Journal of the American Heart Association

SYSTEMATIC REVIEW AND META-ANALYSIS

Clinical Prediction Models for Heart Failure Hospitalization in Type 2 Diabetes: A Systematic Review and Meta-Analysis

Amir Razaghizad, MSc; Emily Oulousian; Varinder Kaur Randhawa, MD, PhD; João Pedro Ferreira, MD, PhD; James M. Brophy, MD, PhD; Stephen J. Greene, MD; Julian Guida, BSc; G. Michael Felker, MD; Marat Fudim, MD, MHS; Michael Tsoukas, MD; Tricia M. Peters, MD, PhD; Thomas A. Mavrakanas, MD; Nadia Giannetti, MD; Justin Ezekowitz, MBChB, MSc; Abhinav Sharma, MD, PhD

BACKGROUND: Clinical prediction models have been developed for hospitalization for heart failure in type 2 diabetes. However, a systematic evaluation of these models' performance, applicability, and clinical impact is absent.

METHODS AND RESULTS: We searched Embase, MEDLINE, Web of Science, Google Scholar, and Tufts' clinical prediction registry through February 2021. Studies needed to report the development, validation, clinical impact, or update of a prediction model for hospitalization for heart failure in type 2 diabetes with measures of model performance and sufficient information for clinical use. Model assessment was done with the Prediction Model Risk of Bias Assessment Tool, and meta-analyses of model discrimination were performed. We included 15 model development and 3 external validation studies with data from 999 167 people with type 2 diabetes. Of the 15 models, 6 had undergone external validation and only 1 had low concern for risk of bias and applicability (Risk Equations for Complications of Type 2 Diabetes). Seven models were presented in a clinically useful manner (eg, risk score, online calculator) and 2 models were classified as the most suitable for clinical use based on study design, external validity, and point-of-care usability. These were Risk Equations for Complications of Type 2 Diabetes (meta-analyzed c-statistic, 0.76) and the Thrombolysis in Myocardial Infarction Risk Score for Heart Failure in Diabetes (meta-analyzed c-statistic, 0.78), which was the simplest model with only 5 variables. No studies reported clinical impact.

CONCLUSIONS: Most prediction models for hospitalization for heart failure in patients with type 2 diabetes have potential concerns with risk of bias or applicability, and uncertain external validity and clinical impact. Future research is needed to address these knowledge gaps.

Key Words: clinical prediction models ■ diabetes ■ heart failure ■ meta-analysis ■ prognostication ■ risk evaluation ■ systematic review

Type 2 diabetes (T2D) contributes to >1.5 million annual deaths worldwide.[1] This high rate of mortality can be partly attributed to the metabolic alterations that precipitate diabetic cardiomyopathy and its attendant complications.[2] Correspondingly, the development of heart failure (HF) has emerged as one of the most common and important manifestations of cardiovascular disease in individuals with diabetes.[3] The concern over HF initially stemmed from safety troubles with certain antidiabetic treatments (eg, thiazolidinediones), and the fact that hospitalizations for heart failure (HHF) portend a poor prognosis.[4,5] Thus, the use of prediction models for identifying patients with diabetes at

Correspondence to: Abhinav Sharma, MD, PhD, Divisions of Cardiology and Experimental Medicine, McGill University, 1001 Decarie Blvrd, Montreal, Qc, H4A 3J1, Canada. Email: abhinav.sharma@mcgill.ca

Supplemental Material for this article is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.121.024833

For Sources of Funding and Disclosures, see page 14.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- Available evidence suggests that several widely available clinical risk tools can robustly prognosticate the risk of heart failure hospitalization in people with type 2 diabetes and identify those who may benefit most from novel guideline-directed medical therapies.
- The effect of most clinical prediction models on clinical outcomes, patient care, provider behaviors has largely been under investigated, as a result, health, economic, and clinical investigations of these tools are warranted.

What Are the Clinical Implications?

- Given that several clinical prediction models have demonstrated robust prognostic accuracy, their implementation in clinical settings might facilitate the identification of high-risk individuals with type 2 diabetes who may benefit most from guideline-directed therapies such as sodium-glucose co-transporter 2 inhibitors.

Nonstandard Abbreviations and Acronyms

ACCORD	Action to Control Cardiovascular Risk in Diabetes trial
HHF	hospitalization for heart failure
PROBAST	Prediction Model Risk of Bias Assessment Tool
RECODe	Risk Equations for Complications of Type 2 Diabetes
TRS-HF$_{DM}$	Thrombolysis in Myocardial Infarction Risk Score for HF in Diabetes
WATCH-DM	(Weight [BMI], Age, Hypertension, Creatinine, HDL-C, Diabetes Control [fasting plasma glucose], QRS Duration, MI, and CABG) risk score

particularly high risk for developing HF and HHF has become clinically important.

Clinical prediction models and risk prediction can facilitate shared decision-making, recruitment into clinical trials, and the selection of patients who may benefit most from therapies that reduce the risk of HHF events (eg, sodium-glucose cotransporter 2 inhibitors).[6,7] In addition, such models may be used to support clinical trial design and cost-effectiveness studies by simplifying risk stratification.[8] Although several studies have developed prediction models for HHF in people with T2D, no models have yet been incorporated in guideline-directed care. Therefore, our review aimed to evaluate the performance, applicability, and clinical impact of existing clinical prediction models for HHF in adults with T2D to facilitate the selection of models for clinical implementation.

