BMJ Open

Prediction models for hospital readmissions in patients with heart disease: a systematic review and meta-analysis

Bastiaan Van Grootven ,[1,2] Patricia Jepma ,[3] Corinne Rijpkema,[4] Lotte Verweij,[3] Mariska Leeflang,[5] Joost Daams,[6] Mieke Deschodt,[7,8] Koen Milisen,[7,9] Johan Flamaing,[10,11] Bianca Buurman[3,5]

To cite: Van Grootven B, Jepma P, Rijpkema C, *et al*. Prediction models for hospital readmissions in patients with heart disease: a systematic review and meta-analysis. *BMJ Open* 2021;**11**:e047576. doi:10.1136/bmjopen-2020-047576

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online. (http://dx.doi.org/10.1136/bmjopen-2020-047576).

BVG and PJ are joint first authors.

Received 02 December 2020
Accepted 30 July 2021

For numbered affiliations see end of article.

Correspondence to
Dr Bastiaan Van Grootven; bastiaan.vangrootven@kuleuven.be

ABSTRACT

Objective To describe the discrimination and calibration of clinical prediction models, identify characteristics that contribute to better predictions and investigate predictors that are associated with unplanned hospital readmissions.

Design Systematic review and meta-analysis.

Data source Medline, EMBASE, ICTPR (for study protocols) and Web of Science (for conference proceedings) were searched up to 25 August 2020.

Eligibility criteria for selecting studies Studies were eligible if they reported on (1) hospitalised adult patients with acute heart disease; (2) a clinical presentation of prediction models with c-statistic; (3) unplanned hospital readmission within 6 months.

Primary and secondary outcome measures Model discrimination for unplanned hospital readmission within 6 months measured using concordance (c) statistics and model calibration. Meta-regression and subgroup analyses were performed to investigate predefined sources of heterogeneity. Outcome measures from models reported in multiple independent cohorts and similarly defined risk predictors were pooled.

Results Sixty studies describing 81 models were included: 43 models were newly developed, and 38 were externally validated. Included populations were mainly patients with heart failure (HF) (n=29). The average age ranged between 56.5 and 84 years. The incidence of readmission ranged from 3% to 43%. Risk of bias (RoB) was high in almost all studies. The c-statistic was <0.7 in 72 models, between 0.7 and 0.8 in 16 models and >0.8 in 5 models. The study population, data source and number of predictors were significant moderators for the discrimination. Calibration was reported for 27 models. Only the GRACE (Global Registration of Acute Coronary Events) score had adequate discrimination in independent cohorts (0.78, 95% CI 0.63 to 0.86). Eighteen predictors were pooled.

Conclusion Some promising models require updating and validation before use in clinical practice. The lack of independent validation studies, high RoB and low consistency in measured predictors limit their applicability.

PROSPERO registration number CRD42020159839.

Strengths and limitations of this study

- ► Largest investigation of unplanned hospital readmission risk to date, including 81 unique prediction models in the systematic review.
- ► Independent and standardised procedures for study selection, data collection and risk of bias (RoB) assessment.
- ► High RoB in current prediction models and unexplained heterogeneity between models limit recommendations for using prediction model in clinical practice.

INTRODUCTION

Hospital readmissions in patients with acute heart disease are associated with a high burden on patients, healthcare and costs.[1] The identification of high-risk hospitalised patients is important to provide timely interventions. Prediction models guide healthcare providers in daily practice to assess patients' probability of readmission within a certain time frame and include candidate variables identified by clinical perspectives, literature or data-driven approaches, for example, using machine learning techniques.[2] Data are often collected from observational cohorts of intervention studies and subsequently analysed to examine what set of predictors best predict the risk of readmission. The clinical applicability of risk prediction models in daily practice is currently limited. Statistical models are often not presented in a clinically useful way or models based on administrative data are considered.[3] These models therefore cannot be readily used in daily practice. In addition, prediction models are often developed for a very specific population, which asks from clinicians to be familiar with several models. Furthermore, patients may belong to multiple populations because of cardiac

comorbidities. Numerous systematic reviews have previously investigated the prediction of unplanned hospital readmissions in several populations.[3–12] While some have included hospitalised patients in general,[11 12] others have focused specifically on patients with heart failure (HF)[4–8 10] or acute myocardial infarction (AMI).[3 9] The conclusion is generally the same, the discrimination is poor to adequate, and there is little consistency in the type of predictors included in the models.

We believe that the state of the art on risk prediction can be improved if more knowledge is available on the performance of clinical risk prediction models and risk predictors across different populations of patients with heart disease. Although heterogeneity in models and predictors is often considered as a limitation, it can inform effect moderators on how predictions can be improved.[13] For example, perhaps we can identify predictors who demonstrate a consistent association with hospital readmission regardless of the underlying disease. If this can be identified, a more general prediction model could be developed that is relevant for the heterogeneous group of patients on cardiac care units. This might contribute to the early recognition and onset of preventive interventions in patients with heart disease at risk of readmission.

We therefore performed a systematic review and meta-analysis on clinical risk prediction models for the outcome unplanned hospital readmission in patients hospitalised for acute heart disease. Our aims were to describe the discrimination and calibration of clinical prediction models, to identify characteristics that contribute to better predictions, and to investigate predictors that are consistently associated with hospital readmissions.

METHODS

A protocol was registered in PROSPERO (registration number: CRD42020159839). The results are reported following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement.[14]

Eligibility criteria

Studies were eligible if (1) the study population included hospitalised adult patients with (symptoms of) heart disease; (2) a prediction model with c-statistic was reported; (3) a clinically useful presentation of the model with risk factors was reported; (4) the outcome was unplanned hospital readmissions within 6 months; (5) the study design was appropriate, that is, (nested) case–control study (prospective and retrospective) cohort study, database and registry study, or secondary analysis of a trial; (6) they were reported in English.

Information sources

A search strategy was designed with an information specialist (PROSPERO protocol and online supplemental text 1). We searched the Medline, EMBASE, WHO ICTPR search portal (for study protocols) and Web of Science (for conference proceedings) databases up to 25 August 2020 without any restrictions for eligible studies. We searched for full-text manuscripts of the identified protocols. After selecting the full-text manuscripts, we screened references lists and prospective citations (using Google Scholar) for additional eligible studies.

Study selection

Three reviewers were involved in the study selection process. Each reviewer independently screened two-thirds of the titles, abstracts and full-text articles of potentially relevant references identified in the literature search. Disagreements were resolved through consensus. Sixteen authors were contacted and six delivered data for readmission when a composite outcome was used. Two authors were also contacted when data were reported combining multiple patient populations. However, no additional data were provided for the population with heart disease and these studies were excluded.

Data extraction

Data extraction was performed based on the 'Critical Appraisal and Data Extraction for Systematic Reviews' of prediction modelling studies checklist using standardised forms in the Distiller Systematic Review Software (see online supplemental text 2 for the data items).[15] The checklist includes items on 11 relevant domains, including source of data, participants, outcomes, candidate predictors, sample size, missing data, model development, model performance, model evaluation, results and interpretation. One reviewer collected the data and the second reviewer verified the extracted data. Disagreements were resolved through consensus. Eight authors were contacted and two delivered data to resolve uncertainties or missing data.

Risk of bias

The Prediction model Risk Of Bias ASsessment Tool (PROBAST) tool[16] was used to assess the risk of bias (RoB) for four 'quality' domains, that is, the participants, predictors, outcome and analysis for each model. One author assessed the RoB as low, high or unclear, and the second author verified the extracted data and RoB conclusion. Disagreements were resolved through consensus. In addition, the applicability of the included studies based on our research question was assessed for three domains, that is, participants, predictors and outcome domains and rated as low concerns, high concerns or uncertain concerns regarding applicability.

Summary measures

The discrimination of the prediction models was described using the concordance (c)-statistic. Missing SEs were derived from the sample data.[17] The calibration was described using the number of observed and expected events, the calibration slope, calibration in large or the Hosmer-Lemeshow test. A definition of the commonly used measures is described in box 1.

The association between risk predictors and hospital readmission was described using regression coefficients.

Box 1 Definitions of commonly used measures

Discrimination:
Refers to the ability of a prediction model to discriminate between a patient with and without the outcome, for example, readmission.

C-statistic:
Is a measure of discrimination. For binary outcomes, the c-statistic is equivalent to the area under the curve: 1 indicates perfect discrimination, and 0.5 indicates that the models does not perform better than chance. Harrell's c-statistic is often used in survival models.

Calibration:
Refers to the agreement between the predicted and the observed probability (or the outcome value for linear models). Calibration is expressed using different measures, for example, calibration slope, calibration in large, Hosmer-Lemeshow test.

Calibration slope:
The slope should be 1, a value <1 indicates extreme predictions, and a value of >1 indicates to moderate predictions.

Calibration in large:
The value should be 0, a negative value indicates overestimation of the prediction, and a positive value indicates underestimation of the prediction.

Hosmer-Lemeshow test:
This is a goodness-of-fit test for binary outcomes. A significant p value, usually <0.05, indicates poor goodness-of-fit.

Derivation/development cohort:
A cohort of patients that is used to estimate the predictor values that are used in a prediction model to estimate a patient's probability for an outcome.

Validation cohort:
A cohort of patients that is used to evaluate how well the developed model performs (in terms of discrimination and calibration).

Internal validation:
Estimates how well the performance of a model will be reproduced in the target population. Several techniques can be used, for example, random-split sample, cross-validation and bootstrapping techniques.

External validation:
Evaluates how well a model performs in a new sample and can consist of temporal validation (sample contains more recently treated patients), geographical validation (sample is from a different centre) of a fully independent validation (validation by an independent team).

Missing SEs for the coefficients were considered missing completely at random and were not imputed. A complete case analysis was performed.

Synthesis of results and analyses

Meta-analyses using random-effects models, with the Hartung-Knapp modification, were performed to describe the distribution of the between-study variance of the different prediction models and their predictors. Because we considered that there would be substantial heterogeneity, conclusions were not based on the precision of the pooled estimates.

