% of “most recent” HF RFs, providing focus for prevention. However, high burden of co-occurring RFs and complexity of interaction between RFs highlights the need for trials across multiple RFs.

The 14 RF groups and 92 RFs are associated with marked differences in mortality after diagnosis, with implications for early diagnosis, risk stratification, management and clinical prioritization. Number and type of comorbidities are related to mortality as per previous studies[51,52], but neither have all RFs been studied together, nor have they been studied by different levels of classification (ESC in this case), nor over the long-term (20 years)[53-56]. For example, in our study, individuals with abnormal loading had worse outcomes than those with

arrhythmias and diseased myocardium, and those with IHD had worse outcomes than hypertension. Our observations may inform future studies of long-term HF pathophysiology by RF clustering[57]. One-year mortality rates are comparable to acute HF, but higher than rates for chronic HF[53], probably reflecting the mixed acute and chronic HF study population.

44.3% of those with HF had ≥4 RFs in the prior 5 years, suggesting major preventive potential. Of all new HF cases, 71.5% had ≥1 of the 7 RCT-HF RFs; 12.9% had ≥1 RCT-CVD RF. By the leading 12 RFs, or by the 14 RF categories, 78-100% of individuals had ≥1 RCT-HF RF, and 65-100% had ≥1 RCT-CVD RF. Most incident HF occurs in presence of hypertension, DM and IHD, highlighting need for primordial prevention. In those without the leading 12 RFs, only 5% had ≥1 RCT-HF RF, 18.1% ≥1 RCT-CVD RF and 84.1% had ≥1 RCT-0 RF.

Strengths and limitations

The key strength of this analysis is to provide a systematic map: RFs for HF have often been studied in isolation[44,45], restricted populations[46,47], or specific subpopulations[48]. Associations between RFs, incidence[22,49] and prognosis[50] (including adjustment for comorbidities[47]) have been investigated, but not across all possible causal RFs. We used national, representative, linked EHR and the most comprehensive list of causes for HF, maximising the external validity of our findings. Incident cases of HF were considered to study causal RFs, and our inclusion criteria enabled the investigation of RFs over a 5-year period prior to diagnosis.

There are inherent limitations. First, there is no ICD-10 code distinguishing “systolic vs diastolic”, “acute vs chronic”, “HF with reduced ejection fraction vs HF with preserved ejection fraction”, and more recent introduction of a new category of “HF with mid-range ejection fraction”[29] (terms to denote these distinctions do however exist in ICD-9-CM and ICD-10-CM which are not used in the UK healthcare system). Furthermore, we lacked echocardiographic data as these events rarely get recorded in structured EHR using ontologies and unstructured data (e.g. clinical text and narrative) as not available for research). Second, the validity of the 91 RF phenotypes, while well-established for some (e.g. hypertension, diabetes, obesity, smoking, heavy alcohol), is not known for the new phenotypes. Coding validity is through the use of comprehensive coding lists across linked EHR data, with review by two cardiologists, and prognosis lends some validity. Third, RFs were analysed by “ever”, “first ever” and “last ever” but neither every permutation and combination nor duration of RFs could be investigated. Therefore, we concentrated on the most common RFs for secondary analyses.

Research implications

First, our findings outline the need for RCTs that examine single and multiple RFs in HF prevention to establish causal inference, and methods such as trial emulation, may have a role where RCTs are unlikely. Second, machine learning may inform distribution and

trajectories of HF by different RF combinations, as well as the impact of longitudinal changes in RFs over time. Third, EHR approaches can be used to define HF subtypes and inform genome-wide approaches, which have led to novel biologic[39] but not translational[40] insights for prevention, to-date. Fourth, prevention strategies may require modification, based on varying prevalence of HF RFs[3], and primary versus secondary prevention. Fifth, novel HF prediction models should account for the interplay of the number and type of RFs, where existing risk prediction models for incident HF have only modest discrimination, partly due to lack of external validation, but also incomplete knowledge of HF causes and classification[46,58].

Clinical implications

Our results have three clinical implications. First, clinician recording and use of better data in EHR is central to understanding and improving HF prevention. Second, in individuals with new and existing HF, RFs by RCT-HF (hypertension, DM and IHD) and RCT-CVD (e.g. smoking, obesity) should be excluded through history, examination and/or investigation and monitored at follow-up, so that evidence-based preventive interventions can be initiated and optimised. Third, HF exemplifies co-occurrence of RFs and multi-morbidity. There are joint clinical guidelines for DM and CVD but more "joined-up" and "cross-disease" thinking is required to emphasise and uptitrate existing treatments in the highest-risk individuals.

Conclusion

In the first systematic and comprehensive map of 92 risk factors for HF, showing that 44.3% of individuals with HF had ≥4 factors recorded by the time of diagnosis, and only 8.0% had no coded risk factor. EHR can be used to study the whole spectrum of causes of HF and should be used to inform future strategies for primary prevention research, diagnostic work-up of individuals with HF as well as treatment of those at highest risk of HF.

Conflict of interests All authors declare no conflict of interest.

DK reports personal fees from Bayer, AtriCure, Amomed, Protherics Medicines Development and Myokardia; all outside the current study.

Acknowledgements

AB is supported by research funding from NIHR (NIHR200937), British Medical Association (TP Gunton award), Astra-Zeneca and UK Research and Innovation. DK is supported by grants from the National Institute for Health Research (NIHR CDF-2015-08-074 RATE-AF; NIHR HTA-130280 DaRe2THINK; NIHR EME-132974 DaRe2THINK-NeuroVascular), the British Heart Foundation (PG/17/55/33087 and AA/18/2/34218), the European Society of Cardiology supported by educational grants from Boehringer Ingelheim/BMS-Pfizer Alliance/Bayer/Daiichi Sankyo/Boston Scientific, the NIHR/University of Oxford Biomedical

Research Centre and British Heart Foundation/University of Birmingham Accelerator Award (STEEER-AF NCT04396418); and Amomed Pharma, IRCCS San Raffaele and Menarini (Beta-blockers in Heart Failure Collaborative Group NCT0083244). HH is an National Institute for Health Research (NIHR) Senior Investigator and funded by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. HH's work is supported by: 1. Health Data Research UK (grant No. LOND1), which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and Wellcome Trust. 2. AB, LP, FA, DEG, DK, SDA, TD, BT, SD, RTL and HH are part of the BigData@Heart Consortium, funded by the Innovative Medicines Initiative-2 Joint Undertaking under grant agreement No. 116074. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA; it is chaired, by DEG and SDA, partnering with 20 academic and industry partners and ESC.