METHODS

The authors declare that all supporting data are available within the article. This review was conducted according to a pre-specified protocol that was developed with clinical experts and registered in the Open Science Framework (osf.io/na26x).[9] The review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses and the Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies guidelines.[10,11]

Search Strategy

We searched Embase (Ovid), MEDLINE (Ovid), Web of Science, Google Scholar (first 200 citations), and Tufts' Clinical Prediction Model Registry from database inception to February 24, 2021. The complete search strategy, detailed in Table S1, was developed and peer-reviewed with 2 academic librarians following the Peer Review of Electronic Search Strategies guidelines.[12] Briefly, a 3-concept search strategy using medical subject headings, Emtree terms, and keywords related to T2D, HF, and prediction modeling were applied. Additional citations were included by consulting content experts and reviewing the reference and citation lists of included studies (Web of Science), related reviews, and conference proceedings. No restrictions were placed on the language, date, or status of publications.

Study Selection and Eligibility Criteria

The titles and abstracts of retrieved citations were screened against prespecified inclusion criteria by 2 independent reviewers (A.R., E.O). References marked as potentially eligible proceeded to secondary full-text assessment. Disagreements were resolved by consensus or when needed, by a third reviewer (A.S.). Eligible studies needed to report the development, validation, or update of a multivariable prediction model for HHF with or without the competing risk of death in patients with T2D (>90% population prevalence). Because our review aimed to identify models apt for clinical use, eligible studies also needed to report measures of model performance (≥1) and sufficient information for clinical use. Eligible performance measures inlcuded but were not limited to assessments of model accuracy,

concordance, Brier score, sensitivity, specificity, discrimination, calibration (eg, Hosmer–Lemeshow test), or R^2. Sufficient information for clinical use constituted regression coefficients or measures of association (eg, hazard ratios) for quantifiable and non-arbitrary variables. Studies not reporting original data (ie, reviews) were excluded.

Data Extraction

Two reviewers (A.R., E.O.) independently extracted the characteristics of the studies and cohorts used to develop, validate, or update eligible models. Extracted study and model characteristics included the first author, year of publication, model data source, model derivation and validation methods, number of predictors screened, variables included in the final model, measures of model performance, outcome details (ie, time horizon, definition), and model presentation. Extracted cohort characteristics included patients' geographic region, age, sex, comorbidities (ie, HF, coronary artery disease), follow-up period, and number of cumulative events. In instances where multiple models were reported in a single study, data were extracted from the best performing model. Discrepancies between the reviewers' results were resolved through consensus or a third reviewer (A.S.).

Model Evaluation

Included studies were evaluated with the Prediction Model Risk of Bias Assessment Tool (PROBAST).[13] PROBAST comprises 4 domains designed to identify methodological limitations in model development and validation with respect to selected participants, predictors, outcomes, and analyses. PROBAST also includes domains designed to assess the applicability of developed models with respect to the included participants, predictors, and outcomes. As the purpose of this review was to identify models apt for clinical use, many variables (>7) and the inclusion of continuous variables were separately considered as barriers to routine use after discussions with clinical experts. Based on the PROBAST domains, potential concern with the risk of bias or applicability of primary studies were classified as either high, low, or unclear. Risk of bias assessments were done in duplicate by 2 independent reviewers (A.R., E.O.), and disagreements were resolved by consensus, or a third reviewer as needed (A.S.).

Performance Measures and Model Validation

A clinical prediction models' performance may be evaluated in internal- or external validation through measures of discrimination and calibration. Discrimination reflects a model's ability to distinguish between patients who do and do not experience an outcome of interest.[14] Calibration conversely reflects a model's predictive accuracy: ie, the agreement between the predicted probability of events and the actual proportion of events observed.[14] A detailed explanation of internal- and external validation and commonly used measures that evaluate these metrics is available in Data S1.

Data Synthesis Statistical Analysis

We constructed evidence tables with details of the identified clinical prediction models and their derivation, validation, or updated (eg, presentation, performance, included variables, risk of bias, and applicability). For the analysis of overall discrimination, meta-analyses were performed using random-effect models in instances where multiple c-statistics (≥3) were available for the same prediction model in internal- or external validation. For primary studies where meta-analysis was not indicated, the results were synthesized qualitatively using a narrative approach. Furthermore, in response to expert review, our protocol was modified to synthesize the modeling studies according to HHF and new-onset HHF, respectively. Statistical analyses were performed using R version 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

The literature search identified 6192 citations (Figure 1). After screening, a total of 18 studies[15–32] published between 2008 and 2021 were included in the review, with data from 999 167 unique patients with T2D. Of these 18 studies,[15–32] 2 were identified by screening reference lists[20] and conference abstracts,[30] respectively. A flow diagram of study selection is shown in Figure 1 and the list of excluded full-text citations is available in Table S2.

Study and Model Characteristics

Together 15 multivariable prediction models[15–29] were assessed as 3 studies[30–32] exclusively performed external validation. Most modeling studies used randomized controlled trials (9 of 15)[15–17,19–22,25,28] for model development, 3 of which were the ACCORD (Action to Control Cardiovascular Risk in Diabetes) trial.[20,21,33] The median number of participants and events used for model development was 8756 (interquartile range, 5184–16 013) and 258 (interquartile range, 223–420), respectively (Table 1). The median number of included variables was 10 (interquartile range, 7–12) with the most common being age (12 of 15), body mass index (8 of 15), and systolic blood pressure (8 of 15; Figure 2). A minority of models incorporated race and ethnicity (4