The c-statistic from each model was pooled and a meta-regression was performed to investigate the moderation effect of age and the number of predictors on the discrimination. A subgroup analysis was performed to investigate the moderation effect of the different patient populations, design, outcome definition and endpoint. The c-statistic of the validated model was used if available; otherwise, the c-statistic from the development phase was used.

The c-statistics of specific prediction models that were evaluated in multiple studies were pooled for the endpoint 30-day follow-up.

Coefficients of predictors that were similarly defined in at least five studies were pooled for the endpoint 30-day follow-up. The patient populations were defined as subgroups to explore consistency and heterogeneity (I^2, tau) in the effect estimates.

Analyses were performed using the 'metan' package in STATA V.15 IC and the 'metamisc package' in Rstudio.

Public and patient involvement

Because of the design of the study and because we did not collect primary date, we did not involve patients or the public in the design, conduct or reporting of our research.

RESULTS

A total of 8588 abstracts were reviewed and 60 studies describing 81 separate models were included (figure 1). Table 1 provides an overview of the included studies and models, which were published between 2001 and 2020. The majority of the studies (n=40) was performed in the USA. The data sources used were mostly retrospective cohort studies (n=15), hospital databases (n=13) and registries (n=13). Included populations were mainly patients with HF(n=29), surgical patients (n=14) and patients with an AMI or acute coronary syndrome (n=10). The average age was between 56.5 and 84 years. The sample size of development cohorts ranged from 182 to 193 899 patients and of the validation cohorts between 104 and 321 088 patients. The outcome of interest was mostly all-cause readmission (n=41) and measured on 30 days (n=55). The incidence of readmission per study ranged from 3% to 43%.

Risk of bias

Figure 2 summarises the RoB and applicability assessment (online supplemental table 1A). The overall RoB was high in 98.9% of the models and only one study[18] showed low RoB in all four domains.

For the domain participants, 82.4% of studies was assessed as high RoB because most studies performed retrospective data analyses or used data from existing sources with large number of candidate predictors that were originally developed for other purposes, for example, administrative databases or registries. The domain predictors were assessed as high RoB in 27.5% of the models, 24.2% as low RoB and 48.4% as unclear RoB. For the domain outcome, 41.8%, 34.1% and 24.2% were assessed as high, low and unclear RoB, respectively.

The domain analysis was assessed as high RoB in 97.8%. Most studies did not use appropriate statistics for the development or validation of prediction models. For example, a

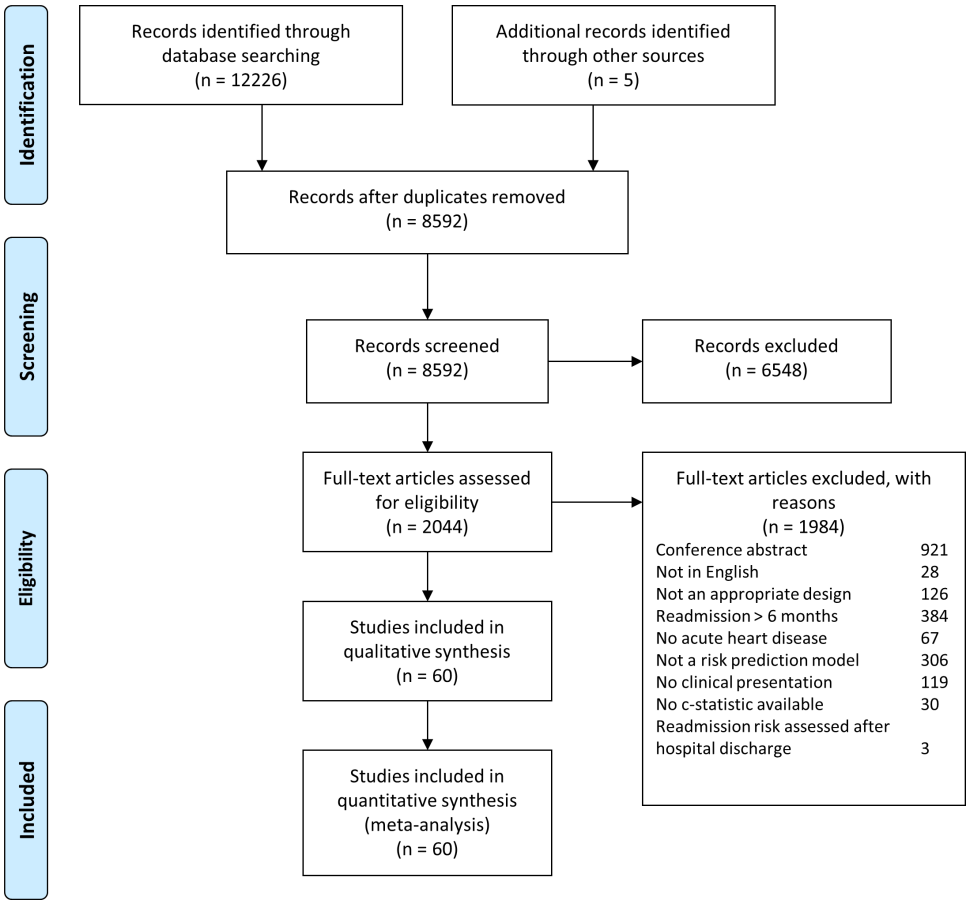

Figure 1 Flowchart. In total, 8592 records were screened and 60 studies with 81 prediction models were included.

description on how complexities in data were handled (eg, competing risk of death) was often missing and relevant performance measures were incomplete (eg, calibration).

The domain participants and predictors were assessed as low concerns regarding applicability in all studies. For the domain outcome, 70.3% of studies used all-cause readmission as the outcome of interest and were therefore assessed as low concerns regarding applicability.

Prediction models

A total of 43 new models were developed for patients with HF (n=15), undergoing surgical procedures (n=12), AMI (n=9), transcatheter aortic valve replacement (TAVR) (n=2), a mixed sample with HF and coronary syndromes (n=2), arrhythmias (n=1), valvular disease (n=1), while one study did not specify the sample (table 1). The c-statistic was lower than 0.6 in 5 models, between 0.6 and 0.7 in 24 models, between 0.7 and 0.8 in 6 models, and between 0.8 and 0.9 in 2 models. In six models, the c-statistic was only reported for a validation cohort (table 2).

A total of 38 separate models were externally validated for patients with HF (n=26), AMI (n=4), surgical patients (n=3), acute coronary syndrome (n=2), arrhythmias (n=2), mixed sample with HF and coronary syndromes (n=1). The discrimination was lower than 0.6 in 16 models, between 0.6 and 0.7 in 15 models, between 0.7 and 0.8 in 5 models, and between 0.8 and 0.9 in 2 models (table 2).

The discrimination of six models was evaluated in multiple independent cohorts and was pooled in meta-analyses (figure 3, online supplemental figures 1–6): the CMS AMI (Centers for Medicare and Medicaid Services Acute Myocardial Infarction) administrative model[19 20] (0.65, 95% CI 0.56 to 0.73); the CMS HF (Heart Failure) administrative model[21–29] (0.60, 95% CI 0.58 to 0.62); the CMS HF medical model[24 27 30–32] (0.60, 95% CI 0.58 to 0.62); the HOSPITAL (Hemoglobin level, discharged from Oncology, Sodium level, Procedure during admission, Index admission Type, Admission, Length of stay) score[33–35] (0.64, 95% CI 0.58 to 0.70); the GRACE (Global Registration of Acute Coronary Events) score[36 37] (0.78, 95% CI 0.63 to 0.86); and the LACE (Length of stay, acuity of the Admission, Comorbidity of the patient and Emergency department use in the duration of 6 months before admission) score[23 28 29 34 38] (0.62, 95% CI 0.53 to 0.70).

On average, models for patients with AMI had the best discrimination (0.67, n=16), followed by patients with TAVR (0.65, n=2), patients with HF (0.64, n=45) and surgical patients (0.63, n=17). The discrimination was highest in studies using secondary analysis (0.70, n=2) and retrospective cohort studies (0.69, n=23), and was lowest