References

1. Conrad N, Judge A, Tran J, Mohseni H, Hedgecott D, Crespillo AP et al. Temporal trends and patterns in heart failure incidence: a population-based study of 4 million individuals. Lancet. 2018 Feb 10;391(10120):572-580.
2. Bhatnagar P, Wickramasinghe K, Wilkins E, Townsend N. Trends in the epidemiology of cardiovascular disease in the UK. Heart. 2016 Dec 15;102(24):1945-1952.
3. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J. 2016;37(27):2129-200.
4. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Jr., Colvin MM, et al. 2017 ACC/AHA/HFSA Focused Update of the 2013 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. J Am Coll Cardiol. 2017;70(6):776-803.
5. Ezekowitz JA, O'Meara E, McDonald MA, Abrams H, Chan M, Ducharme A, et al. 2017 Comprehensive Update of the Canadian Cardiovascular Society Guidelines for the Management of Heart Failure. Can J Cardiol. 2017;33(11):1342-433.
6. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, 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. Circulation. 2019;140(11):e563-e95.
7. Moran AE, Forouzanfar MH, Roth GA, Mensah GA, Ezzati M, Flaxman A, et al. The global burden of ischemic heart disease in 1990 and 2010: the Global Burden of Disease 2010 study. Circulation. 2014;129(14):1493-501.
8. Cook C, Cole G, Asaria P, Jabbour R, Francis DP. The annual global economic burden of heart failure. Int J Cardiol. 2014;171(3):368-76.
9. Khatibzadeh S, Farzadfar F, Oliver J, Ezzati M, Moran A. Worldwide risk factors for heart failure: a systematic review and pooled analysis. Int J Cardiol. 2013;168(2):1186-94.
10. Bragazzi NL, Zhong W, Shu J, Abu Much A, Lotan D, Grupper A, Younis A, Dai H. Burden of heart failure and underlying causes in 195 countries and territories from 1990 to 2017. Eur J Prev Cardiol. 2021 Feb 12:zwaa147.
11. Savarese G, Lund LH. Global Public Health Burden of Heart Failure. Card Fail Rev. 2017 Apr;3(1):7-11.
12. Everett BM, Cornel JH, Lainscak M, Anker SD, Abbate A, Thuren T, et al. Anti-Inflammatory Therapy With Canakinumab for the Prevention of Hospitalization for Heart Failure. Circulation. 2019;139(10):1289-99.
13. Colpani V, Baena CP, Jaspers L, van Dijk GM, Farajzadegan Z, Dhana K, et al. Lifestyle factors, cardiovascular disease and all-cause mortality in middle-aged and elderly women: a systematic review and meta-analysis. Eur J Epidemiol. 2018;33(9):831-45.
14. Kamimura D, Cain LR, Mentz RJ, White WB, Blaha MJ, DeFilippis AP, et al. Cigarette Smoking and Incident Heart Failure: Insights From the Jackson Heart Study. Circulation. 2018;137(24):2572-82.
15. Spahillari A, Talegawkar S, Correa A, Carr JJ, Terry JG, Lima J, et al. Ideal Cardiovascular Health, Cardiovascular Remodeling, and Heart Failure in Blacks: The Jackson Heart Study. Circ Heart Fail. 2017;10(2).

16. Folsom AR, Shah AM, Lutsey PL, Roetker NS, Alonso A, Avery CL, et al. American Heart Association's Life's Simple 7: Avoiding Heart Failure and Preserving Cardiac Structure and Function. Am J Med. 2015;128(9):970-6 e2.
17. Folsom AR, Yamagishi K, Hozawa A, Chambless LE, Atherosclerosis Risk in Communities Study I. Absolute and attributable risks of heart failure incidence in relation to optimal risk factors. Circ Heart Fail. 2009;2(1):11-7.

18. Butler J. Primary prevention of heart failure. ISRN Cardiol. 2012;2012:982417.
19. Bleumink GS, Knetsch AM, Sturkenboom MC, Straus SM, Hofman A, Deckers JW, et al. Quantifying the heart failure epidemic: prevalence, incidence rate, lifetime risk and prognosis of heart failure The Rotterdam Study. Eur Heart J. 2004;25(18):1614-9.
20. Ahmad FS, Ning H, Rich JD, Yancy CW, Lloyd-Jones DM, Wilkins JT. Hypertension, Obesity, Diabetes, and Heart Failure-Free Survival: The Cardiovascular Disease Lifetime Risk Pooling Project. JACC Heart Fail. 2016;4(12):911-9.
21. Khan SS, Ning H, Wilkins JT, Allen N, Carnethon M, Berry JD, et al. Association of Body Mass Index With Lifetime Risk of Cardiovascular Disease and Compression of Morbidity. JAMA Cardiol. 2018;3(4):280-7.
22. Leening MJ, Ferket BS, Steyerberg EW, Kavousi M, Deckers JW, Nieboer D, et al. Sex differences in lifetime risk and first manifestation of cardiovascular disease: prospective population based cohort study. BMJ. 2014;349:g5992.
23. Dunlay SM, Weston SA, Jacobsen SJ, Roger VL. Risk factors for heart failure: a population-based case-control study. Am J Med. 2009;122(11):1023-8.
24. Kalogeropoulos A, Georgiopoulou V, Kritchevsky SB, Psaty BM, Smith NL, Newman AB, et al. Epidemiology of incident heart failure in a contemporary elderly cohort: the health, aging, and body composition study. Arch Intern Med. 2009;169(7):708-15.
25. Clark D, 3rd, Colantonio LD, Min YI, Hall ME, Zhao H, Mentz RJ, et al. Population-Attributable Risk for Cardiovascular Disease Associated With Hypertension in Black Adults. JAMA Cardiol. 2019:1-9.
26. Chatterjee NA, Chae CU, Kim E, Moorthy MV, Conen D, Sandhu RK, et al. Modifiable Risk Factors for Incident Heart Failure in Atrial Fibrillation. JACC Heart Fail. 2017;5(8):552-60.
27. Tison GH, Chamberlain AM, Pletcher MJ, Dunlay SM, Weston SA, Killian JM, et al. Identifying heart failure using EMR-based algorithms. Int J Med Inform. 2018;120:1-7
28. Ng K, Steinhubl SR, deFilippi C, Dey S, Stewart WF. Early Detection of Heart Failure Using Electronic Health Records: Practical Implications for Time before Diagnosis, Data Diversity, Data Quantity and Data Density. Circulation Cardiovascular quality and outcomes. 2016;9(6):649-58.
29. Koudstaal S, Pujades-Rodriguez M, Denaxas S, Gho J, Shah AD, Yu N, et al. Prognostic burden of heart failure recorded in primary care, acute hospital admissions, or both: a population-based linked electronic health record cohort study in 2.1 million people. Eur J Heart Fail. 2017;19(9):1119-27.
30. Lagu T, Pekow PS, Stefan MS, Shieh MS, Pack QR, Kashef MA, et al. Derivation and Validation of an In-Hospital Mortality Prediction Model Suitable for Profiling Hospital Performance in Heart Failure. J Am Heart Assoc. 2018;7(4).