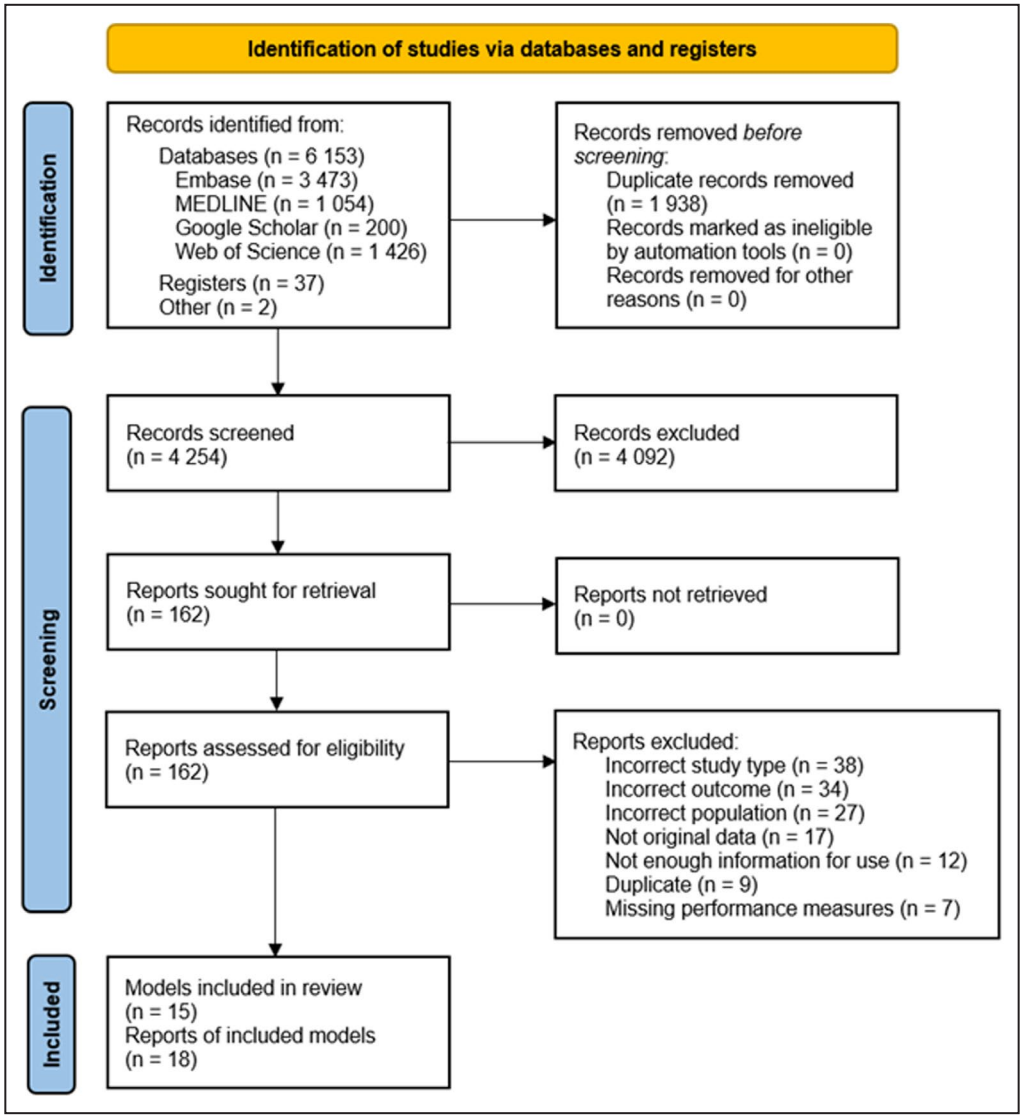

Figure 1. PRISMA flow diagram for study selection.
PRISMA indicates Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

of 15) or sex (3 of 15); estimated glycated hemoglobin was more common in models predicting incident HHF (4 of 6) whereas glomerular filtration rate was more common in models predicting incident or recurrent HHF (5 of 9). Tables 2 and 3 summarize the complete details of the included models and their internal and external performance.

Models With Clinical Utility and External Validation

Of the 15 multivariable models,[15–29] 7 were presented in a useful manner (eg, risk score, online calculator).[17,20,21,24,25,27,28] Of these, 5 had their performance evaluated in external validation (Table 3).[17,20,21,25,27] These were the Thrombolysis in Myocardial Infarction (TIMI) Risk Score for HF in Diabetes (TRS-HF_{DM})[17]; the BRAVO (Building, Relating, Assessing, and Validating Outcomes) risk engine[20]; the Risk Equations for Complications of Type 2 Diabetes (RECODe)[21]; the Weight, Age, Hypertension, Creatinine, HDL-C, Diabetes Control, QRS Duration, MI (myocardial infarction), and CABG (WATCH-DM) risk score[25]; and QDiabetes.[27] Two models strictly predicted incident HHF,[25,27] whereas the other 3 predicted incident or recurrent HHF in patients with and without prevalent HF (Table 1).[17,20,21]

TRS-HF_{DM} Risk Score

The TRS-HF_{DM} was developed in SAVOR-TIMI 53 (Saxagliptin Assessment of Vascular Outcomes Recorded in Patients with Diabetes Mellitus-TIMI 53) trial (Table 1)[17] and externally validated in 3 trials (N=19 081; Table 3).[17,30,31] The development cohort had a mean age of 65.0 years and a median follow-up time of 2.1 years; the prevalence of HF and coronary

Table 1. Characteristics of the Cohorts and Methods Used in Studies Reporting the Development of a Clinical Prediction Model for Heart Failure Hospitalization in Type 2 Diabetes

