Check for updates

ESC European Society of Cardiology

European Journal of Heart Failure (2023)
doi:10.1002/ejhf.2970

RESEARCH ARTICLE

Prediction models for heart failure in the community: A systematic review and meta-analysis

Ramesh Nadarajah[1,2,3]*, Tanina Younsi[3], Elizabeth Romer[3], Keerthenan Raveendra[4], Yoko M. Nakao[1,2], Kazuhiro Nakao[1,2,5], Farag Shuweidhi[6], David C. Hogg[7], Ronen Arbel[8,9], Doron Zahger[10,11], Zaza Iakobishvili[11,12], Gregg C. Fonarow[13], Mark C. Petrie[14], Jianhua Wu[6,15], and Chris P. Gale[1,2,3]

[1]Leeds Institute for Cardiovascular and Metabolic Medicine, University of Leeds, Leeds, UK; [2]Leeds Institute of Data Analytics, University of Leeds, Leeds, UK; [3]Department of Cardiology, Leeds Teaching Hospitals NHS Trust, Leeds, UK; [4]Faculty of Medicine and Health, University of Leeds, Leeds, UK; [5]Department of Cardiovascular Medicine, National Cerebral and Cardiovascular Center, Suita, Japan; [6]School of Dentistry, University of Leeds, Leeds, UK; [7]School of Computing, University of Leeds, Leeds, UK; [8]Community Medical Services Division, Clalit Health Services, Tel Aviv, Israel; [9]Maximizing Health Outcomes Research Lab, Sapir College, Sderot, Israel; [10]Department of Cardiology, Soroka University Medical Center, Beer Sheva, Israel; [11]Faculty of Health Sciences, Ben Gurion University of the Negev, Beer Sheva, Israel; [12]Department of Community Cardiology, Clalit Health Fund, Tel Aviv, Israel; [13]Division of Cardiology, Department of Medicine, University of California at Los Angeles, Los Angeles, CA, USA; [14]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK; and [15]Wolfson Institute of Population Health, Queen Mary University of London, London, UK

Received 15 May 2023; revised 25 May 2023; accepted 1 July 2023

Aims Multivariable prediction models can be used to estimate risk of incident heart failure (HF) in the general population. A systematic review and meta-analysis was performed to determine the performance of models.

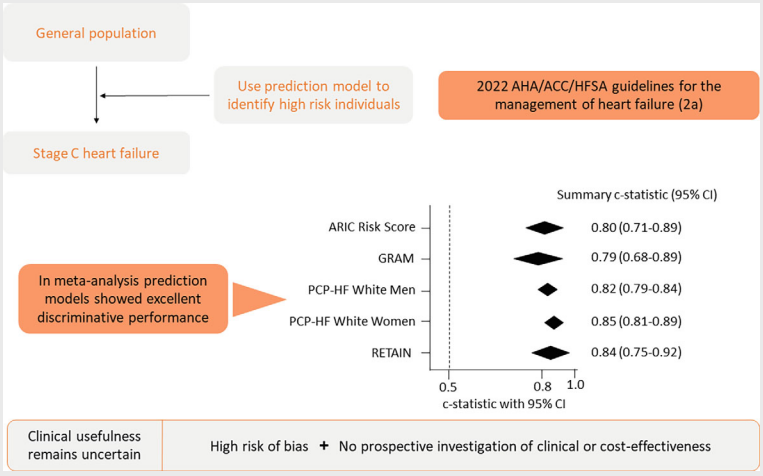
Methods and results From inception to 3 November 2022 MEDLINE and EMBASE databases were searched for studies of multivariable models derived, validated and/or augmented for HF prediction in community-based cohorts. Discrimination measures for models with c-statistic data from ≥3 cohorts were pooled by Bayesian meta-analysis, with heterogeneity assessed through a 95% prediction interval (PI). Risk of bias was assessed using PROBAST. We included 36 studies with 59 prediction models. In meta-analysis, the Atherosclerosis Risk in Communities (ARIC) risk score (summary c-statistic 0.802, 95% confidence interval [CI] 0.707–0.883), GRaph-based Attention Model (GRAM; 0.791, 95% CI 0.677–0.885), Pooled Cohort equations to Prevent Heart Failure (PCP-HF) white men model (0.820, 95% CI 0.792–0.843), PCP-HF white women model (0.852, 95% CI 0.804–0.895), and REverse Time AttentIoN model (RETAIN; 0.839, 95% CI 0.748–0.916) had a statistically significant 95% PI and excellent discrimination performance. The ARIC risk score and PCP-HF models had significant summary discrimination among cohorts with a uniform prediction window. 77% of model results were at high risk of bias, certainty of evidence was low, and no model had a clinical impact study.

Conclusions Prediction models for estimating risk of incident HF in the community demonstrate excellent discrimination performance. Their usefulness remains uncertain due to high risk of bias, low certainty of evidence, and absence of clinical effectiveness research.

*Corresponding author. Leeds Institute for Cardiovascular and Metabolic Medicine, University of Leeds, 6 Clarendon Way, Leeds LS2 9DA, UK. Tel: +44 113 3433241, Email: r.nadarajah@leeds.ac.uk

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Graphical Abstract



Prediction models for risk of incident heart failure in the community demonstrated excellent prediction performance but their usefulness in clinical practice remains uncertain. ARIC, Atherosclerosis Risk in Communities; CI, confidence interval; GRAM, GRaph-based Attention Model; PCP-HF, Pooled Cohort equations to Prevent Heart Failure; RETAIN, REverse Time AttentIoN model.

Keywords Heart failure • Prediction • Community • Prevention • Systematic review • Meta-analysis

Introduction

Heart failure (HF) affects more than 64 million people worldwide,[1] with its incidence and prevalence continuing to increase.[2] Patients with HF are subject to reduced quality of life and premature mortality,[3] and the global economic burden of HF already exceeds $100 billion.[4] Accordingly, interventions to prevent the progression from stage A/B HF (at-risk/pre-HF) to stage C HF (symptomatic HF) are advocated by international guidelines.[5,6]

However, there is substantial heterogeneity of risk within HF stage A/B, and specific groups are not included in current HF staging schema but are nonetheless at increased risk for symptomatic HF.[7] To match the intensity of prevention efforts with the absolute risk of the individual, an accurate quantification of future risk for HF is needed. Incident HF may be predicted from multivariable risk prediction models, and the 2022 American Heart Association/American College of Cardiology/Heart Failure Society of America (AHA/ACC/HFSA) guidelines for the management of HF provide a 2a (level of evidence B, non-randomized) recommendation for their use in clinical practice.[5]

Previous reviews have summarized risk prediction models for incident HF,[8,9] but have not considered whether those models could be applied in a community setting, where it is most likely they would be of use in routine clinical practice, and where the majority of cases of HF locate. Such models should consist of variables commonly available from patient records an not require additional or specialist investigations. Furthermore, previous reviews have not provided a quantitative synthesis of model performance across different cohorts, which limits understanding of model generalizability. They also predate the use of machine learning for risk prediction for HF,[10] and the prediction of HF phenotypes (HF with reduced ejection fraction [HFrEF], and HF with preserved ejection fraction [HFpEF]).[11] Thus, there is an incomplete understanding for whether the use of multivariable risk prediction models for incident HF in the general population could bring about clinical benefit.

We therefore performed a systematic review and meta-analysis to provide an overview of HF risk prediction models that are applicable and have been validated in community-based cohorts. We synthesize the discriminatory abilities of included risk prediction models and investigate the robustness of their development and evaluation to determine which, if any, may be suitable for clinical use.

Methods

Search strategy and inclusion criteria

We searched the MEDLINE and Embase databases through the Ovid platform from inception through 3 November 2022. We used a combination of keywords and subject headings related to HF and prediction models based on previous literature, and the search was limited to the English language (online supplementary material).[8,9,12] We completed forward and backward citation searching for included studies and previous systematic reviews.[8,9] Duplicates were removed using Endnote's duplicate identification strategy and then manually.

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

To be eligible for inclusion a study had to be an original study in human adults (≥18 years of age), develop and/or validate a prediction model(s) for incident HF based on multivariable analysis, and be written in English. Articles were excluded if they included patients with HF at baseline, only reported measures of association between risk factors and incident HF rather than a full prediction model, studied only a subset of the general population (e.g. individuals diagnosed with a particular morbidity), or incorporated variables that would not be routinely available at point of care in the community (e.g. electrocardiographic [ECG] parameters) (online supplementary material). Models developed in clinical trials which were not subsequently validated in a community-setting were also excluded.[5] To be included in meta-analysis a model had to have c-statistic data from ≥3 cohorts.

We uploaded records to a systematic review web application (Rayyan, Qatar Computing Research Institute).[13] Two investigators (TY, ER) independently screened them for inclusion by title, abstract and full text and supplemental materials. Disagreements were resolved by consultation with a third investigator (RN). This review was registered on PROSPERO (CRD42022380892) and informed by the PRISMA statement and CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS; online supplementary material).[14,15]

Data extraction and quality assessment

Two investigators (TY, ER) independently extracted the data from the included studies. All data came from the primary reference, unless otherwise stated. Two investigators (KR, RN) assessed each model in each study for risk of bias and applicability to our review question using the Prediction model Risk Of Bias ASsessment Tool (PROBAST).[16] Discrepancies between reviewers were resolved through additional review during group discussions.

To allow quantitative synthesis of the predictive performance of the models we extracted measures of discrimination and calibration.[17] We extracted data on the c-statistic or area under the receiver operating characteristic curve (AUROC) and corresponding 95% confidence interval (95% CI). When the 95% CI was not reported, we calculated it using methods described by Debray *et al.*[17] We extracted data on the *p*-value of a goodness-of-fit test and the reported ratio for observed to expected (O:E) events or calibration slope. When authors performed augmentation of pre-existing models by adding variables with an aim to enhance predictive value of models, we retrieved the net reclassification improvement (NRI) index of the augmented model compared with the original 'simple' model, as well as the augmented model's performance in terms of discrimination and calibration. We included augmentation data only when the augmentation variables were applicable to primary care settings as outlined previously. We did not extract model performance data when it related to artificial/synthetic data.[18] We also checked for reporting of clinical utility of a model (net benefit in the form of decision curve analysis or decision analytical modelling, which can be used to integrate the benefits and harms of using a model for clinical decision support) and conducted forward citation searching for studies determining the impact (clinical and cost-effectiveness) of using models in real-world clinical practice.