*N.B. References 31-139 are in "**Supplemental References**" in **Web Supplementary Material***

Tables and Figures

Table 1: ESC/AHA risk factors for heart failure: evidence from observational studies and randomized trials, and prevalence in electronic health records. *Factors are ordered by prevalence (high to low) in the population.*

a) Evidence that treating the risk factors reduces risk of incident heart failure (RCT-HF)

Risk factor	Observational Level of evidence[Ref] according to GRADE Strength of association RR(95% CI)	Randomised controlled trial Treatments (incident HF as outcome) RRR(95% CI)	Diseased myocardium	Abnormal loading conditions	Arrhythmias	AHA prevention guidelines	Prevalence N (%)	Number of EHR codes (n)
Hypertension	A[37] 1.61(1.33-1.96)	Antihypertensive 0·72(0·67-0·78)[38]		●			82921 (48.7)	91
Stable angina	A[37] 2.90(1.85-4.54)	Statins 0.91(0.84-0.98)[39] ACEI 0·77(0·67–0·90)[40] Tight BP control 0.76(0.67-0.86)[41]	●				59689 (35.1)	71
Unstable angina	A[42] 1.35(1.02-1.78)	Tight BP control 0.76(0.67-0.86)[41] Clopidogrel 0.82(0.69-0.98)[43] ACEI 0.85(0.78-0.92)[44]	●				28700 (16.9)	16
Myocardial infarction	A[45] 3.80(2.10-6.80)	Clopidogrel 0.82(0.69-0.98)[43] ACEI 0.85(0.78-0.92)[44]	●				26994 (15.9)	74
Diabetes mellitus	A[37] 1.94(1.71-2.19))	ACEI 0.80(0.66-0.96)[46] ARB 0·59 (0·38–0·92)[47] SLGLT2 inhibitors 0·77(0·71-0·84)[48] Tight BP control 0.44 (0.20-0.94)[49]	●				25841 (15.2)	225
Cytostatic drugs	B[50]	Dexrazoxane 0.35(0.27-0.45)[51] Statin 0.31(0.13-0.77)[51] ACEI/ARB 0.11(0.04-0.29)[51] BB 0.31(0.16-0.63)[51]	●				7028 (4.1)	50
Renal failure	B[52] 1.94(1.49-2.53)	ARB 0.67(0.47-0.93)[53]		●			556 (0.33)	44

(b) Evidence that treating the condition reduces risk of cardiovascular disease/mortality or non-RCT evidence for heart failure risk reduction (RCT-CVD)

Risk factor	Observational Level of evidence[Ref]; Strength of association RR(95% CI)	Randomised controlled trial Treatments (incident CVD as outcome) RRR(95% CI)	Diseased myocardium	Abnormal loading conditions	Arrhythmias	AHA prevention guidelines	Prevalence N (%)	Number of EHR codes (n)
Smoking	B[14] 2.82(1.71–4.64)	Smoking cessation 0.72(0.57-0.90)[54]				●	79308 (46.6)	2
Obesity	B[56] 2.12(1.51-2.97)	Bariatric surgery 0.54(0.36-0.82)[57, 58]	●				51068 (30.0)	2
Reduced physical activity	A[55] 1.42(1.37-1.49)	High physical activity 0.74(0.67-0.80)[13]				●	10140 (5.9)	1

Aortic valve disorders	B[37]	1.74(1.07-2.84)	Transcatheter aortic valve implantation 0.55(0.40 to 0.74)[59, 60]		●			5516 (3.2)	70
Amyloidosis	A[63]		Tafamidis 0.70(0.51-0.96)[64]	●				65 (0.04)	23
Reduced cardiorespiratory fitness	B[61]	2.70(2.50-3.57)	High fitness 0.79[0.75-0.83][62]				●	-	0

(c) No evidence of treatment to reduce heart failure risk (RCT-0)