Reference (Model)	Data source	Follow-up (y)	No.*	Country (n)	Age (mean), y	HF (%)	CAD (%)	Model derivation	Variables screened	Outcome	No. events
Heart failure hospitalization											
Willis 2021[15]	CANVAS	3.6	10 142	Multinational (30)	63.3	14.4	56.4	Weibull regression	52	HHF	243
Sharma 2020[16]	EXAMINE	1.6	5154	Multinational (49)	61.0	27.9	100	Cox regression	NR	HHF	195
Berg 2019[17] (TRS-HF_{DM})	SAVOR-TIMI 53	2.1	8212	Multinational (26)	65	12.8	62.4	Cox regression	25	HHF	228
Kim 2019[18]	EMR	5.0	81 091	United States	60.4	7.0	22.0	Multi-task learning	45	HHF	NR
Fraty 2018[19]	SURDIAGENE	5.3	1438	France	65.0	NR	26.7	Fine and Gray regression	24	HHF	206
Shao 2018[20] (BRAVO)	ACCORD	4.7	10 251	United States and Canada	62.8	4.8	35.2	Weibull regression	28	HHF or HF death	454
Basu 2017[21] (RECODe)	ACCORD	4.7	9635	United States and Canada	62.8	4.8	35.2	Cox regression	33	HHF or HF death	454
Wolsk 2017[22]	ELIXA	2.2	5525	Multinational (49)	60.3	22.4	100	Cox regression	45	HHF	221
Kiadaliri 2013[23]	EMR	5.0	21 775	Sweden	56.1	NR	NR	Weibull regression	11	HHF	I: 1366 R: 947
Incident heart failure hospitalization											
Williams 2020[24]	EMR	6.6	54 452	United States	60.0	0.0†	21.0	Cox regression	80	New-onset HHF	1884
Segar 2019[25] (WATCH-DM)	ACCORD	4.9	8756	United States and Canada	62.7	0.0†	35.2	Random survival forests	147	New-onset HHF or HF death	319
Halon 2017[26]	Cohort study	8.4	735	Israel	63.4	0.0	0.0	Cox regression	39	New-onset HHF or cardiovascular death	41
Hippisley-Cox 2015[27] (QDiabetes)	EMR	15.0	437 806	England	60.0	0.0†	17.4	Cox regression	21	New-onset HHF	274
Pfister 2011[28]	PROactive	2.9	5238	Multinational (19)	61.7	0.0†	94.7	Cox regression	34	New-onset HHF or HF death	233
Yang 2008[29]	EMR	5.5	3456	China	57	0.0†	4.4	Cox regression	26	New-onset HHF	274

ACCORD indicates Action to Control Cardiovascular Risk in Diabetes trial; CAD, coronary artery disease; CANVAS, Canagliflozin Cardiovascular Assessment Study; DECLARE-TIMI 58, The Dapagliflozin Effect on Cardiovascular Events–Thrombolysis in Myocardial Infarction 58 trial; ELIXA, Evaluation of Lixisenatide in Acute Coronary Syndrome trial; EMR, electronic medical records; EXAMINE, Examination of Cardiovascular Outcomes With Alogliptin Versus Standard of Care trial; HF, heart failure; HHF, hospitalization for heart failure; I, incident HF; NR, not reported; PROactive, Prospective Pioglitazone Clinical Trsial in Macrovascular Events trial; R, recurrent HF; SURDIAGENE, Survival Diabetes and Genetics cohort; and WATCH-DM, Weight [BMI], Age, Hypertension, Creatinine, HDL-C, Diabetes Control [fasting plasma glucose], QRS Duration, MI, and CABG) risk score.

*Training data set sample size.

†Excluded from training data set.