Data synthesis and statistical analysis

We reported continuous variables as means ± standard deviation and categorical variables as percentages. We evaluated statistical significance in all analyses at the 0.05 level. In individual studies, we assessed the c-statistic/AUROC of a model, where a 95% CI containing 0.5 indicates insufficient discrimination. Calibration of a model was deemed sufficient when authors reported a *p*-value of >0.05 and/or an O:E ratio or calibration slope ranging between 0.95 and 1.05. In assessing augmentation, we defined significant improvement as a positive NRI index with a reported 95% CI that did not contain 0 or with a *p*-value of <0.05. When a study reported on multiple cohorts, and presented separate data for each cohort, we assessed model performance separately for each cohort within that study. Funnel plots were produced as a visual check for publication bias.[19]

We conducted a Bayesian meta-analysis of discrimination through a summary measure of c-statistic and corresponding 95% CI. We calculated the 95% prediction interval (PI) to depict the extent of between-study heterogeneity and to indicate a possible range for prediction model performance in a new validation.[20] When the 95% CI or PI of the summary c-statistic included 0.5, we concluded that there was insufficient evidence that the prediction model has statistically significant discriminatory ability for incident HF in such populations as included in the meta-analysis.[21,22] Summary c-statistics of <0.60, 0.60–0.70, 0.70–0.80, and >0.80 were defined a priori as inadequate, adequate, acceptable and excellent based on prior publications.[23] We conducted meta-analyses in R using the metafor and metamisc package (R Foundation for Statistical Computing 3.6.3).[24–26]

Our primary analysis assessed overall discrimination for models that had ≥3 cohorts with c-statistic data. In the secondary analysis we performed a meta-analysis for each model with ≥3 cohorts reporting c-statistic data while applying a uniform prediction window (e.g. 5 or 10 years) since this is an important methodological consideration when wanting to translate summary risk model performance to clinical settings.[15] We performed sensitivity analyses in which we restricted the primary analyses to only those studies where the participants domain in PROBAST assessment was 'low' or 'unclear' risk of bias, and to only those studies where the overall PROBAST assessment was 'low' or 'unclear' risk of bias.

The Grading of Recommendations, Assessment, Development and Evaluation approach was used to assess the certainty of the evidence (online supplementary material).[27] The certainty of the evidence was graded as 'high', 'moderate', 'low' or 'very low'. One investigator (RN) rated the certainty of the evidence for the primary outcome and this was checked by a second investigator (JW).

Results

Study selection

We identified 12 297 unique records, reviewed 244 full-text reports and included 36 studies (*Figure 1*). A list of excluded studies that met a number of the inclusion criteria is available in online supplementary material.

Characteristics of included studies

The included studies were based on 34 different cohorts (15 prospective cohorts, 19 retrospective electronic health record [EHR] datasets), of which 23 (68%) were located in the United States (US) or Europe (*Table 1*, online supplementary *Table S1*). The number of participants ranged from 747 to 1 904 312, mean age 38–78 years, proportion of women 0–62%, and mean follow-up 4 months to 23.7 years. In 13 studies, baseline characteristics were not reported for the derivation and/or validation cohort.

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

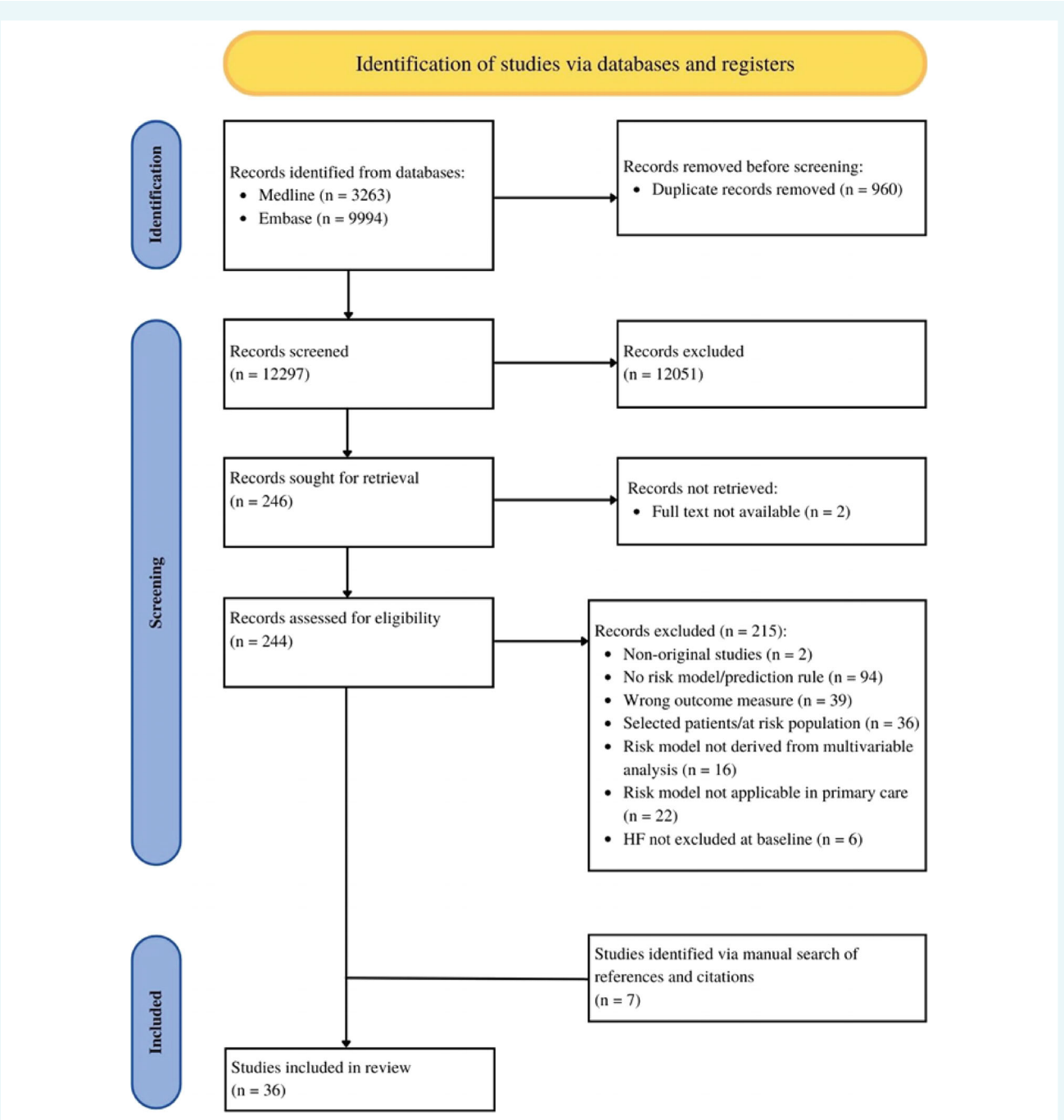


Figure 1 The Preferred Reporting Items for Systematic Review and Meta-Analysis flow-chart of studies included in the meta-analysis. HF, heart failure.

Characteristics of included prediction models

The included studies represented data on 59 multivariable prediction models. Eight models for predicting incident HF were sex-specific,[28–31] eight were sex- and race-specific,[32,33] and three were age-specific (online supplementary *Tables S2* and *S3*).[30] Two models specifically predicted HFpEF,[34,35] one specifically predicted HFrEF,[35] with the rest predicting HF irrespective of ejection fraction. Thirty-one models were developed in prospective community cohorts and 28 were derived in EHR datasets; prediction horizons ranged from 3 months to 30 years.[32,36] Thirty-four models were derived using multivariable Cox or logistic regression, and 25 by various machine learning methods including gradient boosting, random forest and deep learning (online supplementary *Table S3*).

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 Characteristics of included studies

Study	Cohort (country)	Study aim	Cohort type	HF cases/total patients (%)	Age (years, mean±SD)	Female sex (%)	BMI (kg/m^2, mean±SD)	DM (%)	HTN (%)	IHD (%)	Outcome	Outcome coding	Enrolment period (mean f/u in years)	Exclusion criteria
Agarwal, 2012[41]	ARIC (USA)	D	P	1487/13 555 (11.0)	54.1 (5.8)[b]	44.0	27.6 (5.2)[b]	10.3[b]	29.1[b]	4.1[b]	HF	ICD-9/10	1987–1989 (15.5)	Age <45 or >64, missing baseline data, race other than black or white
Arshi, 2021[28]	Rotterdam Study (NL)	D	P	918/6389 (14.4)	69.7 (8.3)	57.1	30 (4)[b]	12.9[b]	32.5[b]	7.6[b]	HF	ESC criteria	1989–2001 (13.0[a])	Incomplete data, lost to f/u
Bavishi, 2020[39]	CLRPP (USA)	D	P	1339/21 240 (6.3)	52.0 (12.0)	58.0	N/S	6.3	24.1[b]	Excluded	HF	ICD-9 and -10	1985–2000 (12)	HF at baseline Incomplete baseline measurement of risk factors
	NWMEDW (USA)	EV	EHR	568/31 256 (1.8)	51.4 (–)	57.0	28.9[b] (–)	11.4[b]	0.5[b]	Excluded	HF	ICD-9 and -10	2005–2013 (5)	Presence of HF, CAD, PAD, stroke, PPM, ICD
Bennis, 2022[69]	Nivel Primary Care Database (NL)	D	EHR	8543/17 086 (50.0)	81.0[a] cases, 76.0[a] controls	60.1	N/S	N/S	N/S	N/S	HF	ICPC	2012–2019 (–)	Data over <3 continuous
Brouwers, 2014[70]	PREVEND (NL)	D	P	168/8569 (2.0)	49.0 (12.7)[b]	50.0	26.4 (4.08)[b]	N/S	13.6	6.0[b]	HF	ESC criteria	1997–1998 (12.5)	Baseline HF, pregnancy, IDDM
Chahal, 2015[42]	MESA (USA)	D	P	176/6814 (2.6)	62.1 (10.0)	52.9	28.3 (5.5)	12.6	44.9	N/S	HF	Expert panel	2000–2002 (4.7[a])	Age <45 or >84, CV disease, AF, pregnancy, active cancer, wt >136 kg
Che, 2017[71]	Unspecified	D	EHR	3357/218 680 (1.5)	N/S	N/S	N/S	1.0	N/S	N/S	HF	ICD-9	N/S	<50 event records
Choi, 2016[72]	Sutter-PAMF (USA)	D	EHR	3884/32 787 (11.8)	N/S	N/S	N/S	N/S	N/S	N/S	HF	ICD-9	N/S	N/S
Choi, 2017[74]	Sutter-PAMF (USA)	D	EHR	3884/32 787 (11.8)	N/S	N/S	N/S	N/S	N/S	N/S	HF	ICD-9	2000–2013	Age <40 or >89 at diagnosis, first encounter outside observation window
Choi, 2018[73]	Sutter-PAMF (USA)	D	EHR	3408/258 555 (1.3)	N/S	N/S	N/S	N/S	N/S	N/S	HF	ICD-9	2000–2013	<2 visits
D'Agostino, 2008[31]	FHS (USA)	D	P	111/8491 (1.3)	49.0 (10.9)[b]	53.3	N/S	5.0	11.0	N/S	HF	Expert panel	1968–1987	Age <30 or >63, baseline CV disease, missing data
Ergatoudes, 2018[37]	The Study of Men Born in 1943 (SE)	D	P	85/747 (11.4)	50.0 (0.0)	0.0	26.2 (3.4)[b]	3.6[b]	39.4[b]	1.3[b]	HF	ESC criteria	1993–1993 (21.0)	Not born in 1943, not male, resident outside Gothenburg