Table 1 Study characteristics

Study	Model	Data source	Development	Validation	Sample size		Population	Average age	Outcome	Readmission (%)	
					Dev	Val				Dev	Val
Moretti *et al*[57]	EuroHeart PCI score	Hospital database	NA	Ext	–	1192	ACS	71 (7)	30d		4.7
Asche *et al*[46]	NR	Retrospective cohort	Yes	Split	2446	612	AMI	65 (15)	30d	8.9	
Cediel *et al*[58]	TARRACO Risk score	Retrospective cohort	Yes	No	611	401	AMI type 2, ischaemia	D: 78 (17) V: 60 (21)	30d	2.6	
		Retrospective cohort	Yes	No	611	401	AMI type 2, ischaemia	D: 78 (17) V: 60 (21)	180d	7.9	
Chotechuang *et al*[36]	GRACE	Retrospective cohort	NA	Ext	–	152	AMI	60.5 (6.3)	30d		5.3
	GRACE	Retrospective cohort	NA	Ext	–	152	AMI	60.5 (6.3)	180d		9.2
Hilbert *et al*[59]	AMI decision tree	Registry	Yes	Ext	10848	10701	AMI	NR	30d	20.6	19.7
Dodson *et al*[18]	SILVER-AMI 30-day readmission calculator	Prospective cohort	Yes	Split	2004	1002	AMI	81.5 (5.0)	30d	18.2	
Kini *et al*[60]	NR	Registry	Yes	Split	60742	26107	AMI	76.5 (8.0)	90d	27.5	
Nguyen *et al*[19]	AMI READMITS score	Retrospective cohort	Yes	Split	661	165	AMI	65.5 (12.8)	30d	13	
	Full-stay AMI model	Retrospective cohort	Yes	Split	661	165	AMI	65.5 (12.8)	30d	13	
	CMS AMI administrative model	Retrospective cohort	NA	Ext	–	826	AMI	65.5 (12.8)	30d		13
Krumholz *et al*[20]	CMS AMI administrative model	Registry	Yes	Split, Ext	100 465	321 088	AMI	78.7 (8.0)	30d	18.9	20.0 (Ext) NR (split)
	CMS AMI medical model	Registry	Yes	Split	130944	130944	AMI	76.2 (7.3)	30d	20	
Rana *et al*[33]	Elixhauser index	Hospital database	NA	Ext	–	1660	AMI	67.9	30d		6.3
	HOSPITAL score	Hospital database	NA	Ext	–	1660	AMI	67.9	30d		6.3
Atzema *et al*[47]	AFTER Part 2 scoring system	Retrospective cohort	Yes	Split	2343	1167	Arrhythmia, AF	D: 68.6 (14.7) V: 68.3 (15.1)	30d	7	7.6
Lahewala *et al*[40]	CHADS2	Administrative	NA	Ext	–	116 450	Arrhythmia, AF	<75	30d		15.8
	CHADS2	Administrative	NA	Ext	–	116 450	Arrhythmia, AF	<75	90d		25.1
	CHA2DS-VASc	Administrative	NA	Ext	–	116 450	Arrhythmia, AF	65–74	30d		15.8
	CHA2DS-VASc	Administrative	NA	Ext	–	116450	Arrhythmia, AF	65–74	90d		25.1
Benuzillo *et al*[61]	CRSS	Hospital database	Yes	Boot, Ext	2589	896 (Ext) 500 (Boot)	CABG	66.7 (9.9)	30d	9.1	8.2 (Ext) 9.1 (Boot)
Deo *et al*[62]	30-day CABG readmission calculator	Administrative	Yes	Boot	155054	1000	CABG	65.4 (10.4)	30d	12.5	
Engoren *et al*[65]	NR	Hospital database	Yes	Split	2644	2711	CABG	NR	30d	7.6	8
Lancey *et al*[63]	NR	Registry	Yes	Split	2341	2520	CABG	64.5 (10.5)	30d	8.8	9.5
Rosenblum *et al*[41]	The STS PROM score	Hospital database	NA	Ext	–	21719	CABG	63.5 (10.7)	30d		9.3
Zitser-Gurevich *et al*[64]	NR	Prospective cohort	Yes	Split	2266.5	2266.5	CABG	65–74	30d	13.3	
	NR	Prospective cohort	Yes	Split	2266.5	2266.5	CABG	65–74	100d	24.1	

Continued

Table 1 Continued

Study	Model	Data source	Development	Validation	Sample size		Population	Average age	Outcome	Readmission (%)	
					Dev	Val				Dev	Val
Zywot *et al*[42]	CABG risk scale	Administrative	Yes	Ext	126519	94318	CABG	D: 70–74 V: 70–74	30d	23	21
Ahmad *et al*[21]	CMS HF administrative model	Prospective cohort	NA	Ext	–	183	HF	61 (18)	30d		22.4
Amarasingham *et al*[22]	ADHERE	Hospital database	NA	Ext	–	1372	HF	56.5	30d		24.1
	CMS HF administrative model	Hospital database	NA	Ext	–	1372	HF	56.5	30d		24.1
	Tabak mortality score	Hospital database	NA	Ext	–	1372	HF	56.5	30d		24.1
Au *et al*[23]	Administrative claims model: HF 30-day mortality	Administrative	NA	Ext	59652	59652	HF	75.8 (12.7)	30d		15.9
	Charlson Comorbidity Score	Administrative	NA	Ext	59652	59652	HF	75.8 (12.7)	30d		15.9
	CMS HF administrative model	Administrative	NA	Ext	59652	59652	HF	75.8 (12.7)	30d		15.9
	LACE	Administrative	NA	Ext	59652	59652	HF	75.8 (12.7)	30d		15.9
Bardhan *et al*[65]	NR	Hospital database	Yes	No	40983	–	HF	69.2 (15.7)	30d	7	
Betihavas *et al*[66]	NR	RCT secondary analysis	Yes	Boot	280	200	HF	74 (64–81)	28d	18	
Cox *et al*[24]	CMS HF administrative model	Hospital database	No	Ext	–	1454	HF	75 (12)	30d		21.5
	CMS HF medical model	Hospital database	No	Ext	–	1454	HF	75 (12)	30d		21.5
Delgado *et al*[67]	15-day CV readmission risk score	Prospective cohort	Yes	Boot	1831	500	HF	72.4 (12.1)	15d	7.1	
	30-day CV readmission risk score	Prospective cohort	Yes	Boot	1831	500	HF	72.4 (12.1)	30d	13.9	
Formiga *et al*[30]	CMS HF medical model	Hospital database	NA	Ext	–	719	HF	78.1 (9)	30d		7.6
	CMS HF medical model	Hospital database	NA	Ext	–	719	HF	78.1 (9)	90d		14.4
Frizzell *et al*[25]	CMS HF administrative model	Registry	NA	External	–	56477	HF	80 (2)	30d		21.2
Hammill *et al*[26]	CMS HF administrative model	Registry	NA	Ext	–	24163	HF	81	30d		21.9
Hilbert *et al*[59]	HF decision tree	Registry	Yes	Ext	39682	38409	HF	NR	30d	25.5	25.2
Hummel *et al*[31]	CMS HF medical model	Prospective cohort	NA	Ext	–	1807	HF	79.8 (7.6)	30d		27
Huynh *et al*[48]	NR	Prospective cohort	Yes	Ext	430	1046	HF	D: 75 (19) V: 67 (17)	30d	21	24
	NR	Prospective cohort	Yes	Ext	430	1046	HF	D: 75 (19) V: 67 (17)	90d	43	42
Ibrahim *et al*[34]	HOSPITAL score	Retrospective cohort	NA	Ext	–	692	HfpEF	68.3 (11.8)	30d		27.3

Continued

Table 1 Continued

Study	Model	Data source	Development	Validation	Sample size		Population	Average age	Outcome	Readmission (%)	
					Dev	Val				Dev	Val
	LACE/LACE + index	Retrospective cohort	NA	Ext	–	692	HfpEF	68.3 (11.8)	30d		27.3
Keenan *et al*[27]	CMS HF administrative model	Registry	Yes	Split, Ext.	28319	845291	HF	79.9 (7.8)	30d	23.6	23.7 (Ext) NR (Split)
	CMS HF medical model	Registry	Yes	Split, Ext.	64329	64329	HF	75–84	30d	23.7	
Kitamura *et al*[53]	FIM	Retrospective cohort	NA	Ext	–	113	HF	80.5 (6.7)	90d		20.4
Leong *et al*[68]	30-day HF readmission risk score	Retrospective cohort	Yes	Split	888	587	HF	D: 70.0 (12.7) V: 69.1 (12.8)	30d	9.9	
Li *et al*[49]	NR	Retrospective cohort	Yes	Split	51783	25887	HF	D: 84 (12) V: 84 (11)	30d	24.2	
Lim *et al*[69]	NR	Registry	Yes	No	4566	–	HF	70.5 (12.0)	30d	6.6 (car) 13 (all)	
Reed *et al*[28]	AH model	Administrative	Yes	Split	NR	NR	HF	NR	30d	NR	
	CMS HF administrative model	Administrative	NA	Split	–	NR	HF	NR	30d		NR
	Hasan	Administrative	NA	Split	–	NR	HF	NR	30d		NR
	LACE	Administrative	NA	Split	–	NR	HF	NR	30d		NR
	PARR-30	Administrative	NA	Split	–	NR	HF	NR	30d		NR
Salah *et al*[70]	ELAN-HF score	Prospective cohort secondary analysis	Yes	No	1301	–	HF	74 (16)	180d	36.1	
Sudhakar *et al*[32]	CMS HF medical model	Hospital database	NA	Ext	–	1046	HF	65.2 (16.6)	30d		35.3
Tan *et al*[71]	NR	Hospital database	Yes	Split	246	104	HF	D: 67.7 (12.3) V: 69.0 (12.9)	90d	24.5	11.7
Wang *et al*[72]	NR	Hospital database	Yes	No	4548	–	HF	68.5 (27.6)	30d	25.1	
Wang *et al*[38]	LACE	Retrospective cohort	NA	Ext	–	253	HF	56.6 (11.5)	30d		24.5
Yazdan-Ashoori *et al*[29]	CMS HF administrative model	Prospective cohort	NA	Ext	–	378	HF	73.1 (13.1)	30d		26
	LACE	Prospective cohort	NA	Ext	–	378	HF	73.1 (13.1)	30d		26
Disdier Moulder *et al*[73]	NR	Prospective cohort	Yes	No	258		HF, ACS, NR	70.5 (23)	30d	17	
	NR	Prospective cohort	Yes	No	258		HF, ACS, NR	70.5 (23)	180d	38	
Raposeiras-Roubín *et al*[37]	GRACE	Retrospective cohort	NA	Ext	–	4229	HF, ACS	68.2 (18.7)	30d		2.6
Burke *et al*[35]	HOSPITAL score	Retrospective cohort	NA	Ext	–	HF: 3189 AMI: 767	HF, AMI	65.8 (16.8)	30d		HF: 18.2 AMI: 17.4
Minges *et al*[74]	NR	Registry	Yes	Split	193899	194179	HF, PCI	65+	30d	11.4	
Pack *et al*[75]	NR	Administrative	Yes	Split	30826	7706	HVD	64.9 (12.2)	90d	12.8	
Oliver-McNeil *et al*[76]	ICD readmission-risk score	Registry	Update	Ext	182	–	ICD	69 (11)	30d		17.6
Wasfy *et al*[52]	Pre-PCI model	Registry	Yes	Split	24052	12008	NR	64.8 (12.5)	30d	10.4	