Risk factor	Observational Level of evidence[Ref]; Strength of association RR(95% CI)		Randomised controlled trial Treatments RRR(95% CI)	Diseased myocardium	Abnormal loading conditions	Arrhythmias	AHA prevention guidelines	Prevalence N (%)	Number of EHR codes (n)
Atrial arrhythmias	A[65]	4.62(3.13-6.83)	-			●		29399 (17.3)	27
Cancer	B[66, 67]	1.94(1.66-2.25)	-		●			28164 (16.6)	1856
Heavy alcohol intake	A[68]	1.20(1.11-1.33)	-	●				25425 (14.9)	5
Severe anaemia	B[69]	2.24(1.15-4.35)	-		●			24352 (14.3)	208
Thyroid disorders	B[70]		-		●			15473 (9.1)	150
Conduction disorders	B[71]	2.29(1.80–2.92)	-			●		12426 (7.3)	96
Iron deficiency	C[72]		-	●				10148 (6.0)	22
Bacteria	B[73]		-	●				9703 (5.7)	270
Sepsis	C[74]		-		●			7703 (4.5)	58
Connective tissue diseases	C[75]	3.17(2.63- 3.83)	-	●				7486 (4.4)	111
Ventricular arrhythmias	B[76]	1.72(1.24–2.37)	-			●		6333 (3.7)	8
Rheumatoid arthritis	B[77]	1.56(1.46–1.66)	-	●				5737 (3.4)	59
Tricuspid valve disorders	B[37]	1.74(1.07-2.84)	-		●			5618 (3.3)	57
Thyrotoxicosis	A[70]	1.94(1.01-3.72)	-		●			3387 (2.0)	39
Fluid overload	C[78]		-		●			3081 (1.8)	4
Mitral valve disorders	B[37]	1.74(1.07-2.84)	-		●			2552 (1.5)	68
Calcium abnormalities	B[79, 80]		-	●				2524 (1.5)	91
Pericardial effusion	C[81]		-		●			1667 (0.98)	6
Sinus node dysfunction	B[45]	3.40(1.10-10.80)	-			●		1530 (0.90)	17
Radiation	B[82-84]	2.70(1.60-4.80)	-	●				1463 (0.86)	34
Left ventricular non-compaction	C[85]		-	●				1461 (0.86)	4
Dilated cardiomyopathy	B[86]		-	●				1395 (0.82)	3
Giant cell arteritis	C[87, 88]	2.40(0.90–6.00)	-	●				1317 (0.77)	9
Parathyroid disorders	C[89]	1.38(1.09-1.74)	-	●				1277 (0.75)	71
Metabolic syndrome	C[90]	1.37(1.02–1.84)	-	●				1061 (0.62)	138
Pregnancy hormonal conditions	B[91]		-	●				818 (0.48)	43
Paget's disease	C[92]		-	●				758 (0.45)	51
Pregnancy(pre-eclampsia)	A[93]	4.19(2.09-8.38)	-		●			664 (0.39)	196
Rickettsia	C[94]		-	●				637 (0.37)	13
Sarcoidosis	C[95]		-	●				535 (0.31)	21
Antidepressant	C[50]		-	●				513 (0.30)	659
Coronary artery aneurysm	C[96]		-	●				476 (0.28)	11
Non-steroidal anti-inflammatory drugs	B[50]		-	●				459 (0.27)	4
Human immunodeficiency virus/Acquired immunodeficiency syndrome	B[97, 98]	2.80(2.00-3.80)	-	●				456 (0.27)	116

Pulmonary valve disorders	B[37]	1.74(1.07-2.84)	-		●			412 (0.24)	11
Malnutrition	C[99]		-	●				408 (0.24)	68
Hypertrophic cardiomyopathy	B[100]	4.31(3.30-5.62)	-	●				322 (0.19)	5
Anabolic	C[101]		-	●				286 (0.17)	10
Arteriovenous fistula	C[102]	2.24(1.15-4.35)	-		●			228 (0.13)	37
Phosphate disorders	C[79]		-	●				219 (0.13)	5
Lupus erythematosus	C[103]		-	●				217 (0.13)	23
Laminopathy	C[104]		-	●				214 (0.13)	18
Addison's disease	C[105]		-	●				209 (0.12)	5
Iron overload	C[106]		-	●				187(0.11)	21
Growth hormone deficiency	C[107, 108]		-	●				164 (0.1)	18
Arrhythmogenic right ventricular cardiomyopathy	C[109]		-	●				135 (0.08)	2
Hypercortisolaemia	C[110]		-	●				117 (0.07)	16
Anorexia nervosa	C[111]		-	●				115 (0.07)	7
Anabolic steroids	B[50]		-	●				115 (0.07)	23
Phaeochromocytoma	C[112]	1.94(1.01-3.72)	-	●				112 (0.07)	9
Haemochromatosis	C[113]		-	●				105 (0.06)	2
Congenital	C[114, 115]		-		●			84 (0.05)	135
Cocaine	C[116]		-	●				62 (0.04)	62
Muscular dystrophies	C[117]		-	●				61 (0.04)	14
Constrictive pericarditis	C[118]		-		●			56 (0.03)	4
Vasospastic angina	C[119]		-	●				56 (0.03)	6
Acromegaly	C[108]		-	●				46 (0.03)	3
Cushin's syndrome	B[120]	2·05(1·11-3·78)	-	●				41 (0.02)	11
Restrictive cardiomyopathy	C[121, 122]		-	●				31 (0.02)	2
Churg-Strauss	C[123]		-	●				29 (0.02)	2
Amphetamine	C[124]		-	●				25 (0.01)	25
Endocardial fibroelastosis	C[125]		-		●			16 (0.01)	5
Grave's disease	C[126]		-	●				15 (0.01)	49
Lead toxicity	C[127]		-	●				11 (0.01)	28
Antiarrhythmic drugs	B[50]		-	●				8 (0)	63
Copper toxicity	C[128]		-	●				7 (0)	9
Spirochaetes	C[129]		-	●				7 (0)	14
Lysosomal storage disease	C[130]		-	●				6 (0)	6
Thiamine deficiency	C[131]		-	●				5 (0)	12
Glycogen storage disease	C[132]		-	●				3 (0)	3
Selenium deficiency	C[133]		-	●				2 (0)	4
Hypereosinophilic syndrome	C[134]		-		●			2 (0)	6
Protozoa	C[135]		-	●				2 (0)	25
Cobalt toxicity	C[127]		-	●				2 (0)	1
Fungi	C[136]		-	●				1 (0)	7
L-carnitine deficiency	C[137]		-	●				1 (0)	3
Chagas disease	C[138]		-	●				0 (0)	19
Immunomodulating drugs	C[50]		-	●				0 (0)	2
Endomyocardial fibrosis	C[139]		-		●			0 (0)	5

Abbreviations: RR: relative risk; RRR: relative risk reduction

Blank cells: Observational evidence- no estimate for strength of association from literature.

- Randomised evidence-no trial evidence of treatments or interventions to reduce incident heart failure

GRADE level of evidence for observational evidence: A(High)- Several high-quality cohort studies with consistent results or, in special cases: one large, high-quality multi-centre trial; B (Moderate)- One high-quality cohort study or several cohort studies with some limitations; C (Low)- One or more cohort studies with severe limitations; OR D (Very low)- Expert opinion, no direct research evidence or one or more studies with very severe limitations.