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 (Continued)

Study	Cohort (country)	Study aim	Cohort type	HF cases/total patients (%)	Age (years, mean±SD)	Female sex (%)	BMI (kg/m², mean±SD)	DM (%)	HTN (%)	IHD (%)	Outcome	Outcome coding	Enrolment period (mean f/u in years)	Exclusion criteria
Gaziano, 2021[35]	VA-CDW (USA)	D, EV	P	138 992/ 190 4312 (7.3)	51.0 (–)	3.5	29.3[b] (–)	22.0	48.8[b]	26.7[b]	HFpEF/rEF	ICD-10, NLP for HFpEF	2002–2007 (11.1[a])	Baseline HIV, CKD, liver disease, hepatitis, dialysis, cancer, major mental health disorder, dementia, cirrhosis, amputation. Non-white or non-black race
Goyal, 2010[29]	Kaiser Permanente Georgia (USA)	D	EHR	4001/359 947 (1.1)	38.2 (14.2)[b]	53.2	N/S	3.4	12.4	1.0	HF	ICD-9	2000–2005	Age <18, baseline HF
Ho, 2016[34]	CHS (USA)	D	P	386/5277 (7.3)	73.0 (6.0)	57.6	26.7 (4.7)	16.0	45.0[b]	25.0[b]	HFpEF	LVEF >50%, expert panel	1989–1993 (15.0)	Age <30 at baseline, missing data
	FHS (USA)	D	P	296/9496 (3.1)	58.0 (14.0)	54.8	26.7 (4.8)	6.0	26.0[b]	11.0[b]	HFpEF	LVEF >50%, expert panel	1979–1998 (15.0)	Age <30 at baseline, missing data
	PREVEND (NL)	D	P	113/7369 (1.5)	49.0 (12.0)	50.2	26.1 (4.2)	4.0	14.0[b]	10.0[b]	HFpEF	LVEF >50%, expert panel	1997–1998 (15.0)	Age <30 at baseline, missing data
	MESA (USA)	EV	P	114/6678 (1.7)	62.0 (10.0)	52.7	28.3 (5.5)	13.0	37.0[b]	0.0[b]	HFpEF	LVEF >50%, expert panel	2000–2002 (15.0)	Age <30 at baseline, missing data
Jin, 2018[75]	Unspecified (CN)	D	EHR	5000/20 000 (25.0)	N/S	N/S	N/S	N/S	N/S	N/S	HF	Diagnostic events	N/S	N/S
Ju, 2021[76]	IBM (country unspecified)	D	EHR	2660/7980 (33.3)	N/S	N/S	N/S	N/S	N/S	N/S	HF	ICD-10	N/S (0.5)	N/S
Khan, 2022[40]	Clalit Health Services (IL)	EV	EHR	16 351/ 1 394 411 (1.2) 5-year window	49.6 (13.2)	55.9	29.6 (4.4) men, 27.0 (5.7) women	9.6 men, 9.5 women	N/S	N/S	HF	ICD-9, new HF medication	2000–2018 (5.0)	Age <30 or >80, baseline, CHD, stroke, AF. Insufficient f/u
				34 505/760 750 (4.5) 10-year window	52.4 (12.1) men, 54.0 (12.4) women	59.3	27.6 (4.4) men, 28.1 (5.7) women	12.7 men, 11.6 women	N/S	N/S	HF	As above	2000–2018 (10.0)	As above
Khan, 2022[32]	ARIC (USA)	D	P	2384/24 838 (9.6)[c]	43.5 (12.3)[c]	54.6[c]	26.6 (5.3)[c]	6.0[c]	15.0[b c]	Excluded	HF	ICD-9 and -10	N/S (23.7[a])[c]	Incomplete baseline measurement of risk factors
	CARDIA (USA)	D	P											
	FHS (USA)	D	P											
	FoF (USA)	D	P											
	MESA (USA)	D	P											

© 2023 The Authors. European Journal of Heart Failure published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 (Continued)

Study	Cohort (country)	Study aim	Cohort type	HF cases/total patients (%)	Age (years, mean±SD)	Female sex (%)	BMI (kg/m^2, mean±SD)	DM (%)	HTN (%)	IHD (%)	Outcome	Outcome coding	Enrolment period (mean f/u in years)	Exclusion criteria
Kwon, 2019[38]	HIRA-NPS (KR)	D, EV	EHR	–/63 030	N/S	N/S	N/S	N/S	N/S	N/S	HF	KCD-9	2014–2015 (–)	<5 visits
Lee, 2020[64]	Brigham & Women's Medicare (USA)	D	EHR	9701/138 388 (7.0)	71.9 (6.9)	61.6	N/S	18.0	55.3	2.9	HF	ICD-9, /NLP for HFpEF	2007–2014 (3.4)	Age <65, use of HF medications
Li, 2022[77]	CPRD (UK)	D, EV	EHR	31 514/954 983 (3.3)	N/S	N/S	N/S	N/S	N/S	N/S	HF	Caliber codes	1985–2014 (5.0)	No IMD data, censored patients
Luo, 2020[78]	Unspecified	D, EV	EHR	3080/12 320 (73.7)	N/S	N/S	N/S	N/S	N/S	N/S	HF	ICD-9	N/S (0.5)	N/S
Ma, 2018 HEALTH ATM[18]	SNOW (country unspecified)	D	EHR	5902/39 050 (15.1)	N/S	N/S	N/S	N/S	N/S	N/S	HF	ICD-9	N/S	Diagnosis in <2 encounters, no HF medications at diagnosis
Ma, 2018 PRIME[79]	Unspecified	D, EV	EHR	2403/7571 (31.7)	N/S	N/S	N/S	N/S	N/S	N/S	HF	ICD-9	N/S (0.8)	<5 visits, ICD-9 codes appearing <5 instances
Nowak, 2020[47]	Biobank (UK)	D, EV	P	1701/450 212 (0.4)	56.5 (8.1)	54.1	27.4 (4.7)	5.2	10.1	7.0	HF / HF death	ICD-9	2006–2010 (10.0)	Missing creatinine or cystatin C
Rao, 2022[80]	CPRD (UK)	D, EV	EHR	13 050/100 071 (13.0)	70.0[a]	58.3	N/S	20.6	65.7	9.3	HF	Caliber codes	N/S (9.0[a])	Age <16, <10 visits, <10 codes, <3 years records, data unlinked with secondary care
Rasmy, 2018[81]	CHF (USA)	EV	EHR	152 790/ 13 05 307 (11.7)	N/S	53.9	N/S	N/S	N/S	N/S	HF	ICD-9	N/S	Age <50 at diagnosis
Smith, 2010[43]	MDCS-CC (SE)	D	P	112/5187 (2.2)	57.6 (5.9)	41.0	25.7 (3.9)	7.7	16.4	1.4	HF	ICD-9/10	1991–1994 (14)	Age <45 or >64, missing data
Stenemo, 2018[82]	PIVUS (SE)	EV	P	170/1586 (10.7)[c]	N/S	50.0	27.0 (±4.2)	12.3[c]	36.6[c]	7.6[c]	HF	ICD-9/10	2001–2004 (10[a])	Missing or low-quality data
	ULSAM (SE)	EV	P		N/S	0.0	26.2 (±3.4)						1970–1973 (8[a])	
Tromp, 2021[30]	FHS (USA)	D	P	1381/24 675 (5.6)[c]	56.0 (14.0)[c]	52.6[c]	27.0 (5.0)[c]	7.1[c]	24.5[c]	N/S	HF	Signs, symptoms, LVEF	1979–1998 (12.7[a])[c]	Prevalent HF, missing data
	MESA (USA)	D	P											
	PREVEND (NL)	D	P											
Velagaleti, 2010[44]	FHS (USA)	D	P	95/2754 (3.4)	58.0 (10.1)[b]	54.0	28.0 (5.2)[b]	10.3[b]	27.3[b]	3.8[b]	HF	Framingham criteria	1995–1998 (9.4)	Missing covariate or biomarker
Goya Wannamethee, 2014[45]	BRHS (UK)	D	P	254/3870 (6.6)	68.6 (5.5)	0.0	26.8 (3.6)	6.9	70.7	25.6[b]	HF	ICD-9	1978-1980 (11)	Age <40 or >59, missing data

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 (Continued)

Study	Cohort (country)	Study aim	Cohort type	HF cases/total patients (%)	Age (years, mean±SD)	Female sex (%)	BMI (kg/m^2, mean±SD)	DM (%)	HTN (%)	IHD (%)	Outcome	Outcome coding	Enrolment period (mean f/u in years)	Exclusion criteria
Willeit, 2016[46]	Multi-cohort (26) (International)	D	P	2212/95 617 (2.3)	61.0 (10.0)	50.8	27.0 (4.5)	8.5	N/S	5.8	HF	ICD-8-10	1936–2006 (7.8[a])	High baseline BNP/NT-proBNP, studies without BNP/NT-proBNP, <1 year f/u, <10 incident events, missing baseline demographic information
Yin, 2019[83]	Unspecified	D, EV	EHR	425/1700 (25.0)	N/S	N/S	N/S	N/S	N/S	N/S	HF	CUI converted to ICD-9	N/S (0.3)	N/S
Zhang, 2019[84]	HF-I (country unspecified)	D, EV	EHR	1228/37 312 (3.3)	N/S	N/S	N/S	N/S	N/S	N/S	HF	ICD-9	N/S (0.5)	N/S