Continued

Table 1 Continued

Study	Model	Data source	Development	Validation	Sample size		Population	Average age	Outcome	Readmission (%)	
					Dev	Val				Dev	Val
Barnett *et al*[77]	NR	Registry	Update	Ext	19964	19964	Surgical	65.3 (12.4)	30d	11.4	
Brown *et al*[43]	STS augmented clinical model	Prospective cohort	Update	Boot	1046	NR	Surgical	65.4 (9.8)	30d	NR	
	STS 30-day readmission model	Prospective cohort	NA	Ext	–	1194	Surgical	73.3 (10.1)	30d		NR
Espinoza *et al*[78]	30-day readmission score after cardiac surgery	Retrospective cohort	Yes	Split	2529	2567	Surgical	65.1 (11.5)	30d	11.9	
Ferraris *et al*[54]	READMIT	Prospective cohort	Yes		2574		Surgical	63 (11)	30d	9.8	
Kilic *et al*[79]	NR	Retrospective cohort	Yes	Split	3898	1295	Surgical	D:61.9 (14.7) V: 61.6 (15.1)	30d	10	11
Stuebe *et al*[80]	NR	Hospital database	Yes	No	4800		Surgical	60–69	30d	12	
Tam *et al*[44]	NR	Retrospective cohort	Yes	Boot	63336	NR	Surgical	66.2 (10.7)	30d	11.3	
Khera *et al*[45]	TAVR 30-Day readmission risk model	Administrative	Yes	Boots, Ext	39305	40 (Boot) 885 (Ext)	TAVR	D: 81.3 V: 81.7	30d	16.2	16.2 (Boot) 18.9 (Ext)
Sanchez *et al*[50]	NR	Registry	Yes	Split	6903	3442	TAVR	D: 81.1 (7.9) V: 81.3 (7.9)	30d	9.8	10.7

Age is reported as mean (SD); median (IQR) or average age as reported in the study.
ACS, acute coronary syndrome; ADHERE, Acute Decompensated Heart Failure Registry; AF, atrial fibrillation; AH, Adventist Health Off-the-shelf model; AMI, acute myocardial infarction; Boot, bootstrapping; CABG, coronary artery bypass grafting; Car, cardiac-related; CHA2DS2-VASc, congestive heart failure, hypertension, age ≥ 75 years, diabetes mellitus, stroke or transient ischemic attack (TIA), vascular disease, age 65 to 74 years, sex category; CMS, Centers for Medicare and Medicaid Services; CRSS, CABG Readmission Risk Score; d, days; Dev, development; ELAN-HF, European Collaboration on Acute Decompensated Heart Failure; Ext, external validation; FIM, motor and cognitive Functional Independence Measure; GRACE, Global Registration of Acute Coronary Events; HF, heart failure; HFpEF, heart failure with preserved ejection fraction; HVD, heart valve disease; ICD, implantable cardioverter defibrillator; LACE, Length of stay, acuity of the Admission, Comorbidity of the patient and Emergency department use in the duration of 6 months before admission; NA, not applicable; NR, not reported; PARR-30, Patients at Risk of Re-admission within 30 days; PCI, percutaneous coronary intervention; READMITS, Renal Function, Elevated Brain Natriuretic Peptide, Age, Diabetes Mellitus, Nonmale Sex, Intervention with Timely Percutaneous Coronary Intervention, and Low Systolic Blood Pressure; SILVER-AMI, Comprehensive Evaluation of Risk Factors in Older Patients with AMI; Split, random split; STS PROM, Society of Thoracic Surgeons Predicted Risk of Mortality; TARRACO, Troponin Assessment for Risk stRatification of patients without Acute COronary atherothrombosis; TAVR, transcatheter aortic valve replacement; Val, validation.

Figure 2 PROBAST (Prediction model Risk Of Bias ASsessment Tool) risk of bias and applicability. The PROBAST tool[16] was used to assess the risk of bias for the participants, predictors, outcome and analysis for each model. Only one study demonstrated low risk of bias on all domains.

in studies using registries (0.61, n=17) and hospital databases (0.61, n=18). The discrimination decreased when the number of predictors increased (beta −0.002, n=90). There were no moderation effects based on the average age of the sample, outcome definition and endpoint of the prediction (online supplemental figures 7–8 and online supplemental table 1B).

The calibration was reported for 27 models using multiple measures and could not be pooled (table 2).

Predictors

A total of 766 predictor values were estimated in the included models. The median number of predictors per model was 15 (IQR=9–28). The predictors were mostly situated in the domains medical comorbidities (n=211),

Table 2 Model discrimination and calibration

Study	Model	Setting	Predictors; n	Cohort	Discrimination	Type calibration	Calibration
Moretti *et al*[57]	EuroHeart PCI score	ACS	16	External	0.59 (0.48–0.71)	NA	
Asche *et al*[46]	NR	AMI	19	Development; random split	0.74; NR	NA	
Cediel *et al*[58]	TARRACO risk score	AMI type 2; ischaemia	7	Development (30d)	0.71 (0.61–0.82)	NA	
		AMI type 2; ischaemia	7	Development (180d)	0.71 (0.64–0.78)	NA	
Burke *et al*[35]	HOSPITAL score	AMI	7	External	0.66 (0.61–0.71)	HLT	p=0.49
Chotechuang *et al*[36]	GRACE	AMI	9	External (30d)	0.77 (0.65–0.88)	NA	
	GRACE	AMI	9	External (180d)	0.63 (0.49–0.77)	NA	
Hilbert *et al*[59]	AMI decision tree	AMI	44	Development; External	0.65 (0.64–0.66) 0.61 (0.61–0.62)	NA	
Dodson *et al*[18]	SILVER-AMI 30-day readmission calculator	AMI	10	Development; random split	0.65; 0.63	HLT	p>0.05; p=0.05
Kini *et al*[60]	NR	AMI	12	Development; random split	NR; 0.66	Slope; in large; plot	0.973 (p=0.330); −0.038 (p=0.221)
Nguyen *et al*[19]	AMI READMITS score	AMI	7	Development; random split	0.75 (0.70–0.80) 0.73 (0.71–0.74)	Plot; plot	
	Full-stay AMI model	AMI	10	Development; random split	0.78 (0.74–0.83) 0.75 (0.74–0.76)	Plot	
	CMS AMI administrative model	AMI	32	External	0.74 (0.69–0.74)	Plot	
Krumholz *et al*[20]	CMS AMI administrative model	AMI	32	Development; external; random split	0.63; 0.63; 0.62	In large; slope	
	CMS AMI medical model	AMI	45	Development; random split	0.58; 0.59	NA	0, 1/0.015; 0.997/0.015; 0.983
Rana *et al*[33]	Elixhauser index	AMI	30	External	0.53 (0.42–0.65)	NA	
	HOSPITAL core	AMI	7	External	0.60 (0.47–0.73)	NA	
Atzema *et al*[47]	AFTER Part 2 scoring system	Arrhythmia; AF	12	Development	0.69; NR	NA	
Lahewala *et al*[40]	CHADS2	Arrhythmia; AF	5	External (30d)	0.64	NA	
	CHADS2	Arrhythmia; AF	5	External (90d)	0.63	NA	
	CHA2DS-VASc	Arrhythmia; AF	9	External (30d)	0.65	NA	
	CHA2DS-VASc	Arrhythmia; AF	9	External (90d)	0.63	NA	
Benuzillo *et al*[61]	CRSS	CABG	5	Development; bootstrapping	0.63; 0.63	HLT	7.13 (p=0.52); 9.31 (p=0.32)
Deo *et al*[62]	30-days CABG readmission calculator	CABG	20	Development	0.65	NA	
Engoren *et al*[55]	NR	CABG	6	Development; random split	0.68 (0.64–0.72) 0.68 (0.64–0.68)	NA	

Continued

Table 2 Continued

Study	Model	Setting	Predictors; n	Cohort	Discrimination	Type calibration	Calibration
Lancey *et al*[63]	NR	CABG	8	Development; random split	0.64; 0.57	NA	
Rosenblum *et al*[41]	The STS PROM score	CABG	40	External	0.59 (0.57–0.60)	NA	
Zitser-Gurevich *et al*[64]	NR	CABG	17	Development; external (30d)	0.63; 0.66/0.63	HLT	7.91 (p=0.44)
	NR	CABG	13	Development (100d)	0.65	HLT	6.76 (p=0.56)
Zywot *et al*[42]	CABG risk scale	CABG	27	Development; external	NR; 0.70	Plot	
Ahmad *et al*[21]	CMS HF administrative model	HF	37	External	0.66 (0.57–0.76)	HLT	p=0.19
Amarasingham *et al*[22]	ADHERE	HF	3	External	0.56 (0.54–0.59)	NA	
	CMS HF administrative model	HF	37	External	0.66 (0.63–0.68)	NA	
	Tabak mortality score	HF	18	External	0.61 (0.59–0.64)	NA	
Au *et al*[23]	Administrative claims model, HF 30-day mortality	HF	17	External	0.58 (0.58–0.59)	NA	
	Charlson Comorbidity Score	HF	32	External	0.55 (0.55–0 56)	NA	
	CMS HF administrative model	HF	37	External	0.59 (0.59–0.60)	NA	
	LACE	HF	18	External	0.58 (0.58–0.59)	NA	
Bardhan *et al*[65]	NR	HF	30	Development	0.56	NA	
Betihavas *et al*[66]	NR	HF	7	Development; bootstrapping	NR; 0.80	NA	
Burke *et al*[35]	HOSPITAL score	HF	7	External	0.67 (0.65–0.70)	HLT	p=0.10
Cox *et al*[24]	CMS HF administrative model	HF	37	External	0.61	NA	
	CMS HF medical model	HF	20	External	0.60	NA	
Delgado *et al*[67]	15-day CV readmission risk score	HF	5	Development; bootstrapping	0.65; 0.63	Plot	
	30-day CV readmission risk score	HF	11	Development; bootstrapping	0.66; 0.64	Plot	
Formiga *et al*[30]	CMS HF medical model	HF	19	External (30d)	0.65 (0.57–0.72)	NA	
	CMS HF medical model	HF	19	External (90d)	0.62 (0.56–0.68)	NA	
Frizzell *et al*[25]	CMS HF administrative model	HF	37	External	0.60	NA	
Hammill *et al*[26]	CMS HF administrative model	HF	37	External	0.59	Plot	
Hilbert *et al*[59]	HF decision tree	HF	44	Development; External	0.59 (0.58–0.60) 0.58 (0.58–0.59)	NA	