Table 2: Co-occurrence of the 12 most prevalent risk factors ever in the 5 years prior to incident heart failure (n=170885 HF cases)

Characteristics at time of heart failure diagnosis	Hypertension	Smoking	Stable angina	Obesity	Atrial arrhythmias	Unstable angina	Cancer	Myocardial infarction	Diabetes	Heavy alcohol intake	Severe anaemia	Thyroid disorders	Other risk factor	0/92 risk factors recorded
N	**82921**	**79308**	**59689**	**51068**	**29399**	**28700**	**28164**	**26994**	**25841**	**25425**	**24352**	**15473**	**4331**	**13661**
RCT evidence for preventative treatment														
RCT-HF	82921 (100)	63529 (80.1)	59689 (100)	42442 (83.1)	23392 (79.6)	28700 (100)	22623 (80.3)	26994 (100)	25841 (100)	20624 (81.1)	18944 (77.8)	12223 (79)	218 (5)	0 (0)
RCT-CVD	59938 (72.3)	79308 (100)	40080 (67.1)	51068 (100)	19341 (65.8)	20739 (72.3)	18608 (66.1)	18694 (69.3)	21697 (84)	19925 (78.4)	15752 (64.7)	10430 (67.4)	785 (18.1)	0 (0)
RCT-O	58408 (70.4)	55671 (70.2)	43465 (72.8)	35687 (69.9)	29399 (100)	22109 (77)	28164 (100)	19269 (71.4)	19479 (75.4)	25425 (100)	24352 (100)	15473 (100)	3678 (84.9)	0 (0)
Demographics														
Age ( )	75.2 (13.1)	73.3 (13.7)	77.1 (11.1)	71.6 (13.4)	80.1 (10)	77.9 (10.9)	80 (10.4)	76.5 (11.3)	75.6 (11.2)	74 (13.6)	78.1 (13.1)	77.9 (12.2)	71.2 (16.8)	67.1 (17.1)
Women	41001 (49.4)	32564 (41.1)	25111 (42.1)	26107 (51.1)	14371 (48.9)	12625 (44)	13758 (48.8)	9542 (35.3)	11608 (44.9)	10175 (40)	15473 (63.5)	11985 (77.5)	2718 (62.8)	6818 (49.9)
Cardiovascular diseases														
S	29809 (35.9)	31366 (39.5)	59689 (100)	19760 (38.7)	12662 (43.1)	25114 (87.5)	10937 (38.8)	23555 (87.3)	12682 (49.1)	10420 (41)	9881 (40.6)	6052 (39.1)	0 (0)	0 (0)
Atrial arrhythmias	15952 (19.2)	14793 (18.7)	12662 (21.2)	9314 (18.2)	29399 (100)	6073 (21.2)	6630 (23.5)	5100 (18.9)	5054 (19.6)	5066 (19.9)	5359 (22)	3646 (23.6)	0 (0)	0 (0)
Unstable angina	15410 (18.6)	16336 (20.6)	25114 (42.1)	10724 (21)	6073 (20.7)	28700 (100)	5649 (20.1)	12072 (44.7)	6827 (26.4)	5458 (21.5)	5348 (22)	3162 (20.4)	0 (0)	0 (0)
Myocardial infarc	13543 (16.3)	15387 (19.4)	23555 (39.5)	8715 (17.1)	5100 (17.3)	12072 (42.1)	4977 (17.7)	26994 (100)	5992 (23.2)	4898 (19.3)	4200 (17.2)	2562 (16.6)	0 (0)	0 (0)
Conduction dis	6703 (8.1)	6472 (8.2)	7300 (12.2)	3900 (7.6)	4403 (15)	4244 (14.8)	2860 (10.2)	3465 (12.8)	2386 (9.2)	2450 (9.6)	2272 (9.3)	1530 (9.9)	416 (9.6)	0 (0)
Cardiovascular risk factors														
Hypertension	82921 (100)	46894 (59.1)	29809 (49.9)	30720 (60.2)	15952 (54.3)	15410 (53.7)	15571 (55.3)	13543 (50.2)	15009 (58.1)	15619 (61.4)	12408 (51)	8519 (55.1)	0 (0)	0 (0)
Smok	46894 (56.6)	79308 (100)	31366 (52.5)	30203 (59.1)	14793 (50.3)	16336 (56.9)	14736 (52.3)	15387 (57)	16118 (62.4)	16624 (65.4)	11591 (47.6)	7562 (48.9)	0 (0)	0 (0)
Obesit	30720 (37)	30203 (38.1)	19760 (33.1)	51068 (100)	9314 (31.7)	10724 (37.4)	8402 (29.8)	8715 (32.3)	15578 (60.3)	9721 (38.2)	7790 (32)	5885 (38)	0 (0)	0 (0)
Cance	15571 (18.8)	14736 (18.6)	10937 (18.3)	8402 (16.5)	6630 (22.6)	5649 (19.7)	28164 (100)	4977 (18.4)	4785 (18.5)	5083 (20)	5588 (22.9)	2899 (18.7)	0 (0)	0 (0)
Diabe s mellitus	15009 (18.1)	16118 (20.3)	12682 (21.2)	15578 (30.5)	5054 (17.2)	6827 (23.8)	4785 (17)	5992 (22.2)	25841 (100)	5033 (19.8)	5307 (21.8)	3026 (19.6)	0 (0)	0 (0)
Heavy alcoh intake	15619 (18.8)	16624 (21)	10420 (17.5)	9721 (19)	5066 (17.2)	5458 (19)	5083 (18)	4898 (18.1)	5033 (19.5)	25425 (100)	3438 (14.1)	2374 (15.3)	0 (0)	0 (0)
Severe anaemia	12408 (15)	11591 (14.6)	9881 (16.6)	7790 (15.3)	5359 (18.2)	5348 (18.6)	5588 (19.8)	4200 (15.6)	5307 (20.5)	3438 (13.5)	24352 (100)	3618 (23.4)	0 (0)	0 (0)
Thyroid disorders	8519 (10.3)	7562 (9.5)	6052 (10.1)	5885 (11.5)	3646 (12.4)	3162 (11)	2899 (10.3)	2562 (9.5)	3026 (11.7)	2374 (9.3)	3618 (14.9)	15473 (100)	0 (0)	0 (0)
Sepsis	4471 (5.4)	4353 (5.5)	3129 (5.2)	3012 (5.9)	1476 (5)	1724 (6)	1918 (6.8)	1467 (5.4)	1942 (7.5)	1371 (5.4)	1654 (6.8)	844 (5.5)	383 (8.8)	0 (0)