AF, atrial fibrillation; AHF, acute heart failure; ARIC, Atherosclerosis Risk in Communities; BMI, body mass index; BNP, B-type natriuretic peptide; BRHS, British Regional Heart Study; CARDIA, Coronary Artery Risk Development in Young Adults; CHD, coronary heart disease; CHF, Cerner Health Facts; CKD, chronic kidney disease; CLRPP, Cardiovascular Lifetime Risk Pooling Project; CPRD, Clinical Practice Research Datalink; CUI, concept unique identifier; CV, cardiovascular; D, derivation; EHR, electronic health record; ESC, European Society of Cardiology; EV, external validation; FHS, Framingham Heart Study; FoF, Framingham Offspring cohort; f/u, follow-up; HF, heart failure; HFpEF, heart failure with preserved ejection fraction; HFrEF, heart failure with reduced ejection fraction; HIRA-NPS, Health Insurance Review and Assessment-National Patient Samples; ICD, International Classification of Diseases; ICPC, International Classification of Primary Care; IDDM, insulin-dependent diabetes mellitus; IHD, ischaemic heart disease; IMD, index of multiple deprivation; KCD, Korean Standard Classification of Disease; LVEF, left ventricular ejection fraction; MDCS-CC, Malmö Diet and Cancer Study Cardiovascular Cohort; MESA, Multi-Ethnic Study of Atherosclerosis; NLP, natural language processing; N/S, not specified; NT-proBNP, N-terminal pro-B-type natriuretic peptide; NWMEDW, Northwestern Medicine Enterprise Data Warehouse; P, prospective cohort; PAMF, Palo Alto Medical Foundation; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; PREVEND, Prevention of Renal and Vascular End-stage Disease; ULSAM, Uppsala Longitudinal Study of Adult Men; VA-CDW, Veterans Affairs Corporate Data Warehouse; wt, weight.

[a]Median reported.

[b]Derived number.

[c]Value for combined cohorts.

Online supplementary *Tables S4* and *S5*, and *Figure S1* summarize the variables and variable types used. Hypertension, age and diabetes mellitus were included in ≥90% of models. Types of codes used for deep learning techniques were predominantly for diagnoses (24/25, 96.0%), and also medications (12/25, 48.0%) and demographics (10/25, 40.0%).

Risk prediction model performance among included cohorts

All studies reported a measure of discrimination, with reported c-statistics/AUROCs from 0.61 (95% CI 0.55–0.66)[37] to 0.95 (95% CI 0.95–0.95)[38] (online supplementary *Table S6*). Twenty-one regression models had calibration estimated (by goodness-of-fit test, calibration plot, or both), but none of the machine learning models. Eighteen models were externally validated, and 11 were externally validated in more than one other cohort/dataset. Nine models had acceptable calibration and good discrimination performance (0.76–0.85),[28,31,39,40] with only the Pooled Cohort equations to Prevent Heart Failure (PCP-HF) sex- and race-specific models reporting acceptable calibration and good discrimination on external validation.[39,40]

Clinical utility and clinical impact of included risk prediction models

None of the included studies conducted a clinical utility analysis, and forward citation searching did not find any studies of clinical impact for included risk prediction models.

Augmentation of included risk prediction models

We identified augmentation data applicable to primary care settings for 10 models (online supplementary *Table S7*), with the augmentation value of B-type natriuretic peptides (BNP or N-terminal proBNP [NT-proBNP]) tested for seven models,[28,37,41–46] C-reactive protein (CRP) for three models,[37,43,45] urine albumin-to-creatinine ratio (ACR)[44] for two models,[44,47] and high-density lipoprotein (HDL) cholesterol for one model.[46] Significant improvement by the addition of BNP alone was demonstrated for six models,[28,41–45] and also in combination with CRP[43] or urine ACR.[44]

Risk of bias assessment

Overall, 77% of model results were at high risk of bias (*Figure 2*, online supplementary *Table S8*, online supplementary *Figure S2*), predominantly driven by high risk of bias in the analysis domain (72%), especially due to the handling of missing data (67%).

Meta-analysis

Six models were eligible for the primary meta-analysis, three regression models (ARIC [Atherosclerosis Risk in Communities] risk score, PCP-HF white men and white women models) and

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

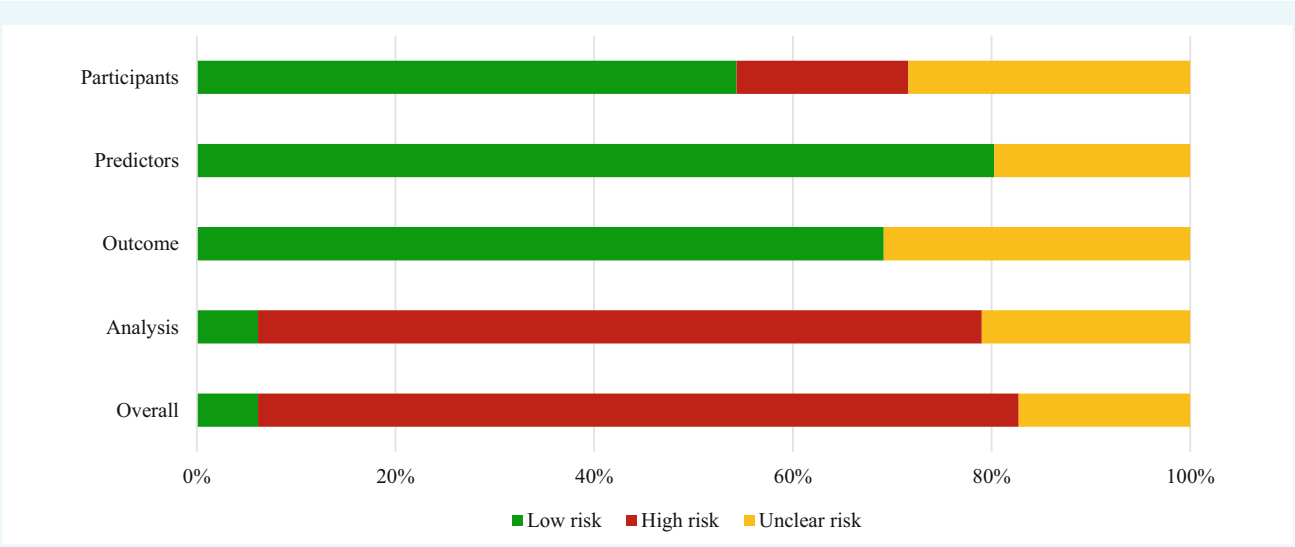


Figure 2 Risk of bias across all included studies.

three deep learning models (GRAM [GRaph-based Attention Model], RETAIN [REverse Time AttentIoN model] and RETAINEX [RETAIN with extra time dimensions and embedding matrices]; *Figure 3*). Despite high heterogeneity, five models had a statistically significant 95% PI and excellent discriminative performance: ARIC risk score, GRAM, PCP-HF white men and white women models, and RETAIN. The international collaboration for heart failure subtypes (ICHFS) HFpEF model showed acceptable summary discriminative performance for HFpEF prediction specifically (*Figure 4*).

Only the ARIC risk score (10-year) and the PCP-HF white men and white women models (5-year), had a statistically significant 95% PI among cohorts that had a uniform risk prediction window, and discriminative performance remained excellent (*Figure 5*). In our sensitivity analysis of restricting primary and secondary analyses to studies with 'low' or 'unclear' risk of bias for the participants domain of PROBAST, we found only the RETAIN and PCP-HF white men and white women models had a statistically significant 95% PI (online supplementary *Figure S3*), and only the PCP-HF white men and white women models at a uniform risk prediction window (5 years; online supplementary *Figure S4*). When restricting primary and secondary analysis to models with 'low' or 'unclear' risk of bias for overall PROBAST assessment, no models met eligibility for inclusion. Funnel plots were symmetrical but with additional horizontal scatter (online supplementary *Figure S5-S7*), consistent with the presence of between-study heterogeneity.

Certainty of evidence

The initial certainty level of the included prediction modelling studies was set at 'high' because the association between the predictors and outcomes was considered irrespective of any causal connection.[48] The overall certainty level was, however, downgraded to 'moderate', then 'low' because of inconsistent results given high heterogeneity and the high overall risk of bias in included studies. The final overall certainty of evidence was 'low', implying that our confidence in the effect estimates is limited and further research is very likely to change the effect estimate.

Discussion

This systematic review and meta-analysis provides an overview of 59 models in the community setting for estimating subsequent risk of incident HF in the general population. In the meta-analysis, five models had excellent discriminative performance for HF incidence, but only the ARIC risk score and PCP-HF models at a uniform prediction window. However, no model met eligibility for inclusion in meta-analysis if studies at overall high risk of bias were excluded, certainty of evidence was low, and none of the models underwent prospective investigation of clinical or cost-effectiveness, suggesting that their clinical usefulness remains uncertain (*Graphical Abstract*).

Clinical relevance

Whilst there have been advances in the treatment of stage C HF across the ejection fraction spectrum, prognosis still remains poor,[3] emphasizing the need to shift upstream to prevent irreversible myocardial damage that heralds symptomatic HF.[49] The 2022 AHA/ACC/HFSA HF guidelines recommend estimation of risk of incident HF in the general population,[5] and specifically example the Framingham Heart Failure risk score,[50] Health ABC Heart Failure score,[51] ARIC risk score,[41] and PCP-HF.[33] The Framingham Heart Failure risk score and Health ABC Heart Failure score were not included in this study as they require ECG interpretation for left ventricular hypertrophy, but ECGs are missing in over three-quarters of routinely-collected community-based medical records.[39] Here, we found that both the ARIC risk score and PCP-HF models, without ECG variables, showed excellent discrimination performance in community-dwelling individuals and they are both available as web-based tools to facilitate clinical application.