Continued

Table 2 Continued

Study	Model	Setting	Predictors; n	Cohort	Discrimination	Type calibration	Calibration
Hummel *et al*[31]	CMS HF medical model	HF	28	External	0.61	NA	
Huynh *et al*[48]	NR	HF	12	Development; external (30d)	0.82 (0.76–0.87) 0.73 (0.69–0.77)	NA	
	NR	HF	12	Development; external (90d)	NR; 0.65	NA	
Ibrahim *et al*[34]	HOSPITAL score	HfpEF	7	External	0.60 (0.55–0.64)	NA	
	LACE	HfpEF	18	External	0.55 (0.50–0.60)	NA	
	LACE + index	HfpEF	24	External	0.57 (0.52–0.62)	NA	
Keenan *et al*[27]	CMS HF administrative model	HF	37	Development; external; random split	0.60; 0.60; 0.61	In large; slope	0, 1/0.02; 1.01/0.09; 1.05
	CMS HF medical model	HF	30	Development; random split	0.58; 0.61	In large; slope	0, 1/0, 1
Kitamura *et al*[53]	FIM	HF	13	External	0.78	NA	
Leong *et al*[68]	30-day HF readmission risk score	HF	7	Development; random split	0.76; 0.76	NA	
Li *et al*[49]	NR	HF	10	Development; random split	0.63 (0.62–0.63) 0.63 (0.62–0.63)	HLT; plot	0.15 (p>0.005)
Lim *et al*[69]	NR	HF	13	Development	0.68 (car); 0.62 (all)	HLT	27.5 (p=0.001) (car) 8.0 (p=0.429) (all)
Reed *et al*[28]	AH model	HF	14	Development; random split	0.86 (0.85–0.86) 0.85 (0.84–0.86)	NA	
	CMS HF administrative model	HF	37	Random split	0.55 (0.54–0.56) 0.55 (0.54–0.57)	NA	
	Hasan	HF	9	Random split	0.80 (0.79–0.81) 0.80 (0.80–0.82)	NA	
	LACE	HF	18	Random split	0.75 (0.74–0.81) 0.74 (0.73–0.76)	NA	
Reed *et al* (continued)[28]	PARR-30	HF	10	Random split	0.82 (0.81–0.83) 0.81 (0.80–0.82)	NA	
Salah *et al*[70]	ELAN-HF Score	HF	10	Development	0.60 (0.56–0.64)	NA	
Sudhakar *et al*[32]	CMS HF medical model	HF	20	External	0.61 (0.57–0.64) ≥65 y, 0.59 (0.53–0.64) Random patient-level, 0.58 (0.50–0.65)	NA	
Tan *et al*[71]	NR	HF	3	Random split	0.73	HLT; plot	p=0.62
Wang *et al*[72]	NR	HF	12	Development	0.65	NA	
Wang *et al*[38]	LACE	HF	18	External	0.56 (0.48–0.64)	NA	
Yazdan-Ashoori *et al*[29]	CMS HF administrative model	HF	37	External	0.61 (0.55–0.67)	NA	
	LACE	HF	18	External	0.59 (0.52–0.65)	HLT	p=0.73

Continued

Table 2 Continued

Study	Model	Setting	Predictors; n	Cohort	Discrimination	Type calibration	Calibration
Disdier Moulder *et al*[73]	NR	HF; ACS; NR	4	Development (30d)	0.68	NA	
	NR	HF; ACS; NR	5	Development (180d)	0.69	NA	
Raposeiras-Roubín *et al*[37]	GRACE	HF; ACS	9	External	0.74 (0.73–0.80)	HLT	p=0.14
Minges *et al*[74]	NR	HF; PCI	35	Development; random split	0.67; 0.66	NA	
Pack *et al*[75]	NR	HVD	28	Development; random split	0.67 (full dev)/ 0.65 (nomogram); 0.67 (full val)	Harrell's E; O,E; Harrell's E; plot	0.1%; 1.9%; 1.6%
Oliver-McNeil *et al*[76]	ICD readmission-risk score	ICD	4	Update; External	0.69 (0.58–0.79)	HLT; plot	3.44 (p=0.49)
Wasfy *et al*[52]	Pre-PCI model	NR	23	Development; random split	0.68; 0.67	HLT; plot	p=0.59
Barnett *et al*[77]	NR validation	Surgical	15	External	0.59	NA	
	NR update	Surgical	18	Update	0.60 (0.59–0.62)	NA	
Brown *et al*[43]	STS augmented clinical model	Surgical	27	Update (bootstrap); random split; external (bootstrap)	0.66 (0.61–0.72); 0.56; 0.47 (0.42–0.53)	HLT	p=1.0
	STS 30-day readmission model	Surgical	21	Update (bootstrap); random split; external (bootstrap)	0.66 (0.62–0.71), 0.58, 0.47 (0.41–0.52)	HLT	p=0.492
Espinoza *et al*[78]	30-day readmission score after cardiac surgery	Surgical	5	Development; random split	0.66 (0.63–0.70) 0.64 (0.61–0.67)	NA	
Ferraris *et al*[54]	READMIT	Surgical	9	Development	0.70	HLT	5.966 (p=0.651)
Kilic *et al*[79]	NR	Surgical	15	Development; random split	NR; 0.64	HLT; plot	p=0.45; p=0.57
Stuebe *et al*[80]	NR	Surgical	7	Development	0.63	NA	
Tam *et al*[44]	NR	Surgical	29	Development; bootstrapping	0.63; 0.65	Plot	
Khera *et al*[45]	TAVR 30-Day readmission risk model	TAVR	11	Development; random split; external	NR; 0.63; 0.69	HLT; RMSE; RMSE; plot	p=0.33; 0.978; 0.928
Sanchez *et al*[50]	NR	TAVR	10	Development; random split	0.61; 0.60	HLT	p=0.749; p=0.403

ACS, acute coronary syndrome; ADHERE, Acute Decompensated Heart Failure Registry; AF, atrial fibrillation; AH, Adventist Health Off-the-shelf model; AMI, acute myocardial infarction; CABG, coronary artery bypass grafting; Car, cardiac-related; CHADS2, Congestive heart failure, Hypertension, Age, Diabetes, previous Stroke/transient ischemic attack; CHADS2-VASc, congestive heart failure, hypertension, age ≥ 75 years, diabetes mellitus, stroke or transient ischemic attack (TIA), vascular disease, age 65 to 74 years, sex category; CMS, Centers for Medicare and Medicaid Services; CRSS, CABG Readmission Risk Score; d, days; dev, development; FIM, motor and cognitive Functional Independence Measure; GRACE, Global Registry of Acute Coronary Events; HF, heart failure; HFpEF, heart failure with preserved ejection fraction; HLT, Hosmer-Lemeshow test; HOSPITAL, Hemoglobin level, discharged from Oncology, Sodium level, Procedure during admission, Index admission Type, Admission, Length of stay; HVD, heart valve disease; ICD, implantable cardioverter defibrillator; LACE, Length of stay, acuity of the Admission, Comorbidity of the patient and Emergency department use in the duration of 6 months before admission; NA, not applicable; NR, not reported; O, E, observed, expected; PARR-30, Patients at Risk of Re-admission within 30 days; PCI, percutaneous coronary intervention; plot, calibration plot; READMITS, Renal Function, Elevated Brain Natriuretic Peptide, Age, Diabetes Mellitus, Nonmale Sex, Intervention with Timely Percutaneous Coronary Intervention, and Low Systolic Blood Pressure; SILVER-AMI, Comprehensive Evaluation of Risk Factors in Older Patients with AMI; STS, Society of Thoracic; STS, Society of Thoracic; STS PROM, Society of Thoracic Surgeons Predicted Risk of Mortality; TAVR, transcatheter aortic valve replacement; val, validation.

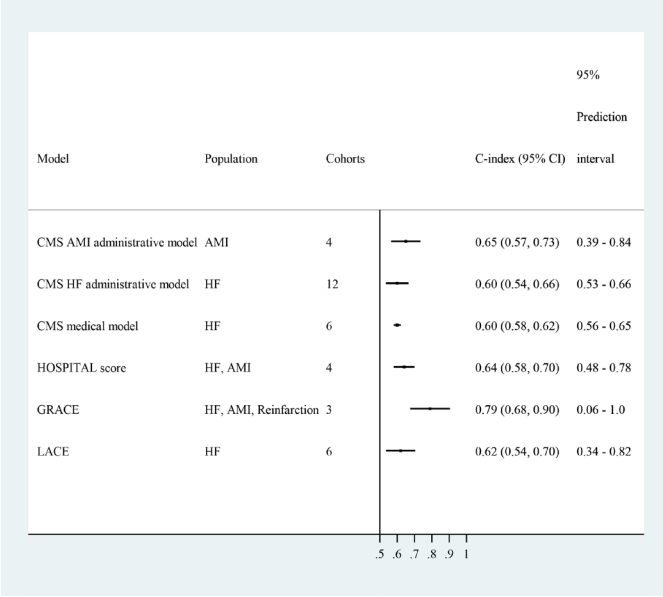

Figure 3 Meta-analysis of prediction models. Random-effect models were used to pool similar models reported in independent cohorts. For the HOSPITAL score, the discriminations for the HF and AMI samples were similar (0.65 and 0.64). For GRACE, the discriminations for the AMI and reinfarction samples were similar (0.77 and 0.74), and was higher for the HF sample (0.83). Only GRACE demonstrated adequate discrimination in external cohorts. AMI, acute myocardial infarction; CMS, Centers for Medicare and Medicaid Services; CF, heart failure.