Medication														
Antiplatelet	44857 (54.1)	44219 (55.8)	43882 (73.5)	28296 (55.4)	20450 (69.6)	23414 (81.6)	16371 (58.1)	21630 (80.1)	18724 (72.5)	14422 (56.7)	14378 (59)	8922 (57.7)	793 (18.3)	1678 (12.3)
Statin	41279 (49.8)	42231 (53.2)	37653 (63.1)	28771 (56.3)	14442 (49.1)	20858 (72.7)	13212 (46.9)	19022 (70.5)	20049 (77.6)	13961 (54.9)	11466 (47.1)	7677 (49.6)	319 (7.4)	599 (4.4)
Warfarin	15304 (18.5)	14328 (18.1)	13336 (22.3)	9464 (18.5)	20049 (68.2)	6342 (22.1)	6570 (23.3)	5467 (20.3)	5344 (20.7)	4881 (19.2)	5378 (22.1)	3522 (22.8)	267 (6.2)	409 (3)
Beta blocker	38242 (46.1)	36276 (45.7)	33334 (55.8)	25275 (49.5)	17240 (58.6)	18339 (63.9)	13241 (47)	16430 (60.9)	13579 (52.5)	12255 (48.2)	11362 (46.7)	7458 (48.2)	785 (18.1)	1814 (13.3)
CCB	44505 (53.7)	42203 (53.2)	37242 (62.4)	29956 (58.7)	17030 (57.9)	20649 (71.9)	15208 (54)	16920 (62.7)	17608 (68.1)	14253 (56.1)	13581 (55.8)	8508 (55)	822 (19)	1760 (12.9)
ACEI	42843 (51.7)	40964 (51.7)	34891 (58.5)	29980 (58.7)	17644 (60)	17991 (62.7)	14431 (51.2)	17640 (65.3)	19567 (75.7)	13647 (53.7)	13250 (54.4)	8101 (52.4)	766 (17.7)	1673 (12.2)
ARB	14595 (17.6)	13395 (16.9)	10857 (18.2)	10917 (21.4)	5867 (20)	6034 (21)	5086 (18.1)	5112 (18.9)	6891 (26.7)	4764 (18.7)	4808 (19.7)	3143 (20.3)	214 (4.9)	367 (2.7)

CCB-Calcium channel blocker; ACEI-Angiotensin-converting enzyme inhibitor; Angiotensin receptor blockers-ARB.
Other aetiologic factor – patients with a risk factor not in the top 12. No recognized risk factor- No history of any of the 91 risk factors in the 5 years preceding incident HF.

Key questions

For 92 HF risk factors (RFs):
(i)What is the evidence for association and effective RCT-based interventions (for HF: *RCT-HF*; CVD prevention, but unknown impact on HF: *RCT-CVD*; or no preventive intervention: *RCT-0*)?

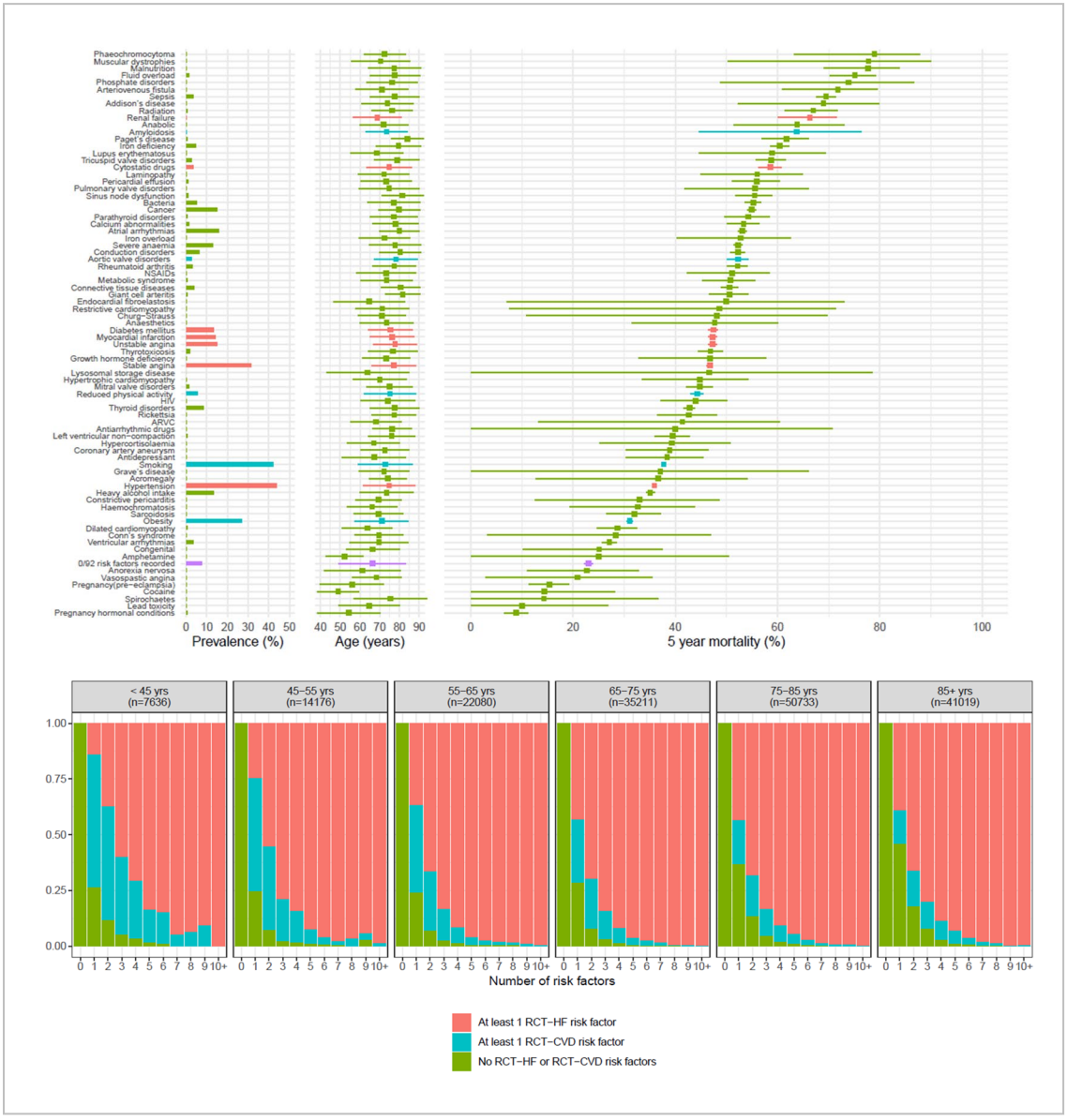
(ii)What is the preventable burden at population level?