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

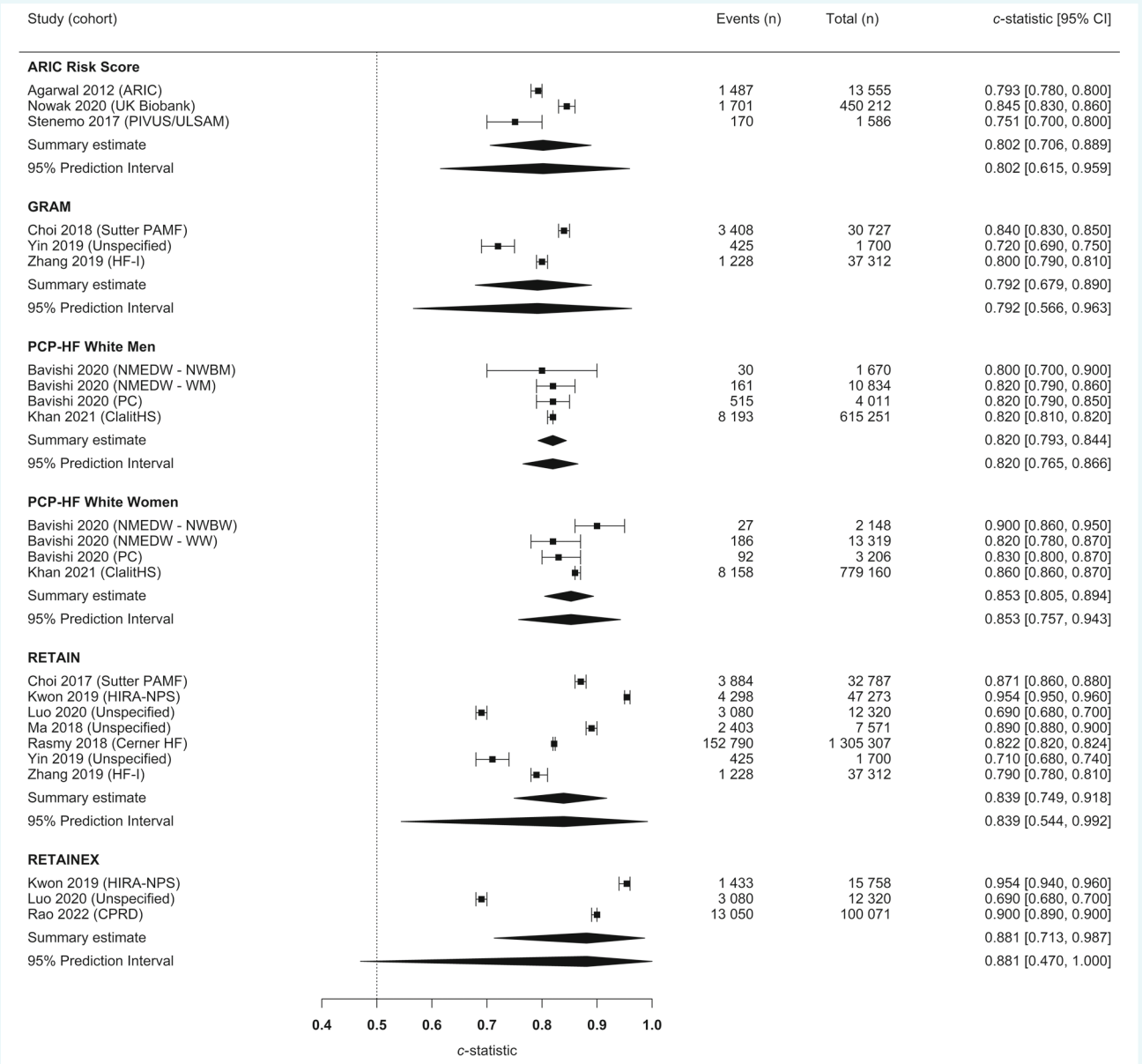


Figure 3 Primary analysis: meta-analysis of c-statistics. ARIC, Atherosclerosis Risk in Communities; Cerner HF, Cerner Health Facts; CI, confidence interval; Clalit HS, Clalit Health Services; CPRD, clinical practice research datalink; GRAM, GRaph-based Attention Model; HF-I, heart failure – I; HIRA-NPS, Health Insurance Review and Assessment Service National Patients Sample; NMEDW, Northwestern Medicine Enterprise Data Warehouse; NWBM – non-white and non-black men; PC, pooled cohort; PCP-HF, Pooled Cohort equations to Prevent Heart Failure; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; RETAIN, REverse Time AttentIoN model; RETAINEX, RETAIN with extra time dimensions and embedding matrices; Sutter PAMF, Palo Alto Medical Foundation; USLAM, Uppsala Longitudinal Study of Adult Men; WM, white men.

Furthermore, though the PCP-HF white men and white women models were derived as race-specific equations, they have also shown excellent performance in non-white and non-black individuals.[39]

Nonetheless the feasibility of implementing the PCP-HF and ARIC risk scores in routine clinical practice remains unknown. They both include systolic blood pressure and body mass index, but these may only be available for between a fifth and a third of the population in European countries in routinely-collected community-based medical records.[52,53] Furthermore, PCP-HF requires a complete dataset for glucose, total cholesterol and HDL-cholesterol, which may not be routinely tested in a large proportion of asymptomatic community-dwelling individuals. Accordingly, the use of each model may require additional appointments or investigations, placing extra burden on both healthcare professionals and patients.

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

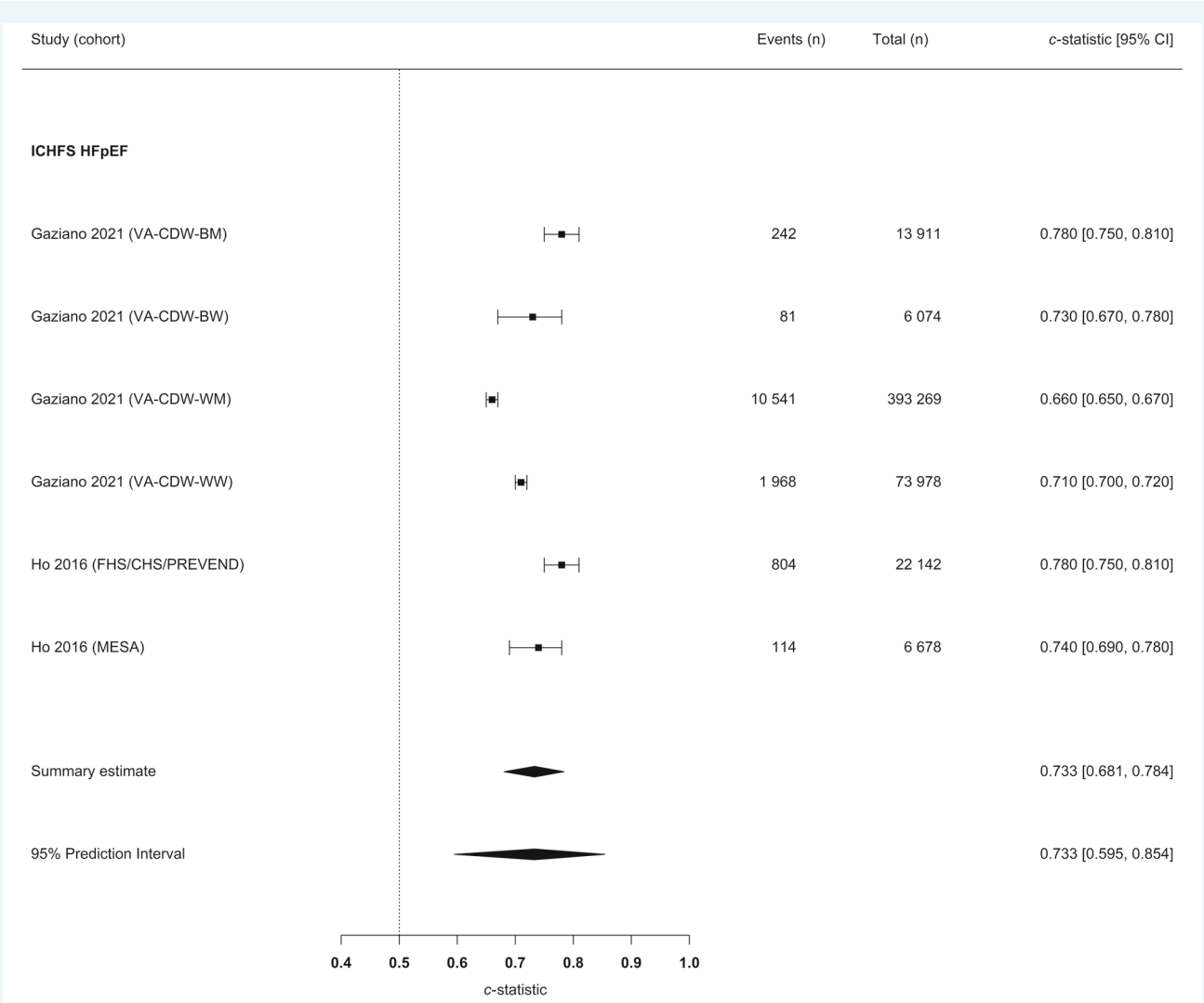


Figure 4 Meta-analysis of c-statistics for models for incident heart failure with preserved ejection fraction (HFpEF). BM, black men; BW, black women; CHS, Cardiovascular Health Study; CI, confidence interval; FHS, Framingham Heart Study; ICHFS, international collaboration for heart failure subtypes; MESA, Multi-Ethnic Study of Atherosclerosis; PREVEND, Prevention of Renal and Vascular End-stage Disease; VA-CDW, Veterans Health Administration corporate data warehouse; WM, white men; WW, white women.

Multiple randomized controlled trials (RCTs) demonstrate that incident HF can be reduced in hypertensive patients with blood pressure control, and in type 2 diabetic patients with sodium–glucose cotransporter 2 inhibitor treatment.[54–56] The value of systematic screening and prevention for HF beyond guideline-adherent therapy for these narrow subgroups is unknown. The STOP-HF RCT demonstrated that BNP-guided collaborative care in a broad community cohort reduced *de novo* asymptomatic left ventricular systolic dysfunction and emergency cardiovascular hospitalisation.[57] Cost-effectiveness analysis of this programme suggests that savings in cardiovascular hospitalizations offsets increased outpatient and primary care costs.[58] However >85% of patients treated did not have a clinical event over more than 4 years of follow-up, and hospitalizations for HF were not reduced by the intervention.[57,58] Identifying a higher risk cohort with a multivariable risk score may improve clinical and cost-effectiveness of primary prevention for HF. However, there has yet to be a study assessing the scale to which HF risk scores can be implemented into clinical practice, whether interventions based on predicted risk reduce the later occurrence of HF, and whether this reduces costs at a health system level. Overall, primary prevention programmes for HF have yet to become routine practice, in contrast to prevention of vascular disease,[59] and the usefulness of risk prediction models for HF to improve patient outcomes and cost-effectiveness of care remains uncertain.