Abbreviations: CMS = Centers for Medicare and Medicaid Services; AMI = Acute Myocardial Infarction; HF = Heart Failure; HOSPITAL = Hemoglobin level, discharged from Oncology, Sodium level, Procedure during admission, Index admission Type, Admission, Length of stay; GRACE = Global Registry of Acute Coronary Events; LACE = length of stay (L), acuity of the admission (A), comorbidity of the patient (C) and emergency department use in the duration of 6 months before admission.

disease and hospital characteristics (n=128), demographic data (n=128), laboratory values (n=97) and medical history characteristics (n=51). Age (n=47), presence of diabetes (n=26), insurance status (n=24), length of stay (n=28) and gender (n=23) were the most prevalent predictors. There was little consistency in the definition of predictors, and most studies did not report how they were measured.

Only 18 predictors were similarly defined in multiple studies and could be pooled for the outcome readmission at 30 days (figure 4, online supplemental table 2A and online supplemental figures 9–26). The coefficients of four predictors demonstrated a consistent and significant association across the different samples: chronic obstructive pulmonary disease (COPD), HF or history of HF, and valvular disease. The coefficients of 11 predictors demonstrated an overall significant association, that is, age, female gender, arrhythmias, chronic lung disease, diabetes mellitus, cerebrovascular disease, cardiovascular accident, anaemia, peripheral vascular disease, urgent admission and infection, but this was not consistent across the samples and the prediction intervals were not significant. The effect of these predictors was mostly smaller in the HF samples.

The coefficients for most predictors could not be pooled because they had different definitions, cut-off values or reference categories. However, renal disease, including dialysis, a longer length of stay, creatinine, NT-proBNP (N-Terminal-PRO hormone Brain Natriuretic Peptide) and previous hospital admissions demonstrated a consistent association with readmissions.

DISCUSSION

In this systematic review, we included 60 studies that reported the results from 81 separate clinical risk prediction models and 766 risk predictors for unplanned readmission in patients with acute heart disease. We found some promising prediction models, however, no clinical model demonstrated good discrimination (ie, c-statistic >0.8) in independently externally validated cohorts, regardless of the underlying patient populations. GRACE was the only model that demonstrated adequate discrimination in multiple cohorts in patients with acute coronary syndromes[36 37] and HF.[37] There was little consistency in the measurement of risk predictors.

The results of our review are in line with previous systematic reviews which have mainly focused on samples of patients with HF, AMI or focused on generic prediction models. All reviews confirm that the discrimination is generally low. Our review confirms the importance of previous HF[5 6] and previous hospital admissions[6 8] as consistent predictors of the risk of readmission. In addition, two prevalent comorbidities, COPD and valve disease, were also consistent predictors across the different populations. Other reviews also identified the importance of age, gender, comorbidities and certain laboratory values. These were also significant in our review but the association was not always consistent across the different populations or heterogeneously measured making comparisons difficult. As a result, no clinical risk prediction model or set of predictors that is relevant for different populations of heart disease could be identified.

Our review focused specifically on prediction models with a clinical presentation that can be used in daily practice, for example, risk scores or nomograms. These simple models do not consider interactions between predictor values or non-linear link functions in their predictions. This may partially explain the poor discrimination.[39] Using web applications or electronic patient records to run more complex prediction algorithms can likely offer a solution for future models. A recent systematic review observed an average c-statistic of 0.74 for models using electronic patient records and machine learning algorithms.[11] Our review included 11 studies[18 29 32 36 37 40–45] that developed or validated electronic tools for risk prediction and their discrimination ranged between 0.59 and 0.77.

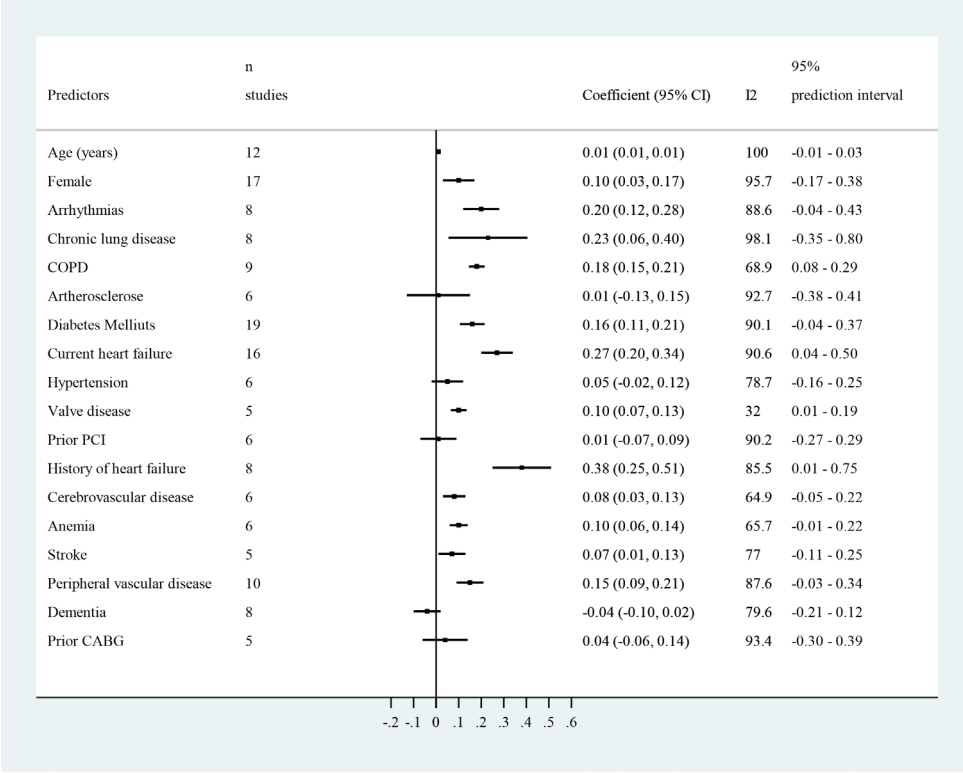

Figure 4 Predictors of unplanned hospital readmission. The plot provides an overview of the random-effects meta-analyses that were performed for predictors who were similarly defined for the outcome unplanned hospital readmission at 30-day follow-up. See online supplemental table 2A and online supplemental figures 9–26 for more details. CABG, coronary artery bypass grafting; COPD, chronic obstructive pulmonary disease; PCI, percutaneous coronary intervention.

However, these electronic tools were mostly derived from score charts and nomograms.

There are also concerns about the generalisability of the prediction models. The median age of patients included in the samples was 68 years (IQR=65–75). However, older and frail patients suffer more multimorbidity and geriatric syndromes, and the distribution of predictor and outcome values will also be different than in younger samples. It is therefore unlikely that the majority of the current models will hold their value in daily clinical practice where there is a high prevalence of older patients. Only eight studies[18 20 27 46–50] included one or more geriatric risk factors (eg, physical performance, dementia) as predictors for readmission. The performance of models including geriatric conditions was similar to models without these conditions. This might be explained by the relative young mean age of the samples in our review. Mahmoudi *et al*[11] reported that functional and frailty status are important predictors, but were only included in a small number of studies. Frailty was not identified in any of the models in our review. It might be valuable to examine the additive value of these predictors in prediction models for patients with heart disease.

We observed high RoB in almost all clinical risk prediction models (98.8%). This was mainly because the calibration was lacking or not fully reported (eg, only p value of Hosmer-Lemeshow test). Furthermore, most studies performed retrospective data analyses or used data from existing sources. However, our results demonstrate that studies using these data sources had the lowest c-statistic, and that the c-statistic decreased when more predictors were tested. Databases often have missing data, misclassification bias and random measurement error, which likely explains their average poor performance.[51] Only the SILVER-AMI (Comprehensive Evaluation of Risk Factors in Older Patients with AMI) study[18] demonstrated low RoB on all domains. However, their readmission risk calculator for older patients with AMI only discriminated modestly (c-statistic=0.65).

Our review shows the current state-of-the art of risk prediction in patients with acute heart disease. The timely identification of patients with acute heart disease at risk of readmission remains challenging with the prediction models identified in this systematic review. Therefore, further research in risk prediction remains important and some recommendations for further research can be derived from this review. First, consistency is needed in the definition and measurement of predictors. More homogeneity might improve the identification of important predictors and their effect on readmission. Based on our insights, we believe that models could be improved by incorporating some key predictors, that is, age, gender, comorbidity scores (or at least heart failure, COPD, cardiovascular disease, diabetes mellitus), admission status, readmission history and the geriatric profile (eg, functional status, cognitive

status). Because there are a still a large number of potential predictors, a large sample size is needed to estimate the coefficients with sufficient precision, and to prevent against overfitting the models. Some selection of predictors may still be warranted, and penalised techniques (eg, lasso regression) should be preferred over traditional selection based on p values. Second, the results suggest that multiple predictors are associated with readmissions regardless of the underlying population. Therefore, attention might be shifted from developing new risk prediction models to updating and externally validating existing prediction models in different populations with heart disease. For example, the Adventist Health Off-the-shelf model[28] showed high discrimination rates in both the development (0.86) and the validation cohorts (0.85). External validation is recommended to examine the generalisability of this model in other settings. In addition, the AMI READMITS (Acute Myocardial Infarction Renal Function, Elevated Brain Natriuretic Peptide, Age, Diabetes Mellitus, Nonmale Sex, Intervention with Timely Percutaneous Coronary Intervention, and Low Systolic Blood Pressure) score,[19] full-stay AMI readmission model,[19] pre-PCI model,[52] motor and cognitive Functional Independence Measure (FIM),[53] READMIT,[54] 30-day readmission model of Huynh *et al*,[48] and the model of Engoren *et al*[55] were examined in one study and showed reasonable c-statistics in the development (0.68–0.82) and validation cohorts (0.64–0.78). For these studies, model updating recalibration and external validation is recommended to improve the predictive performance and generalisability of these prediction models. Third, the applicability of current prediction models in daily practice is an important concern as most models had poor performance, were not replicated and had high RoB. More high-quality studies are needed that evaluate the discrimination, calibration and clinical usefulness. To limit the RoB as much as possible, future studies should adhere to the relevant reporting guidelines[56] and could use PROBAST[16] as a guidance to plan their study. Fourth, more complex models integrated in electronic patient records may results in better predictions.