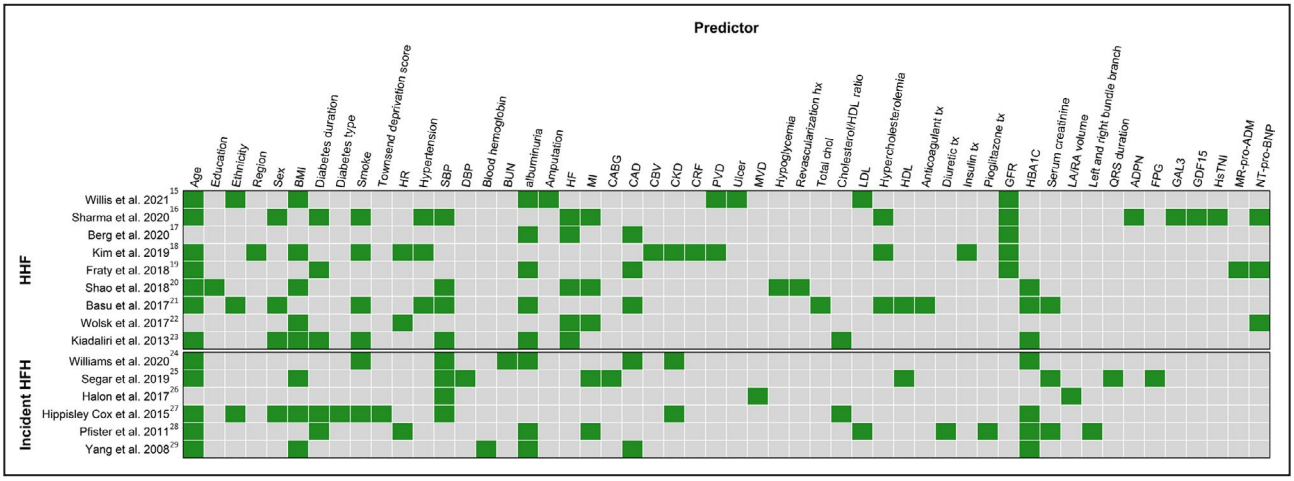

Figure 2. Matrix of risk predictors for heart failure hospitalization in included model development studies.
Afib indicates atrial fibrillation; BMI, body mass index; BUN, blood urea nitrogen; CABG, coronary artery bypass grafting; CAD, coronary artery disease; CBV, cerebrovascular disease; CKD, chronic kidney disease; Cr, Creatinine; CRF, chronic renal failure; DBP, diastolic blood pressure; FPG, free plasma glucose; GAL3, Galectin-3; GDF-15, Growth-Differentiation-Factor-15; GFR, Glomerular Filtration Rate; HBA1c, glycated hemoglobin A1c; HDL, high-density lipoprotein; HHF, hospitalization for heart failure; HF, heart failure; HsTNI, high-sensitivity troponin; LA/RA, left atrium / right atrium; LDL, low-density lipoprotein; MR-pro-ADM, Mid-regional pro-ADM; NT-proBNP, N-terminal pro-B-type natriuretic peptide; PVD, peripheral vascular disease; and SBP, systolic blood pressure.

artery disease was 12.8% and 62.4%, respectively. The TRS-HF$_{DM}$ is an integer-based risk score between 0 to 7 that uses 5 clinical variables (Table 2) and predicts the 4-year risk of incident or recurrent HHF. Proteomic biomarkers were purposely excluded to prioritize parsimony and ease-of-use. Good discriminative performance was observed in internal validation (c-statistic: 0.81; Table 2).[17] In addition, external validation studies showed moderate-to-good discrimination and calibration for 0.5-, 2.5-, 4- and 7-year event predictions (Table 3).[30,31] The overall c-statistic was 0.78 (95% CI, 0.76–0.80; Figure 3). However, 1 external validation study demonstrated the model modestly overestimated the risk of events for individuals with diabetes and acute coronary syndrome who are classified at severe risk.[30]

WATCH-DM Risk Score

The Weight, Age, Hypertension, Creatinine, HDL-C, Diabetes Control, QRS Duration, MI, and CABG risk score was developed in the ACCORD trial (Table 1) and externally validated in 1 trial (N= 10 819; Table 3).[25] The development cohort had a mean age of 62.7 and a median follow-up time of 4.9 years; the prevalence of HF and coronary artery disease was 0% and 35.2%, respectively. The model predicts the 5-year risk of incident HHF and is available as an online calculator as well as an integer-based risk score that can range between 0 to 34. The model used machine learning for development, and it included 7 multinomous clinical variables (Table 2). In external validation, the online calculator had a c-statistic of 0.74 and Hosmer-Lemeshow χ^2 of 11.1 (P=0.20) while the risk score had a c-statistic of 0.70 and Hosmer-Lemeshow χ^2 of 10.0 (P=0.29). Markedly, the risk score showed better performance for incident HF with reduced ejection fraction versus preserved ejection fraction in external validation (c-statistic, 0.72 versus 0.64, respectively, P<0.001).

Building, Relating, Assessing, and Validating Outcomes Risk Engine

The Building, Relating, Assessing, and Validating Outcomes risk engine was developed in ACCORD and externally validated in 3 trials (N = 16 388; Table 3).[20] The development cohort had a mean age of 62.8 and a median follow-up time of 4.7 years; the prevalence of HF and coronary artery disease was 4.8% and 35.2%, respectively. Building, Relating, Assessing, and Validating Outcomes predicts the annual risk of incident or recurrent HHF alongside stroke, MI, angina, revascularization, severe pressure loss, end-stage renal disease, blindness, all-cause mortality, and cardiovascular death. The model is accessible as an online calculator that requires 18 variables, 9 of which are for HHF (Table 2). Internal validation demonstrated good performance: for HHF the Brier score was 0.008 and the c-statistic was 0.79 (95% CI, 0.77 to 0.81). In external validation, the risk engine demonstrated good calibration across the 28 end points (calibration slope: 1.07; R^2: 0.86), however, HHF specific calibration was not reported.[20]

RECODe Risk Equations

RECODe was developed in ACCORD and validated in 3 cohorts (N = 8 061; Table 3).[21,32] The development

Table 2. Characteristics and Internal Performance of Clinical Prediction Models for Heart Failure Hospitalization in Type 2 Diabetes