Previous work

In concordance with previous reviews we observed sub-optimal conduct in model development,[9] and a failure to progress risk scores to impact studies.[8] We provide a number of further advances. Previously summaries exclusively referenced models

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

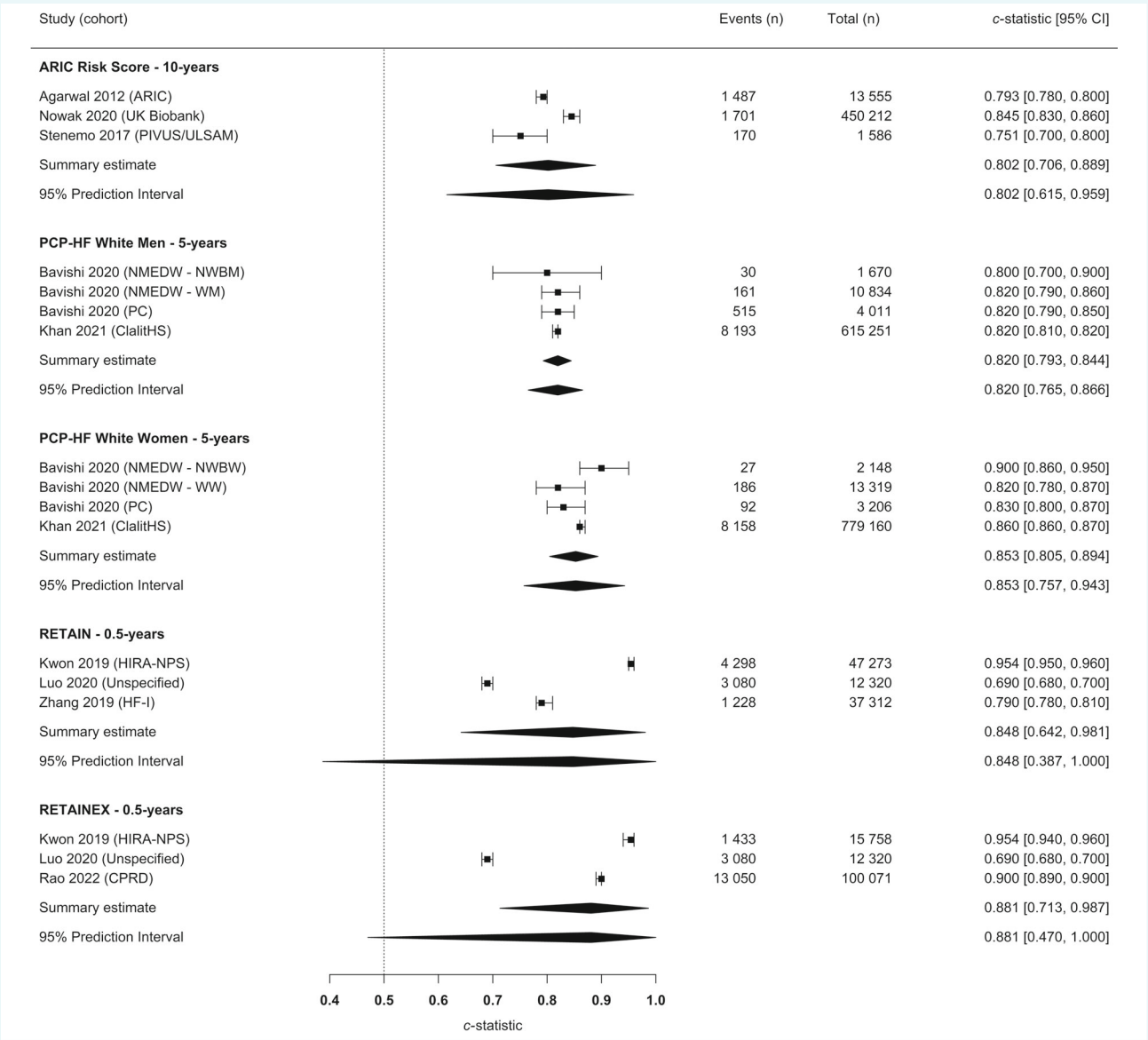


Figure 5 Secondary analysis: meta-analysis of C-statistics grouped according to application of a uniform prediction window within a model. ARIC, Atherosclerosis Risk in Communities; CI, confidence interval; Clalit HS, Clalit Health Services; CPRD, clinical practice research datalink; HF-I, heart failure – I; HIRA-NPS, Health Insurance Review and Assessment Service National Patients Sample; NMEDW, Northwestern Medicine Enterprise Data Warehouse; NWBM – non-white and non-black men; PC, pooled cohort; PCP-HF, Pooled Cohort equations to Prevent Heart Failure; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; RETAIN, REverse Time AttentIoN model; RETAINEX, RETAIN with extra time dimensions and embedding matrices; USLAM, Uppsala Longitudinal Study of Adult Men; WM, white men.

derived by regression, but more than 40% of the included models here were derived though machine learning approaches, highlighting the trend towards data-driven computational modelling for prediction in cardiovascular disease.[12] We observed a proliferation of models that were sex- and/or race-specific, demonstrating the increasing understanding of different epidemiology, clinical characteristics and pathophysiology of HF between men and women and across different racial groups.[60,61] Additionally, we provide a quantitative synthesis which demonstrates how models may perform when applied to a new population.

Machine learning for prediction of heart failure risk

We found that unsupervised deep learning models had equally excellent discriminative performance as the ARIC risk score and PCP-HF, though barriers exist for such models to be implemented.[62] Deep learning studies summarized here also frequently provided insufficient detail of the training and test cohorts, and there was an absence of calibration reporting, meaning the applicability of the results remains uncertain. Adherence to reporting

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

guidelines designed specifically for risk prediction machine learning studies, currently under development,[63] would improve the quality of reporting and increase confidence in the translatability of these models to clinical practice.[62]

Predicting specific heart failure phenotypes

We observed a transition in recent years for investigators to test performance of HF prediction models for HFpEF and HFrEF phenotypes,[64] and to develop phenotype-specific models.[34,35,64] On meta-analysis the discriminatory performance for HFpEF was inferior to overall HF. This may be due to challenges in defining and identifying HFpEF cases (online supplementary *Table S2*), but also suggests that HFpEF prediction is more difficult and may require additional variables.

Strengths and limitations

We had a comprehensive search strategy, thorough analysis approach, diverse domain expertise amongst the reviewers, and only included models that had been tested in the general population, which ensures the applicability of our results for primary prevention in a primary care setting. However, we acknowledge limitations in our study. Meta-analysis of model calibration performance was prohibited by the lack of reporting of such analyses. We restricted our search to studies written in English, though this has not been found to lead to significant bias.[65] We also did not present meta-regression or subgroup meta-analysis to investigate heterogeneity between studies based on study-level characteristics or subgroups in the absence of available individual patient data given that such analyses would be prone to ecological bias,[66] and are inferior to subgroup results derived from individual participant data (IPD).[17] An IPD meta-analysis, however, was not the scope of the current study. Between-study heterogeneity can occur due to differences in study characteristics, differences in study quality, or differences between studied populations. We included both prospective and retrospective cohort studies, which may have introduced bias as mildly symptomatic patients may be less likely to undergo cardiac testing in routine practice. Study populations varied in mean age, proportion who were women, comorbidity burden, and percentage of observed HF cases. More than three quarters of included studies were at high risk bias, predominantly related to improper handling of missing data. This is a commonly observed shortfall in prediction modelling research,[67] even in models recommended for use in healthcare.[68]

Conclusion

This systematic review and meta-analysis identified 59 risk prediction models for incident HF applicable in the community. We observed that machine learning and specification by sex, race and HF phenotype are increasingly common in HF risk prediction modelling. Prediction models showed excellent prediction performance for incident HF. However high risk of bias, missingness of requisite variables in routinely collected data, low certainty of evidence, and a lack of impact studies means that the usefulness of integrating HF prediction models into clinical practice remains uncertain.

Supplementary Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Conflict of interest: Y.M.N. reports a study grant from Bayer. Z.I. reports consulting fees and payments from Pfizer, Bayer, AstraZeneca, Boehringer Ingelheim, Sanofi, Medison and Novo Nordisk; reports support for attending meetings from Bayer; has participated in Data Safety Monitoring Boards or Advisory Boards for Bayer, Boehringer Ingelheim, AstraZeneca, Novo Nordisk, and Sanofi. G.C.F. reports consulting fees from Abbott, Amgen, AstraZeneca, Bayer, Cytokinetics, Eli Lilly, Janssen, Medtronic, Merck, Novartis and Pfizer. M.C.P. reports grants from Boehringer Ingelheim, Roche, SQ Innovations, AstraZeneca, Novartis, Novo Nordisk, Medtronic, Boston Scientific, and Pharmacosomos; consulting fees and payments from Boehringer Ingelheim, AstraZeneca, Novartis, Novo Nordisk, Pharmacosomos, Abbvie, Bayer, Takeda, Corvia, Cardorentis, Seimens, and Vifor; has participated in Data Safety Monitoring Boards and Advisory Boards for Teikoku and AstraZeneca and is Director of Global Clinical Trials Partners. C.P.G. has received grants for research from Abbott and Bristol Myers Squibb; consulting fees from AstraZeneca, Bayer and Daiichi Sankyo; honoraria for speaking at meetings and educational events from AstraZeneca, Wondr Medical and Menarini; support for attending meetings from Bayer and Bristol Myers Squibb; and has acted as an advisory board member for Amgen, AstraZeneca, Bayer, Daiichi Sankyo and Menarini. All other authors have nothing to disclose.