Key findings

High-quality evidence for association (n=10) and interventions (7 RCT-HF; 6 RCT-CVD) were lacking for HF RFs. Among 170,885 incident HF cases, 28.5% had no RCT-HF RFs and 38.6% had no RCT-CVD RFs. 15.6% had no RF or only RCT-0 RFs.

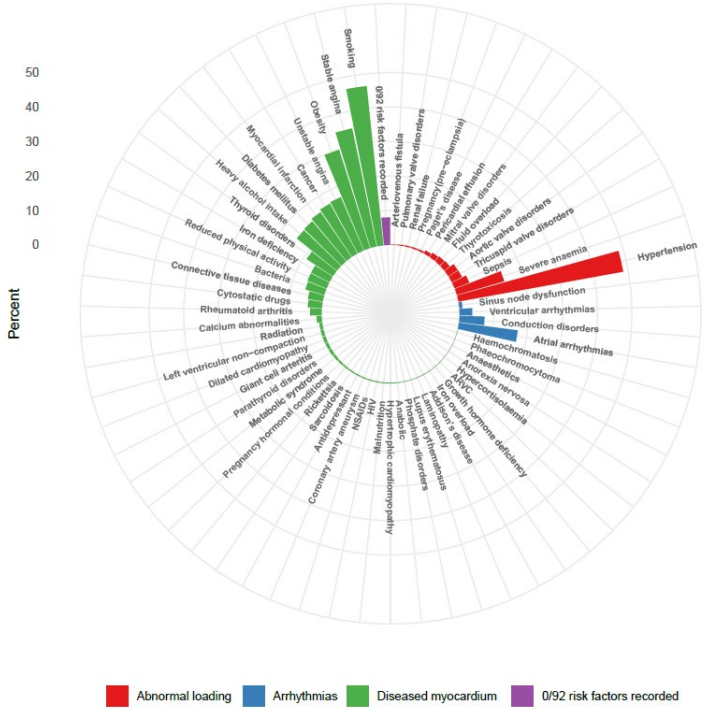
Take-home message

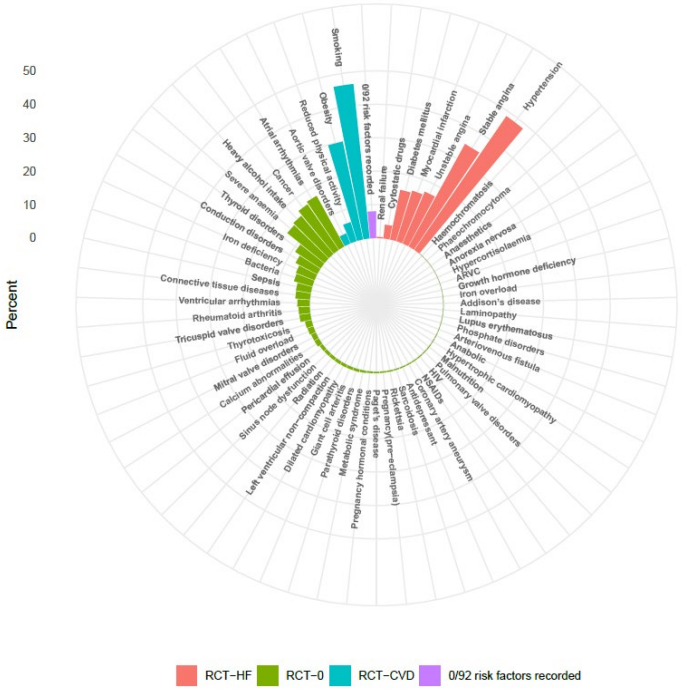
1 in 6 individuals with HF have no recorded RFs or RFs without trials. In a systematic map of primary preventive opportunities across 92 RFs for HF, we demonstrate frequent co-occurrence and need for trials tackling multiple RFs.



Graphical abstract

Figure 1: Prevalence of risk factors recorded any time in the 5 years before first diagnosis of heart failure in 170885 patients, classified by mode of action (diseased myocardium, abnormal loading, arrhythmic and other) and evidence for preventive treatment (RCT-HF, RCT-CVD, RCT-0 or 0/92 risk factors).





Factors with <100 patients are excluded from this plot.

Figure 2. Five-year all-cause mortality from time of incident HF diagnosis: by risk factors (n=89) in 170855 individuals with incident heart failure

Figure 3. Five-year mortality in patients with incident HF (n=170885) by the 12 most common risk factors at any time in the preceding 5 years