Reference (Model)	Model presentation	Model variables (n)	Internal validation	Time horizon	Internal Model Performance		Heart failure hospitalization definition†	Risk of bias|applicability‡
					Calibration	Discrimination*		
Heart Failure Hospitalization								
Willis 2021[15]	Regression coefficients	Age, race and ethnicity, BMI, eGFR, LDL, ulcer history, AF, PVD, micro-or macroalbuminuria, amputation history (10)	Apparent	7-y risk	O/E ratio: 0.98	0.83	Hospitalization >24 hours with new or worsening clinical and physical signs of HF with need for therapy.	High|High§
Sharma 2020[16]	Regression coefficients	NT-proBNP for consistency , HF history, hsTNI, GDF15, Hypercholesterolemia, hypertension, MI history, DM duration, GAL3, SBP, age, sex, smoke, adiponectin, eGFR (15)	Apparent	6-mo risk	NR	0.83	Hospitalization >12 hours with new or worsening clinical, radiological, or physical signs of HF with need for therapy.	High|High
Berg 2019 (TRS-HF_{DM})[17]	Integer score and online calculator	HF history, AF, CAD, eGFR, uACR (5)	Bootstrapping	4-y risk	NR	0.81	Hospitalization >12 hours with evidence of new or worsening HF and need for additional or increased therapy.	High|Low
Kim 2019[18]	Regression coefficients	GFR, normal GFR, BMI, HR, smoke, age, region, hypertension therapy, Hypercholesterolemia treatment, insulin treatment, CAD, CBV, PVD, CKD, CRF (15)	Bootstrapping	5-y risk	NR	0.81	Hospital discharge with *ICD-9* codes: 428.0, 428.1, 428.2, 428.3, 428.4, 428.9.	High|High
Fraty 2018[19]	Regression coefficients	Age, DM duration, eGFR, uACR, CAD, MR-proADM, NT-proBNP (7)	Apparent	NR	NR	0.84	Definition as provided by the ESC guidelines 2012.	High|High
Shao 2018[20] (BRAVO)	Online calculator	HbA1c, SBP, BMI, age, hypoglycemia, education, MI history, HF history, revascularization history (9)	Cross-validation	NA¶	Brier score: 0.008	0.79	See definition provided in Segar 2019 above.	High|Low
Basu 2017[21] (RECODe)	Online calculator	Age, sex, race and ethnicity, smoke, SBP, CAD, hypertension therapy, statin treatment, anticoagulant treatment, HbA1c, TC, HDL, Serum creatinine, uACR (14)	Cross-validation	10-y risk	Calibration slope/ intercept/p: 1.01/–0.0004/0.93	0.75	See definition provided in Segar 2019 above.	Low|Low
Wolsk 2017[22]	Regression coefficients	NT-proBNP, BMI, NSTEMI, HF history, MI history (5)	Apparent	NR	NR	0.77	Hospitalization with new or worsening clinical and physical signs of HF with need for therapy.	High|High
Kiadaliri 2013[23]	Regression coefficients	Sex, age, HbA1c, SBP, TC/HDL, BMI, micro/macroalbuminuria, smoke, HF history, DM duration (10)	Split-sample	5-y risk	HL test Training: *P*=0.30 Test: *P*=0.14	Training: 0.84 Test: 0.84	Fatal or non-fatal HF hospitalization with *ICD-10* code I50.	High|High§
Incident heart failure hospitalization								
Williams 2020[24]	Integer score	Age, CAD, BUN, AF, HbA1c, albumin, SBP, CKD, smoke (9)	Apparent	1, 3, 5-y risk	NR	0.78	EMR documented hospital admission with HF as the primary diagnosis in the absence of prior HF diagnosis.	High|High

(*Continued*)

Table 2. Continued

Reference (Model)	Model presentation	Model variables (n)	Internal validation	Time horizon	Internal Model Performance		Heart failure hospitalization definition†	Risk of bias|applicability‡
					Calibration	Discrimination*		
Segar 2019[25] (WATCH-DM)	Online calculator & integer score	Age, BMI, SPB, FPG, QRS, Serum Cr, DBP, HDL-C, MI history, CABG history for consistency (10)	Split-sample	5-y risk	HL test: *P*=0.29	Training: 0.74 Test: 0.77	Hospitalization with clinical and radiologic evidence of HF; or death due to HF or cardiogenic shock with evidence of HF in the absence of acute ischemic event	High|High
Halon 2017[26]	Regression coefficients	LA/RA volume ratio, microvascular disease, SBP (3)	Apparent	NR	NR	0.79	Hospitalization requiring >1 of: typical HF symptoms and findings on examination, dyspnea and radiological evidence, or dyspnea and HF diagnosis requiring IV therapy with furosemide.	High|High
Hippisley-Cox 2015[27] (QDiabetes)	Online calculator	Age, sex, deprivation, race and ethnicity, smoke, DM type, DM duration, CKD, AF, HbA1c, TC/HDL, SBP, BMI (13)	Split-sample	10-y risk	R^2: 39.8 - 40.0‖	0.77 - 0.77‖	Primary care record codes: G58%, G5yy9, G5yyA, 662f, 662g, 662h, 662i; ICD-10 codes: I110, I130, I42, I50 for cases of HF from hospital and mortality records.	High|High
Pfister 2011[28]	Integer score	Age, total creatinine, diuretic treatment, HbA1c, DM duration, LDL, HR, left BBB, right BBB, MI also history for consistency, microalbuminuria, pioglitazone treatment (12)	Bootstrapping	NR	HL test *P*=0.62	0.75	Serious HF event requiring hospitalization or prolongation of stay, was fatal or life-threatening, or resulted in significant disability.	High|High
Yang 2008[29]	Regression coefficients	Age, BMI, HBA1c, uACR, hemoglobin, CAD (6)	Split-sample	5-y risk	HL test *P*>0.20	0.85	Hospital discharge with ICD-9 code 428.X.	High|High

Model validation methods and model performance measures are delineated according to internal- and external validation. AF indicates atrial fibrillation; BBB, bundle branch block; BMI, body mass index; BUN, blood urea nitrogen; CABG, coronary artery bypass grafting; CAD, coronary artery disease; CBV, cerebrovascular disease; CKD, chronic kidney disease; CRF, chronic renal failure; DBP, diastolic blood pressure; EMR, electronic medical records; FPG, free plasma glucose; GAL3, Galectin-3; GDF-15, growth-differentiation factor-15; GFR, glomerular filtration rate; HBA1c, glycated hemoglobin A1c; HDL, high-density lipoprotein; HF, heart failure; HL, Hosmer-Lemeshow; HsTNI, high-sensitivity troponin; *ICD-9, International Classification of Diseases, Ninth Revision*; *ICD-10, International Classification of Diseases, Tenth Revision;* LA/RA, left atrium/right atrium; LDL, low-density lipoprotein; MI, myocardial infarction; MR-pro-ADM, Mid-regional pro-ADM; NSTEMI, non–ST-segment–elevation myocardial infarction; NT-proBNP, N-terminal pro-B-type natriuretic peptide; O/E ratio, predicted mean event rate/observed mean event rate; PVD, peripheral vascular disease; SBP, systolic blood pressure; TC, total cholesterol; uACR, urine albumin to creatinine ratio; and WATCH-DM, (Weight [BMI], Age, hyperTension, Creatinine, HDL-C, Diabetes control [fasting plasma glucose], QRS Duration, MI, and CABG) risk score.