References

1. Roth GA, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, *et al.* Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018;**392**:1736–1788. https://doi.org/10.1016/S0140-6736(18)32203-7
2. 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;**391**:572–580. https://doi.org/10.1016/S0140-6736(17)32520-5
3. Conrad N, Judge A, Canoy D, Tran J, Pinho-Gomes AC, Millett ER, *et al.* Temporal trends and patterns in mortality after incident heart failure: A longitudinal analysis of 86 000 individuals. *JAMA Cardiol* 2019;**4**:1102–1111. https://doi.org/10.1001/jamacardio.2019.3593
4. Cook C, Cole G, Asaria P, Jabbour R, Francis DP. The annual global economic burden of heart failure. *Int J Cardiol* 2014;**171**:368–376. https://doi.org/10.1016/j.ijcard.2013.12.028
5. Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM, *et al.* 2022 AHA/ACC/HFSA Guideline for the management of heart failure: A report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *J Am Coll Cardiol* 2022;**79**:e263–e421. https://doi.org/10.1016/j.jacc.2021.12.012
6. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, *et al.* 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). With the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail* 2022;**24**:4–131. https://doi.org/10.1002/ejhf.2333
7. Sinha A, Gupta DK, Yancy CW, Shah SJ, Rasmussen-Torvik LJ, McNally EM, *et al.* Risk-based approach for the prediction and prevention of heart failure. *Circ Heart Fail* 2021;**14**:e007761. https://doi.org/10.1161/CIRCHEARTFAILURE.120.007761
8. Echouffo-Tcheugui JB, Greene SJ, Papadimitriou L, Zannad F, Yancy CW, Gheorghiade M, *et al.* Population risk prediction models for incident heart failure:

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

A systematic review. *Circ Heart Fail* 2015;**8**:438–447. https://doi.org/10.1161/CIRCHEARTFAILURE.114.001896
9. Sahle BW, Owen AJ, Chin KL, Reid CM. Risk prediction models for incident heart failure: A systematic review of methodology and model performance. *J Card Fail* 2017;**23**:680–687. https://doi.org/10.1016/j.cardfail.2017.03.005
10. Averbuch T, Sullivan K, Sauer A, Mamas MA, Voors AA, Gale CP, *et al.* Applications of artificial intelligence and machine learning in heart failure. *Eur Heart J Digit Health* 2022;**3**:311–322. https://doi.org/10.1093/ehjdh/ztac025
11. Withaar C, Lam CSP, Schiattarella GG, de Boer RA, Meems LMG. Heart failure with preserved ejection fraction in humans and mice: Embracing clinical complexity in mouse models. *Eur Heart J* 2021;**42**:4420–4430. https://doi.org/10.1093/eurheartj/ehab389
12. Nadarajah R, Alsaeed E, Hurdus B, Aktaa S, Hogg D, Bates MG, *et al.* Prediction of incident atrial fibrillation in community-based electronic health records: A systematic review with meta-analysis. *Heart* 2022;**108**:1020–1029. https://doi.org/10.1136/heartjnl-2021-320036
13. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan – A web and mobile app for systematic reviews. *Syst Rev* 2016;**5**:210. https://doi.org/10.1186/s13643-016-0384-4
14. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, *et al.* The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 2021;**372**:n71. https://doi.org/10.1136/bmj.n71
15. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, *et al.* Critical appraisal and data extraction for systematic reviews of prediction modelling studies: The CHARMS checklist. *PLoS Med* 2014;**11**:e1001744. https://doi.org/10.1371/journal.pmed.1001744
16. Wolff RF, Moons KG, Riley RD, Whiting PF, Westwood M, Collins GS, *et al.* PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;**170**:51–58. https://doi.org/10.7326/M18-1376
17. Debray TP, Damen JA, Snell KI, Ensor J, Hooft L, Reitsma JB, *et al.* A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017;**356**:i6460. https://doi.org/10.1136/bmj.i6460
18. Ma T, Xiao C, Wang F. Health-atm: A deep architecture for multifaceted patient health record representation and risk prediction. *Proceedings of the 2018 SIAM International Conference on Data Mining*. SIAM; 2018. p261–269.
19. Bridge J, Blakey JD, Bonnett LJ. A systematic review of methodology used in the development of prediction models for future asthma exacerbation. *BMC Med Res Methodol* 2020;**20**:22. https://doi.org/10.1186/s12874-020-0913-7
20. Debray TP, Damen JA, Riley RD, Snell K, Reitsma JB, Hooft L, *et al.* A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. *Stat Methods Med Res* 2019;**28**:2768–2786. https://doi.org/10.1177/0962280218785504
21. Himmelreich JC, Veelers L, Lucassen WA, Schnabel RB, Rienstra M, van Weert HC, *et al.* Prediction models for atrial fibrillation applicable in the community: A systematic review and meta-analysis. *Europace* 2020;**22**:684–694. https://doi.org/10.1093/europace/euaa005
22. He Y, Ong Y, Li X, Din FV, Brown E, Timofeeva M, *et al.* Performance of prediction models on survival outcomes of colorectal cancer with surgical resection: A systematic review and meta-analysis. *Surg Oncol* 2019;**29**:196–202. https://doi.org/10.1016/j.suronc.2019.05.014
23. Lloyd-Jones DM. Cardiovascular risk prediction: Basic concepts, current status, and future directions. *Circulation* 2010;**121**:1768–1777. https://doi.org/10.1161/CIRCULATIONAHA.109.849166
24. Debray T, de Jong V. metamisc: Meta-Analysis of Diagnosis and Prognosis Research Studies. https://cran.r-project.org/web/packages/metamisc/metamisc.pdf. Accessed 5 June 2023.
25. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2019.
26. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;**36**:1–48. https://doi.org/10.18637/jss.v036.i03
27. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, *et al.* GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;**336**:924–926. https://doi.org/10.1136/bmj.39489.470347.AD
28. Arshi B, van den Berge JC, van Dijk B, Deckers JW, Ikram MA, Kavousi M. Implications of the ACC/AHA risk score for prediction of heart failure: The Rotterdam study. *BMC Med* 2021;**19**:43. https://doi.org/10.1186/s12916-021-01916-7
29. Goyal A, Norton CR, Thomas TN, Davis RL, Butler J, Ashok V, *et al.* Predictors of incident heart failure in a large insured population: A one million person-year follow-up study. *Circ Heart Fail* 2010;**3**:698–705. https://doi.org/10.1161/CIRCHEARTFAILURE.110.938175
30. Tromp J, Paniagua SM, Lau ES, Allen NB, Blaha MJ, Gansevoort RT, *et al.* Age dependent associations of risk factors with heart failure: Pooled population based cohort study. *BMJ* 2021;**372**:n461. https://doi.org/10.1136/bmj.n461
31. D'Agostino RB Sr, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, *et al.* General cardiovascular risk profile for use in primary care: The Framingham Heart Study. *Circulation* 2008;**117**:743–753. https://doi.org/10.1161/CIRCULATIONAHA.107.699579
32. Khan SS, Ning H, Allen NB, Carnethon MR, Yancy CW, Shah SJ, *et al.* Development and validation of a long-term incident heart failure risk model. *Circ Res* 2022;**130**:200–209. https://doi.org/10.1161/CIRCRESAHA.121.319595
33. Khan SS, Ning H, Shah SJ, Yancy CW, Carnethon M, Berry JD, *et al.* 10-year risk equations for incident heart failure in the general population. *J Am Coll Cardiol* 2019;**73**:2388–2397. https://doi.org/10.1016/j.jacc.2019.02.057
34. Ho JE, Enserro D, Brouwers FP, Kizer JR, Shah SJ, Psaty BM, *et al.* Predicting heart failure with preserved and reduced ejection fraction: The international collaboration on heart failure subtypes. *Circ Heart Fail* 2016;**9**:e003116. https://doi.org/10.1161/CIRCHEARTFAILURE.115.003116
35. Gaziano L, Cho K, Djousse L, Schubert P, Galloway A, Ho YL, *et al.* Risk factors and prediction models for incident heart failure with reduced and preserved ejection fraction. *ESC Heart Fail* 2021;**8**:4893–4903. https://doi.org/10.1002/ehf2.13429
36. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inform Assoc* 2017;**24**:361–370. https://doi.org/10.1093/jamia/ocw112
37. Ergatoudes C, Thunström E, Hansson PO, Morales D, Mandalenakis Z, Rosengren A, *et al.* Natriuretic and inflammatory biomarkers as risk predictors of heart failure in middle-aged men from the general population: A 21-year follow-up. *J Card Fail* 2018;**24**:594–600. https://doi.org/10.1016/j.cardfail.2018.07.005
38. Kwon BC, Choi MJ, Kim JT, Choi E, Kim YB, Kwon S, *et al.* Retainvis: Visual analytics with interpretable and interactive recurrent neural networks on electronic medical records. *IEEE Trans Vis Comput Graph* 2019;**25**:299–309. https://doi.org/10.1109/TVCG.2018.2865027
39. Bavishi A, Bruce M, Ning H, Freaney PM, Glynn P, Ahmad FS, *et al.* Predictive accuracy of heart failure-specific risk equations in an electronic health record-based cohort. *Circ Heart Fail* 2020;**13**:e007462. https://doi.org/10.1161/CIRCHEARTFAILURE.120.007462
40. Khan SS, Barda N, Greenland P, Dagan N, Lloyd-Jones DM, Balicer R, *et al.* Validation of heart failure-specific risk equations in 1.3 million Israeli adults and usefulness of combining ambulatory and hospitalization data from a large integrated health care organization. *Am J Cardiol* 2022;**168**:105–109. https://doi.org/10.1016/j.amjcard.2021.12.017
41. Agarwal SK, Chambless LE, Ballantyne CM, Astor B, Bertoni AG, Chang PP, *et al.* Prediction of incident heart failure in general practice: The Atherosclerosis Risk in Communities (ARIC) study. *Circ Heart Fail* 2012;**5**:422–429. https://doi.org/10.1161/CIRCHEARTFAILURE.111.964841
42. Chahal H, Bluemke DA, Wu CO, McClelland R, Liu K, Shea SJ, *et al.* Heart failure risk prediction in the Multi-Ethnic Study of Atherosclerosis. *Heart* 2015;**101**:58–64. https://doi.org/10.1136/heartjnl-2014-305697
43. Smith JG, Newton-Cheh C, Almgren P, Struck J, Morgenthaler NG, Bergmann A, *et al.* Assessment of conventional cardiovascular risk factors and multiple biomarkers for the prediction of incident heart failure and atrial fibrillation. *J Am Coll Cardiol* 2010;**56**:1712–1719. https://doi.org/10.1016/j.jacc.2010.05.049
44. Velagaleti RS, Gona P, Larson MG, Wang TJ, Levy D, Benjamin EJ, *et al.* Multimarker approach for the prediction of heart failure incidence in the community. *Circulation* 2010;**122**:1700–1706. https://doi.org/10.1161/CIRCULATIONAHA.109.929661
45. Goya Wannamethee S, Welsh P, Whincup PH, Lennon L, Papacosta O, Sattar N. N-terminal pro brain natriuretic peptide but not copeptin improves prediction of heart failure over other routine clinical risk parameters in older men with and without cardiovascular disease: Population-based study. *Eur J Heart Fail* 2014;**16**:25–32. https://doi.org/10.1093/eurjhf/hft124
46. Willeit P, Kaptoge S, Welsh P, Butterworth AS, Chowdhury R, Spackman SA, *et al.* Natriuretic peptides and integrated risk assessment for cardiovascular disease: An individual-participant-data meta-analysis. *Lancet Diabetes Endocrinol* 2016;**4**:840–849. https://doi.org/10.1016/S2213-8587(16)30196-6
47. Nowak C, Ärnlöv J. Kidney disease biomarkers improve heart failure risk prediction in the general population. *Circ Heart Fail* 2020;**13**:e006904. https://doi.org/10.1161/CIRCHEARTFAILURE.120.006904
48. Van Remoortel H, Scheers H, De Buck E, Haenen W, Vandekerckhove P. Prediction modelling studies for medical usage rates in mass gatherings: A systematic review. *PLoS One* 2020;**15**:e0234977. https://doi.org/10.1371/journal.pone.0234977
49. Hammond MM, Everitt IK, Khan SS. New strategies and therapies for the prevention of heart failure in high-risk patients. *Clin Cardiol* 2022;**45**:S13–S25. https://doi.org/10.1002/clc.23839
50. Kannel WB, D'Agostino RB, Silbershatz H, Belanger AJ, Wilson PW, Levy D. Profile for estimating risk of heart failure. *Arch Intern Med* 1999;**159**:1197–1204. https://doi.org/10.1001/archinte.159.11.1197