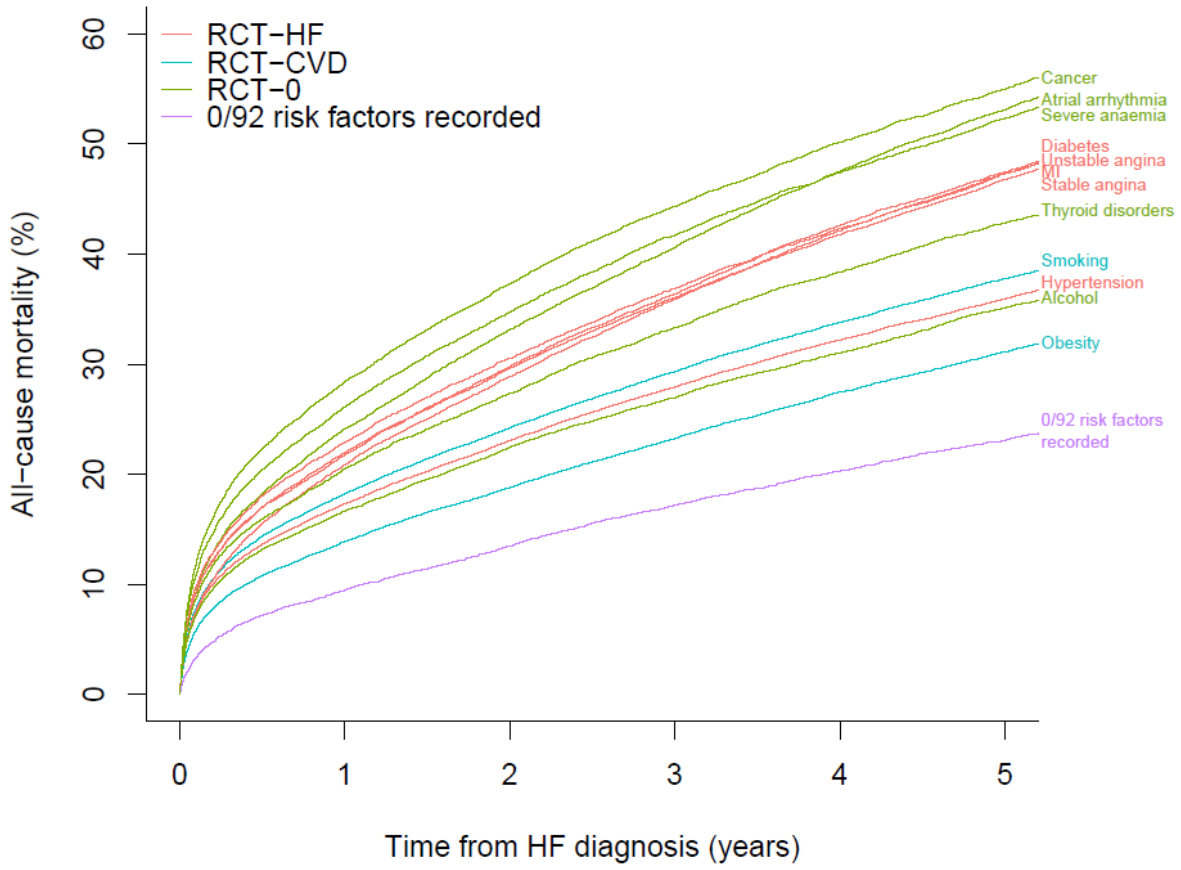
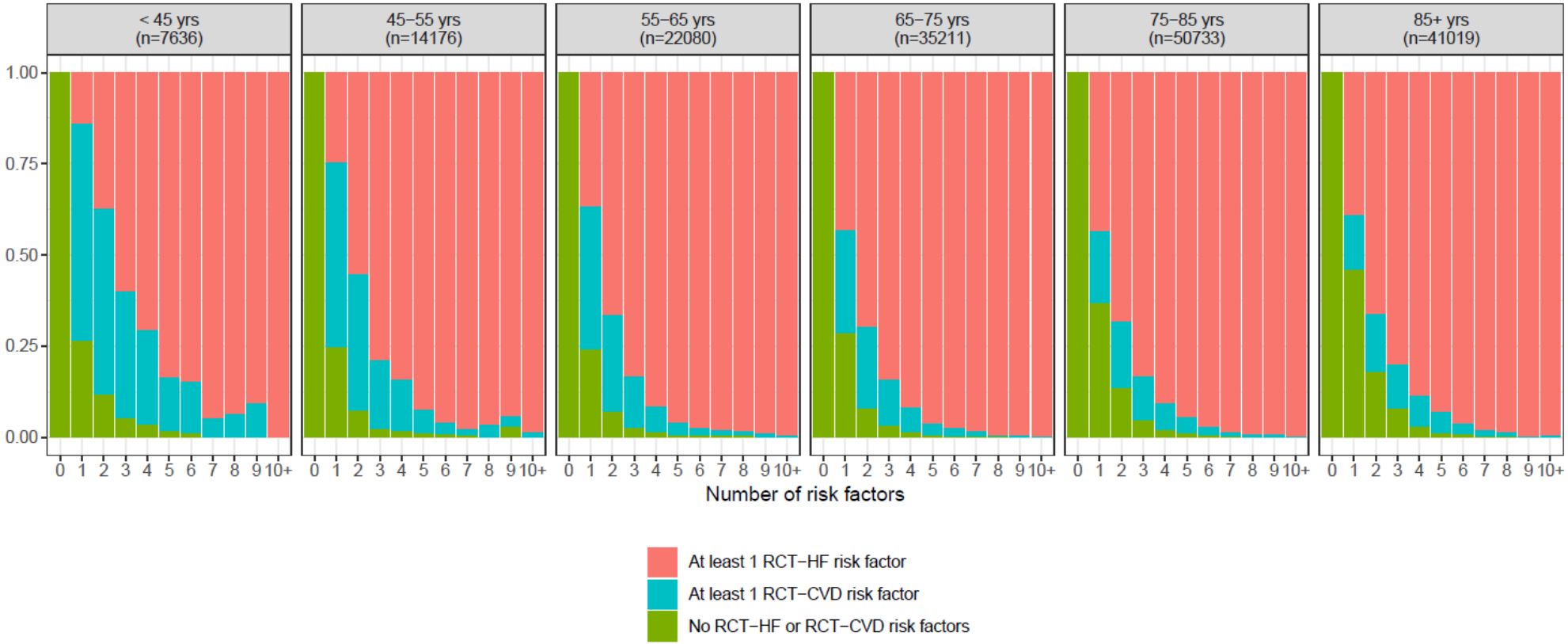


Figure 4: Number of risk factors co-occurring in patients and proportion of patients with at least 1 risk factor treatable for HF prevention or CVD prevention, stratified by age group (n=170855)



All relevant permissions have been given to use figures and tables in this manuscript.