*Discrimination is measured as c-statistics unless specified otherwise.

†The definition for heart failure hospitalization derived from the training set used for model development.

‡High and low ratings relate to the degree of concern associated with the risk of bias or applicability. These domains were assessed with the Prediction Model Risk of Bias Assessment Tool.

§Downgraded to high concern with applicability because the study only reported regression coefficients.

‖Range across men and women, respectively.

¶Time horizon is not applicable as the model is a discrete-time patient-level microsimulation.

cohort had a mean age of 62.8 and a median follow-up time of 4.7 years; the prevalence of HF and coronary artery disease was 4.8% and 35.2%, respectively. The risk equation is available as an online calculator that predicts incident or recurrent HHF alongside the 10-year risk of nephropathy, retinopathy, neuropathy, myocardial infarction or stroke, and all-cause mortality. The online calculator requires 16 clinical variables, 14 of which are for HHF (Table 2). RECODe showed moderate-to-good discrimination and calibration in internal and external validation. In ACCORD, the c-statistic was 0.75 and the calibration slope was 1.01 (*P*=0.93 [insignificant p-values indicate acceptable calibration]).[21] In external validation, the c-statistic ranged between 0.73 to 0.80 and the calibration slopes ranged between 0.72 to 1.13 for the 10-year estimated risk of HHF (Table 3).[32] The overall c-statistic for RECODe was 0.76 (95% CI, 0.73−0.79; Figure 3).

QDiabetes Risk Calculator

QDiabetes was developed using primary care data from 437 806 people and then validated in 2 cohorts (N = 334 933, Table 3).[27] The development cohort had a mean age of 60.0 and a follow-up time of 15 years; the prevalence of HF and coronary artery disease was 0% and 17.4%, respectively. The risk calculator is available as a sex-specific online calculator that predicts the 10-year risk of incident HHF. The risk calculator uses 12 variables (Table 2). The model demonstrated moderate performance in internal and external validation. In internal validation the c-statistic was 0.76 and 0.77 in men and women, respectively, whilst in external validation the c-statistic was 0.77 and 0.78. Calibration, which was assessed visually with calibration plots, was satisfactory as there was good agreement between the predicted and observed risks in each validation cohort. The overall c-statistic for the QDiabetes Risk Calculator was 0.77 (95% CI, 0.76−0.78; Figure 3).

Models With Incomplete Evaluation

The 10 remaining clinical prediction models either never underwent external validation (9 of 10)[15,16,19,22–24,26,28,29] or were externally validated but not presented in a clinically practical form (1 of 10; Table 1).[18] Therefore, these models' clinical utility was judged as less applicable for the review. As an exception, 2 models were presented as integer-based risk scores that could be relevant in certain settings.[24,28] The first model by William and colleagues used routine care data to predict the 1-, 3- and 5-year risk of HHF.[24] The model used 9 variables (Table 2) and had moderate discrimination in patients without prior HF (c-statistic: 0.78). The second model by Pfister and colleagues used data from PROactive (Prospective Pioglitazone Clinical Trial in Macrovascular Events) to develop an 11 variable risk score (Table 2).[28] Internal model performance with bootstrap re-sampling was moderate (c-statistic: 0.75): however, there were some concerns with applicability (eg, inclusion of variables with limited availability).

Risk of Bias and Applicability

The most prevalent sources of potential bias in model development included use of unblinded outcome adjudication (12 of 15),[15–25,29] handling of missing data (11 of 15),[1] management of competing risk (8 of 15),[15,17,22–24,27–29] and consideration of model overfitting (8 of 15).[15,16,19,22–24,26,27] Applicability concerns were model predictor accessibility (6 of 15)[16,19,22,25,26,28] and generalizability of the derivation cohorts (6 of 15).[16,18,24,25,28,29] In validation, sources of bias included unblinded outcome adjudication (8 of 9)[17,18,20,21,25,30–32] and evaluation of model performance (4 of 9).[17,18,25,31] All but one model (RECODe) by Basu and colleagues[21] had potential for risk of bias during development, and all but 3 models[17,20,21] had potential applicability concerns (Figures S1 through S4).

DISCUSSION

This review was designed to evaluate existing clinical prediction models for HHF in adults with T2D to facilitate clinical model selection. Only studies reporting performance measures and sufficient information for model use were included to highlight tools with the most clinical utility. Altogether 15 models were identified, of which only 5 were externally validated and presented in a clinically practical form. Most identified models (n=15) included >7 variables (75%) or only reported regression coefficients (53%), limiting clinical applicability. Moreover, most identified clinical prediction models had potential concerns with risk of bias (93%) or applicability (67%), highlighting the need for improved methods in modeling studies.