Limitations

Although we performed an extensive literature search, we might have missed some eligible studies, particularly those published in non-English languages. We were able to perform meta-analysis for predictors that were often (≥5 models) reported. However, it might be possible that some less frequently mentioned predictors (eg, geriatric predictors) are a valuable addition in clinical practice. The review included a large number of results and statistical tests which may result in an inflated alpha error. The meta-regression identified that models with less predictors had a better discrimination, but this could also be explained by overfitting models; this could not be tested.

CONCLUSION

A large number of clinical models have recently been developed. Although some models are promising as they demonstrated adequate to good discrimination, no model can currently be recommended for clinical practice. The lack of independently validated studies, high RoB and low consistency in measured predictors limit their applicability. Model updating and external validation is urgently needed.

Author affiliations
[1]Department of Public Health and Primary Care, Katholieke Universiteit Leuven, Leuven, Belgium
[2]Research Foundation Flanders, Brussel, Belgium
[3]Center of Expertise Urban Vitality, Faculty of Health, Amsterdam University of Applied Sciences, Amsterdam, Netherlands
[4]Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam, Noord-Holland, Netherlands
[5]Faculty of Science, Amsterdam UMC Locatie AMC, Amsterdam, Netherlands
[6]Medical Library, Amsterdam UMC Location AMC, Amsterdam, North Holland, Netherlands
[7]Department of Public Health and Primary Care, KU Leuven – University of Leuven, Leuven, Belgium
[8]Department of Public Health, University of Basel, Basel, Switzerland
[9]Department of Geriatric Medicine, University Hospitals Leuven, Leuven, Belgium
[10]Department of Public Health and Primary Care, University Hospitals Leuven, Leuven, Belgium
[11]Department of Geriatric Medicine, KU Leuven – University of Leuven, Leuven, Belgium

Twitter Patricia Jepma @Patricia_Jepma and Mieke Deschodt @mieke_deschodt

Contributors BVG and PJ had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. BVG and PJ contributed equally as first authors. Concept and design: All authors. Acquisition, analysis or interpretation of data: BVG, PJ, CR, ML, JD. Drafting the manuscript: BVG, PJ. Critical revision of the manuscript: All authors. Analysis: BVG, PJ. Supervision: BB.

Funding This work was partly supported by the Research Foundation Flanders (FWO) fellowship grant (grant number 1165518N (BVG)), and by the Dutch Research Council (NWO) (grant number 023.009.036 (PJ)). The funders had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs
Bastiaan Van Grootven http://orcid.org/0000-0002-3182-573X
Patricia Jepma http://orcid.org/0000-0003-1271-6869

REFERENCES

1 Virani SS, Alonso A, Benjamin EJ, *et al*. Heart disease and stroke statistics-2020 update: a report from the American heart association. *Circulation* 2020;141:e139–596.
2 Shipe ME, Deppen SA, Farjah F, *et al*. Developing prediction models for clinical use using logistic regression: an overview. *J Thorac Dis* 2019;11:S574–84.
3 Smith LN, Makam AN, Darden D, *et al*. Acute myocardial infarction readmission risk prediction models: a systematic review of model performance. *Circ Cardiovasc Qual Outcomes* 2018;11:e003885.
4 Di Tanna GL, Wirtz H, Burrows KL, *et al*. Evaluating risk prediction models for adults with heart failure: a systematic literature review. *PLoS One* 2020;15:e0224135.
5 Mahajan SM, Heidenreich P, Abbott B, *et al*. Predictive models for identifying risk of readmission after index hospitalization for heart failure: a systematic review. *Eur J Cardiovasc Nurs* 2018;17:675–89.
6 O'Connor M, Murtaugh CM, Shah S, *et al*. Patient characteristics predicting readmission among individuals hospitalized for heart failure. *Med Care Res Rev* 2016;73:3–40.
7 Rahimi K, Bennett D, Conrad N, *et al*. Risk prediction in patients with heart failure: a systematic review and analysis. *JACC Heart Fail* 2014;2:440–6.
8 Betihavas V, Davidson PM, Newton PJ, *et al*. What are the factors in risk prediction models for rehospitalisation for adults with chronic heart failure? *Aust Crit Care* 2012;25:31–40.
9 Desai MM, Stauffer BD, Feringa HHH, *et al*. Statistical models and patient predictors of readmission for acute myocardial infarction: a systematic review. *Circ Cardiovasc Qual Outcomes* 2009;2:500–7.
10 Ross JS, Mulvey GK, Stauffer B, *et al*. Statistical models and patient predictors of readmission for heart failure: a systematic review. *Arch Intern Med* 2008;168:1371–86.
11 Mahmoudi E, Kamdar N, Kim N, *et al*. Use of electronic medical records in development and validation of risk prediction models of hospital readmission: systematic review. *BMJ* 2020;369:m958.
12 Zhou H, Della PR, Roberts P, *et al*. Utility of models to predict 28-day or 30-day unplanned spital readmissions: an updated systematic review. *BMJ Open* 2016;6:e011060.
13 Song F, Sheldon TA, Sutton AJ, *et al*. Methods for exploring heterogeneity in meta-analysis. *Eval Health Prof* 2001;24:126–51.
14 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
15 Distiller. Distiller systematic review software. Available: https://www.evidencepartners.com/products/distillersr-systematic-review-software/ [Accessed 25 Aug 2020].
16 Wolff RF, Moons KGM, Riley RD, *et al*. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;170:51–8.
17 Newcombe RG. Confidence intervals for an effect size measure based on the mann-whitney statistic. part 2: asymptotic methods and evaluation. *Stat Med* 2006;25:559–73.
18 Dodson JA, Hajduk AM, Murphy TE, *et al*. Thirty-day readmission risk model for older adults hospitalized with acute myocardial infarction. *Circ Cardiovasc Qual Outcomes* 2019;12:e005320.
19 Nguyen OK, Makam AN, Clark C, *et al*. Predicting 30-Day Myocardial Infarction: The AMI "READMITS" (Renal Function, elevated brain natriuretic peptide, age, diabetes mellitus, nonmale sex, intervention with timely percutaneous coronary intervention, and low systolic blood pressure) score. *J Am Heart Assoc* 2018;7:e008882.
20 Krumholz HM, Lin Z, Drye EE, *et al*. An administrative claims measure suitable for profiling hospital performance based on 30-day all-cause readmission rates among patients with acute myocardial infarction. *Circ Cardiovasc Qual Outcomes* 2011;4:243–52.
21 Ahmad FS, French B, Bowles KH, *et al*. Incorporating patient-centered factors into heart failure readmission risk prediction: a mixed-methods study. *Am Heart J* 2018;200:75–82.
22 Amarasingham R, Moore BJ, Tabak YP, *et al*. An automated model to identify heart failure patients at risk for 30-day readmission or death using electronic medical record data. *Med Care* 2010;48:981–8.
23 Au AG, McAlister FA, Bakal JA, *et al*. Predicting the risk of unplanned readmission or death within 30 days of discharge after a heart failure hospitalization. *Am Heart J* 2012;164:365–72.
24 Cox ZL, Lai P, Lewis CM, *et al*. Customizing national models for a medical center's population to rapidly identify patients at high risk of 30-day all-cause Hospital readmission following a heart failure hospitalization. *Heart Lung* 2018;47:290–6.
25 Frizzell JD, Liang L, Schulte PJ, *et al*. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. *JAMA Cardiol* 2017;2:204–9.
26 Hammill BG, Curtis LH, Fonarow GC, *et al*. Incremental value of clinical data beyond claims data in predicting 30-day outcomes after heart failure hospitalization. *Circ Cardiovasc Qual Outcomes* 2011;4:60–7.
27 Keenan PS, Normand S-LT, Lin Z, *et al*. An administrative claims measure suitable for profiling hospital performance on the basis of 30-day all-cause readmission rates among patients with heart failure. *Circ Cardiovasc Qual Outcomes* 2008;1:29–37.
28 Reed J, Bokovoy J, Doram K. Unplanned readmissions after hospital discharge among heart failure patients at risk for 30-day readmission using an administrative dataset and "off the shelf" readmission models. *Internet J Cardiovasc Res* 2014;9:07–15.
29 Yazdan-Ashoori P, Lee SF, Ibrahim Q, *et al*. Utility of the lace index at the bedside in predicting 30-day readmission or death in patients hospitalized with heart failure. *Am Heart J* 2016;179:51–8.
30 Formiga F, Masip J, Chivite D, *et al*. Applicability of the heart failure readmission risk score: a first European study. *Int J Cardiol* 2017;236:304–9.
31 Hummel SL, Katrapati P, Gillespie BW, *et al*. Impact of prior admissions on 30-day readmissions in medicare heart failure inpatients. *Mayo Clin Proc* 2014;89:623–30.
32 Sudhakar S, Zhang W, Kuo Y-F, *et al*. Validation of the readmission risk score in heart failure patients at a tertiary hospital. *J Card Fail* 2015;21:885–91.
33 Rana S, Tran T, Luo W, *et al*. Predicting unplanned readmission after myocardial infarction from routinely collected administrative hospital data. *Aust Health Rev* 2014;38:377–82.
34 Ibrahim AM, Koester C, Al-Akchar M, *et al*. Hospital score, lace index and lace+ index as predictors of 30-day readmission in patients with heart failure. *BMJ Evid Based Med* 2020;25:166-167.
35 Burke RE, Schnipper JL, Williams MV, *et al*. The hospital score predicts potentially preventable 30-day readmissions in conditions targeted by the hospital readmissions reduction program. *Med Care* 2017;55:285–90.
36 Chotechuang Y, Phrommintikul A, Muenpa R, *et al*. The prognostic utility of grace risk score in predictive cardiovascular event rate in STEMI patients with successful fibrinolysis and delay intervention in non PCI-capable Hospital: a retrospective cohort study. *BMC Cardiovasc Disord* 2016;16:212.
37 Raposeiras-Roubín S, Abu-Assi E, Cambeiro-González C, *et al*. Mortality and cardiovascular morbidity within 30 days of discharge following acute coronary syndrome in a contemporary European cohort of patients: how can early risk prediction be improved? the six-month grace risk score. *Rev Port Cardiol* 2015;34:383–91.
38 Wang H, Robinson RD, Johnson C, *et al*. Using the lace index to predict hospital readmissions in congestive heart failure patients. *BMC Cardiovasc Disord* 2014;14:97.
39 Kerr KF, Pepe MS. Joint modeling, covariate adjustment, and interaction: contrasting notions in risk prediction models and risk prediction performance. *Epidemiology* 2011;22:805–12.
40 Lahewala S, Arora S, Patel P, *et al*. Atrial fibrillation: Utility of $CHADS_2$ and CHA_2DS_2-VASc scores as predictors of readmission, mortality and resource utilization. *Int J Cardiol* 2017;245:162–7.
41 Rosenblum JM, Lovasik BP, Hunting JC, *et al*. Predicted risk of mortality score predicts 30-day readmission after coronary artery bypass grafting. *Gen Thorac Cardiovasc Surg* 2019;67:661–8.
42 Zywot A, Lau CSM, Glass N, *et al*. Preoperative scale to determine all-cause readmission after coronary artery bypass operations. *Ann Thorac Surg* 2018;105:1086–93.
43 Brown JR, Jacobs JP, Alam SS, *et al*. Utility of biomarkers to improve prediction of readmission or mortality after cardiac surgery. *Ann Thorac Surg* 2018;106:1294–301.
44 Tam DY, Fang J, Tran A, *et al*. A clinical risk scoring tool to predict readmission after cardiac surgery: an Ontario administrative and clinical population database study. *Can J Cardiol* 2018;34:1655–64.
45 Khera S, Kolte D, Deo S, *et al*. Derivation and external validation of a simple risk tool to predict 30-day hospital readmissions after transcatheter aortic valve replacement. *EuroIntervention* 2019;15:155–63.
46 Asche CV, Ren J, Kirkness CS. A prediction model to identify acute myocardial infarction (AMI) patients at risk for 30-day readmission. *SCSC* 2016;1:1–8 https://dl.acm.org/doi/
47 Atzema CL, Dorian P, Fang J, *et al*. A clinical decision instrument to predict 30-day death and cardiovascular hospitalizations after an emergency department visit for atrial fibrillation: the atrial