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

51. Butler J, Kalogeropoulos A, Georgiopoulou V, Belue R, Rodondi N, Garcia M, *et al.* Incident heart failure prediction in the elderly: The Health ABC Heart Failure score. *Circ Heart Fail* 2008;**1**:125–133. https://doi.org/10.1161/CIRCHEARTFAILURE.108.768457
52. Himmelreich JC, Lucassen WA, Harskamp RE, Aussems C, van Weert HC, Nielen MMJ. CHARGE-AF in a national routine primary care electronic health records database in the Netherlands: Validation for 5-year risk of atrial fibrillation and implications for patient selection in atrial fibrillation screening. *Open Heart* 2021;**8**:e001459. https://doi.org/10.1136/openhrt-2020-001459
53. Szymanski T, Ashton R, Sekelj S, Petrungaro B, Pollock KG, Sandler B, *et al.* Budget impact analysis of a machine learning algorithm to predict high risk of atrial fibrillation among primary care patients. *Europace* 2022;**24**:1240–1247. https://doi.org/10.1093/europace/euac016
54. Thomopoulos C, Parati G, Zanchetti A. Effects of blood pressure-lowering treatment. 6. Prevention of heart failure and new-onset heart failure – Meta-analyses of randomized trials. *J Hypertens* 2016;**34**:373–384. https://doi.org/10.1097/HJH.0000000000000848
55. Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, *et al.*; DECLARE-TIMI 58 Investigators. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. *N Engl J Med* 2019;**380**:347–357. https://doi.org/10.1056/NEJMoa1812389
56. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, *et al.*; EMPA-REG OUTCOME Investigators. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med* 2015;**373**:2117–2128. https://doi.org/10.1056/NEJMoa1504720
57. Ledwidge M, Gallagher J, Conlon C, Tallon E, O'Connell E, Dawkins I, *et al.* Natriuretic peptide-based screening and collaborative care for heart failure: The STOP-HF randomized trial. *JAMA* 2013;**310**:66–74. https://doi.org/10.1001/jama.2013.7588
58. Ledwidge MT, O'Connell E, Gallagher J, Tilson L, James S, Voon V, *et al.* Cost-effectiveness of natriuretic peptide-based screening and collaborative care: A report from the STOP-HF (St Vincent's Screening TO Prevent Heart Failure) study. *Eur J Heart Fail* 2015;**17**:672–679. https://doi.org/10.1002/ejhf.286
59. National Institute for Health and Care Excellence. Cardiovascular disease: risk assessment and reduction, including lipid modification. Clinical guideline [CG181] 18 July 2014. https://www.nice.org.uk/guidance/cg181/chapter/Recommendations#identifying-and-assessing-cardiovascular-disease-risk-for-people-without-established-cardiovascular. Accessed 24 May 2023
60. Lam CS, Arnott C, Beale AL, Chandramouli C, Hilfiker-Kleiner D, Kaye DM, *et al.* Sex differences in heart failure. *Eur Heart J* 2019;**40**:3859–3868c. https://doi.org/10.1093/eurheartj/ehz835
61. Piña IL, Jimenez S, Lewis EF, Morris AA, Onwuanyi A, Tam E, *et al.* Race and ethnicity in heart failure: JACC focus seminar 8/9. *J Am Coll Cardiol* 2021;**78**:2589–2598. https://doi.org/10.1016/j.jacc.2021.06.058
62. Barker J, Li X, Khavandi S, Koeckerling D, Mavilakandy A, Pepper C, *et al.* Machine learning in sudden cardiac death risk prediction: A systematic review. *Europace* 2022;**24**:1777–1787. https://doi.org/10.1093/europace/euac135
63. Collins GS, Dhiman P, Navarro CLA, Ma J, Hooft L, Reitsma JB, *et al.* Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open* 2021;**11**:e048008. https://doi.org/10.1136/bmjopen-2020-048008
64. Lee MP, Glynn RJ, Schneeweiss S, Lin KJ, Patorno E, Barberio J, *et al.* Risk factors for heart failure with preserved or reduced ejection fraction among Medicare beneficiaries: Application of competing risks analysis and gradient boosted model. *Clin Epidemiol* 2020;**12**:607–616. https://doi.org/10.2147/CLEP.S253612
65. Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, *et al.* The effect of English-language restriction on systematic review-based meta-analyses: A systematic review of empirical studies. *Int J Technol Assess Health Care* 2012;**28**:138–144. https://doi.org/10.1017/S0266462312000086
66. Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Individual patient-versus group-level data meta-regressions for the investigation of treatment effect modifiers: Ecological bias rears its ugly head. *Stat Med* 2002;**21**:371–387. https://doi.org/10.1002/sim.1023
67. Nijman S, Leeuwenberg A, Beekers I, Verkouter I, Jacobs J, Bots M, *et al.* Missing data is poorly handled and reported in prediction model studies using machine learning: A literature review. *J Clin Epidemiol* 2022;**142**:218–229. https://doi.org/10.1016/j.jclinepi.2021.11.023
68. Tsvetanova A, Sperrin M, Peek N, Buchan I, Hyland S, Martin GP. Missing data was handled inconsistently in UK prediction models: A review of method used. *J Clin Epidemiol* 2021;**140**:149–158. https://doi.org/10.1016/j.jclinepi.2021.09.008
69. Bennis FC, Hoogendoorn M, Aussems C, Korevaar JC. Prediction of heart failure 1 year before diagnosis in general practitioner patients using machine learning algorithms: a retrospective case–control study. *BMJ Open* 2022;**12**:e060458. https://doi.org/10.1136/bmjopen-2021-060458
70. Brouwers FP, van Gilst WH, Damman K, van den Berg MP, Gansevoort RT, Bakker SJ, *et al. Clinical risk stratification optimizes value of biomarkers to predict new-onset heart failure in a community-based cohort*. Vol **7**. Circulation: Heart Failure; 2014. p723–731.
71. Che Z, Cheng Y, Sun Z, Liu Y. Exploiting convolutional neural network for risk prediction with medical feature embedding. *arXiv preprint arXiv* 2017;**1701**:07474.
72. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *Journal of the American Medical Informatics Association* 2016a;**24**:361–370. https://doi.org/10.1093/jamia/ocw112
73. Choi E, Bahadori MT, Song L, Stewart WF, Sun J. GRAM: graph-based attention model for healthcare representation learning. *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*. New York, NY: Association for Computing Machinery; 2017. p787–795.
74. Choi E, Bahadori MT, Kulas JA, Schuetz A, Stewart WF, Sun J. RETAIN: An Interpretable Predictive Model for Healthcare using Reverse Time Attention Mechanism (Version 4). *arXiv* 2016b. https://doi.org/10.48550/ARXIV.1608.05745
75. Jin B, Che C, Liu Z, Zhang S, Yin X, Wei X. Predicting the Risk of Heart Failure With EHR Sequential Data Modeling. *IEEE Access* 2018;**6**:9256–9261. https://doi.org/10.1109/access.2017.2789324
76. Ju R, Zhou P, Wen S, Wei W, Xue Y, Huang X, *et al.* 3D-CNN-SPP: A Patient Risk Prediction System From Electronic Health Records via 3D CNN and Spatial Pyramid Pooling. *IEEE Transactions on Emerging Topics in Computational Intelligence* 2021;**5**:247–261. https://doi.org/10.1109/tetci.2019.2960474
77. Li Y, Salimi-Khorshidi G, Rao S, Canoy D, Hassaine A, Lukasiewicz T, *et al.* Validation of risk prediction models applied to longitudinal electronic health record data for the prediction of major cardiovascular events in the presence of data shifts. *European Heart Journal-Digital Health* 2022;**3**:535–547. https://doi.org/10.1093/ehjdh/ztac061
78. Luo J, Ye M, Xiao C, Ma F. *HiTANet: Hierarchical Time-Aware Attention Networks for Risk Prediction on Electronic Health Records*. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York NY, United States: Association for Computing Machinery; 2020. https://doi.org/10.1145/3394486.3403107
79. Ma F, You Q, Xiao H, Chitta R, Zhou J, Gao J. Kame: Knowledge-based attention model for diagnosis prediction in healthcare. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. New York NY: Association for Computing Machinery; 2018. p743–752.
80. Rao S, Li Y, Ramakrishnan R, Hassaine A, Canoy D, Cleland J, *et al.* An Explainable Transformer-Based Deep Learning Model for the Prediction of Incident Heart Failure. *IEEE Journal of Biomedical and Health Informatics* 2022;**26**:3362–3372. https://doi.org/10.1109/jbhi.2022.3148820
81. Rasmy L, Wu Y, Wang N, Geng X, Zheng WJ, Wang F, *et al.* A study of generalizability of recurrent neural network-based predictive models for heart failure onset risk using a large and heterogeneous EHR data set. *Journal of Biomedical Informatics* 2018;**84**:11–16. https://doi.org/10.1016/j.jbi.2018.06.011
82. Stenemo M, Nowak C, Byberg L, Sundström J, Giedraitis V, Lind L, *et al.* Circulating proteins as predictors of incident heart failure in the elderly. *European Journal of Heart Failure* 2018;**20**:55–62.
83. Yin C, Zhao R, Qian B, Lv X, Zhang P. Domain knowledge guided deep learning with electronic health records. *In: 2019 IEEE International Conference on Data Mining (ICDM)*. IEEE; 2019. p738–747.
84. Zhang X, Qian B, Cao S, Li Y, Chen H, Zheng Y, *et al.* INPREM: An interpretable and trustworthy predictive model for healthcare. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York, NY: Association for Computing Machinery; 2020. p450–460.

© 2023 The Authors. *European Journal of Heart Failure* published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.

18790844, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ejhf.2970 by Test, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License