fibrillation in the emergency room, part 2 (AFTER2) study. *Am Heart J* 2018;203:85–92.
48 Huynh Q, Negishi K, De Pasquale CG, *et al*. Validation of predictive score of 30-day hospital readmission or death in patients with heart failure. *Am J Cardiol* 2018;121:322–9.
49 Li L, Baek J, Jesdale BM, *et al*. Predicting 30-day mortality and 30-day re-hospitalization risks in Medicare patients with heart failure discharged to skilled nursing facilities: development and validation of models using administrative data. *J Nurs Home Res Sci* 2019;5:60–7.
50 Sanchez CE, Hermiller JB, Pinto DS, Jr PDS, *et al*. Predictors and risk calculator of early unplanned hospital readmission following contemporary self-expanding transcatheter aortic valve replacement from the STS/ACC TVT registry. *Cardiovasc Revasc Med* 2020;21:263–70.
51 Jordan K, Moons K. Electronic healthcare records and prognosis research. In: *Prognosis research in healthcare. concepts, methods and impact*. Oxford: Oxford University press, 2019.
52 Wasfy JH, Rosenfield K, Zelevinsky K, *et al*. A prediction model to identify patients at high risk for 30-day readmission after percutaneous coronary intervention. *Circ Cardiovasc Qual Outcomes* 2013;6:429–35.
53 Kitamura M, Izawa KP, Taniue H, *et al*. Relationship between activities of daily living and readmission within 90 days in hospitalized elderly patients with heart failure. *Biomed Res Int* 2017;2017:7420738.
54 Ferraris VA, Ferraris SP, Harmon RC, *et al*. Risk factors for early Hospital readmission after cardiac operations. *J Thorac Cardiovasc Surg* 2001;122:278–86.
55 Engoren M, Habib RH, Dooner JJ, *et al*. Use of genetic programming, logistic regression, and artificial neural nets to predict readmission after coronary artery bypass surgery. *J Clin Monit Comput* 2013;27:455–64.
56 et alCollins G, Reitsma J, Altman D. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement, 2020. Available: https://www.equator-network.org/reporting-guidelines/tripod-statement/ [Accessed 20 Aug 2020].
57 Moretti C, D'Ascenzo F, Omedè P, *et al*. Thirty-day readmission rates after PCI in a metropolitan center in Europe: incidence and impact on prognosis. *J Cardiovasc Med* 2015;16:238–45.
58 Cediel G, Sandoval Y, Sexter A, *et al*. Risk estimation in type 2 myocardial infarction and myocardial injury: the tarraco risk score. *Am J Med* 2019;132:217–26.
59 Hilbert JP, Zasadil S, Keyser DJ, *et al*. Using decision trees to manage Hospital readmission risk for acute myocardial infarction, heart failure, and pneumonia. *Appl Health Econ Health Policy* 2014;12:573–85.
60 Kini V, Peterson PN, Spertus JA, *et al*. Clinical model to predict 90-day risk of readmission after acute myocardial infarction. *Circ Cardiovasc Qual Outcomes* 2018;11:e004788.
61 Benuzillo J, Caine W, Evans RS, *et al*. Predicting readmission risk shortly after admission for CABG surgery. *J Card Surg* 2018;33:163–70.
62 Deo SV, Raza S, Altarabsheh SE, *et al*. Risk calculator to predict 30-day readmission after coronary artery bypass: a strategic decision support tool. *Heart Lung Circ* 2019;28:1896–903.
63 Lancey R, Kurlansky P, Argenziano M, *et al*. Uniform standards do not apply to readmission following coronary artery bypass surgery: a multi-institutional study. *J Thorac Cardiovasc Surg* 2015;149:e1:850–7.
64 Zitser-Gurevich Y, Simchen E, Galai N, *et al*. Prediction of readmissions after CABG using detailed follow-up data: the Israeli CABG study (ISCAB). *Med Care* 1999;37:625–36.
65 Bardhan I, Oh J, Zhiqiang Z, *et al*. Predictive analytics for readmission of patients with congestive heart failure. *Information Systems Research* 2015;26:19–39.
66 Betihavas V, Frost SA, Newton PJ, *et al*. An absolute risk prediction model to determine unplanned cardiovascular readmissions for adults with chronic heart failure. *Heart Lung Circ* 2015;24:1068–73.
67 Delgado JF, Ferrero Gregori A, Fernández LM, *et al*. Patient-associated predictors of 15- and 30-day readmission after hospitalization for acute heart failure. *Curr Heart Fail Rep* 2019;16:304–14.
68 Leong KTG, Wong LY, Aung KCY, *et al*. Risk stratification model for 30-day heart failure readmission in a multiethnic South East Asian community. *Am J Cardiol* 2017;119:1428–32.
69 Lim N-K, Lee SE, Lee H-Y, *et al*. Risk prediction for 30-day heart failure-specific readmission or death after discharge: data from the Korean acute heart failure (KorAHF) registry. *J Cardiol* 2019;73:108–13.
70 Salah K, Kok WE, Eurlings LW, *et al*. A novel discharge risk model for patients hospitalised for acute decompeNsated heart failure incorporating N-terminal pro-B-type natriuretic peptide levels: a European coLlaboration on acute decompensated heart failure: ELAN-HF score. *Heart* 2014;100:115–25.
71 Tan B-Y, Gu J-Y, Wei H-Y, *et al*. Electronic medical record-based model to predict the risk of 90-day readmission for patients with heart failure. *BMC Med Inform Decis Mak* 2019;19:193.
72 Wang LE, Shaw PA, Mathelier HM, *et al*. Evaluating risk-prediction models using data from electronic health records. *Ann Appl Stat* 2016;10:286–304.
73 Disdier Moulder MP, Larock JM, Garofoli A, *et al*. Family help with medication management: a predictive marker for early readmission. *Mayo Clin Proc Innov Qual Outcomes* 2017;1:211–8.
74 Minges KE, Herrin J, Fiorilli PN, *et al*. Development and validation of a simple risk score to predict 30-day readmission after percutaneous coronary intervention in a cohort of medicare patients. *Catheter Cardiovasc Interv* 2017;89:955–63.
75 Pack QR, Priya A, Lagu T, *et al*. Development and validation of a predictive model for short- and medium-term hospital readmission following heart valve surgery. *J Am Heart Assoc* 2016;5:e003544.
76 Oliver-McNeil S, Templin TN, Haines DE. Preoperative ICD risk score variables predict 30-day readmission after implantable cardioverter defibrillator implantation in patients with heart failure. *Heart Lung* 2016;45:29–33.
77 Barnett SD, Sarin E, Kiser AC, *et al*. Examination of a proposed 30-day readmission risk score on discharge location and cost. *Ann Thorac Surg* 2020;109:1797–803.
78 Espinoza J, Camporrontondo M, Vrancic M, *et al*. 30-day readmission score after cardiac surgery. *Clin Trials Regul Sci Cardiol* 2016;20:1–5.
79 Kilic A, Magruder JT, Grimm JC, *et al*. Development and validation of a score to predict the risk of readmission after adult cardiac operations. *Ann Thorac Surg* 2017;103:66–73.
80 Stuebe J, Rydingsward J, Lander H, *et al*. A pragmatic preoperative prediction score for nonhome discharge after cardiac operations. *Ann Thorac Surg* 2018;105:1384–91.