Prieto-Merino and Pocock have proposed that 3 features of risk models should be favored: (1) relative simplicity with reasonably easy-to-obtain variables; (2) clinical relevance in the context of the disease state; and (3) overall generalizability to other settings. Models with these characteristics should be favored compared with complex mathematical models.[34] Based on these criteria and our appraisal of the included models' risk of bias, the RECODe risk equations[21] and the TRS-HF_{DM}[17] demonstrated the most clinical potential. Both models were easy-to-use (eg, as risk scores and online calculators) and were externally validated (ie, generalized) in 3 large cardiovascular safety trials. Although RECODe was the only model with low potential for risk of bias, it requires the input of 16 variables, which restricts its potential for routine use. Nevertheless, RECODe provides risk estimates for several diabetes-specific

*References 15,16,18–20,22,23,25,26,28,29

Table 3. External Validity of Identified Clinical Prediction Model for Heart Failure Hospitalization in Type 2 Diabetes

Reference	Data Source	No.	Countries (n, sites)	Age, y*	HF (%)	CAD (%)	External validation	Event time horizon	External model performance		Heart failure hospitalization definition‡	Risk of bias|applicability§
									Calibration	Discrimination†		
TRS-HF$_{DM}$												
Berg 2020[17]	DECLARE-TIMI 58	8578	Multinational (882)	63.9	10.0	40.6	Temporal	4-y risk	Nam D'Agostino: *P*=0.20	0.78	Hospitalization >24 hours with new or worsening clinical and physical signs of HF with need for therapy.	High|Low
Elharram 2020[31]	ACCORD	5123	United States and Canada (77)	62.7	4.9	35.2	Independent	7-y risk	Nam D'Agostino: *P*=0.13	0.78	See ACCORD's definition in Table 2.	High|Low
Razaghizad 2021[30]	EXAMINE	5380	Multinational (898)	61.0	27.9	100	Independent	6- and 30-mo risk	Calibration slope/ intercept/p:0.81/–0.17/0.33; 0.77/–0.18 /0.06	0.75	See EXAMINE's definition in Table 2.	High|High
RECODe												
Basu 2017[21]	Look AHEAD	4760	United States (16)	57.5	NR	14.0	Independent	10-y risk	Calibration slope/ intercept/p: 1.13/–0.011/0.07	0.76	Hospitalization with clinical and radiologic evidence of HF with need for therapy or ventricular dysfunction.	Low|Low
Basu 2018[32]	MESA	1555	United States (6)	63.0	0.0‖	NR	Independent	10-y risk	Calibration slope/ intercept/p: 1.01/0.005/0.42	0.80	*ICD-9* code 428 discharge or underlying cause of death I50; radiologic evidence of HF; or autopsy finding of pulmonary edema or HF.	Low|Low
Basu 2018[32]	JHS	1746	United States (3)	57.5	NR	NR	Independent	10-y risk	Calibration slope/ intercept/p: 0.72/0.091/0.07	0.73	Hospitalization with clinical and radiologic evidence of HF with need for therapy, and dilated ventricle or poor heart function assessed by echocardiography or ventriculography.	
QDiabetes												
Hippisley-Cox 2015[27]	CPRD	197 905	United Kingdom (254)	61.0	0.0‖	19.1	Geographic	10-y risk	R^2: 41.1 – 38.7¶	0.77 – 0.78¶	See Hippisley-Cox's definition in Table 2.	High|Low
WATCH-DM												
Segar 2019[25]	ALLHAT	10 819	United States and Canada (623)	67.0	0.0‖	NR	Temporal	5-y risk	HL test: *P*=0.20	0.74	HHF was primarily based on clinic investigator reports.	High|Low

(*Continued*)

Table 3. Continued

Reference	Data Source	No.	Countries (n, sites)	Age, y*	HF (%)	CAD (%)	External validation	Event time horizon	External model performance		Heart failure hospitalization definition‡	Risk of bias|applicability§
									Calibration	Discrimination†		
BRAVO												
Shao 2018[20]	ASPEN	2410	Multinational (70)	61.1	0.0‖	NR	Geographic	NR	NR	NR	NR	High|High
Shao 2018[20]	ADVANCE	11 140	Multinational (215)	66.3	NR	NR	Geographic	NR	NR	NR	Death attributable to HF, HHF, or worsening New York Heart Association class.	
Shao 2018[20]	CARDS	2838	United Kingdom (132)	61.7	NR	NR	Geographic	NR	NR	NR	NR	

ACCORD indicates Action to Control Cardiovascular Risk in Diabetes trial; ADVANCE, the Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified Release Controlled Evaluation trial; ALLHAT, the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack trial; ASPEN, Atorvastatin Study for Prevention of Coronary Heart Disease Endpoints in Non-Insulin-Dependent Diabetes Mellitus trial; CARDS, the Collaborative Atorvastatin Diabetes Study trial; CPRD, Clinical Research Practice Datalink database; DECLARE-TIMI 58, The Dapagliflozin Effect on Cardiovascular Events–Thrombolysis in Myocardial Infarction 58 trial; EXAMINE, Examination of Cardiovascular Outcomes With Alogliptin Versus Standard of Care trial; HF, heart failure; HHF, hospitalization for heart failure; HL, Hosmer-Lemeshow; JHS, Jackson Heart Study; Look AHEAD, Action for Health in Diabetes trial; MESA, The Multi-Ethnic Study of Atherosclerosis trial; NR, not reported; and WATCH-DM, (Weight [BMI], Age, Hypertension, Creatinine, HDL-C, Diabetes Control [fasting plasma glucose], QRS Duration, MI, and CABG) risk score.

*Mean or median as available in the publication.

†Discrimination is measured as c-statistics unless specified otherwise.

‡The definition for heart failure hospitalization derived from the data set used for model validation.

§High and low ratings relate to the degree of concern associated with the risk of bias or applicability. These domains were assessed with the Prediction Model Risk of Bias Assessment Tool.

‖Patients with prevalent heart failure excluded.

¶Range across men and women, respectively.

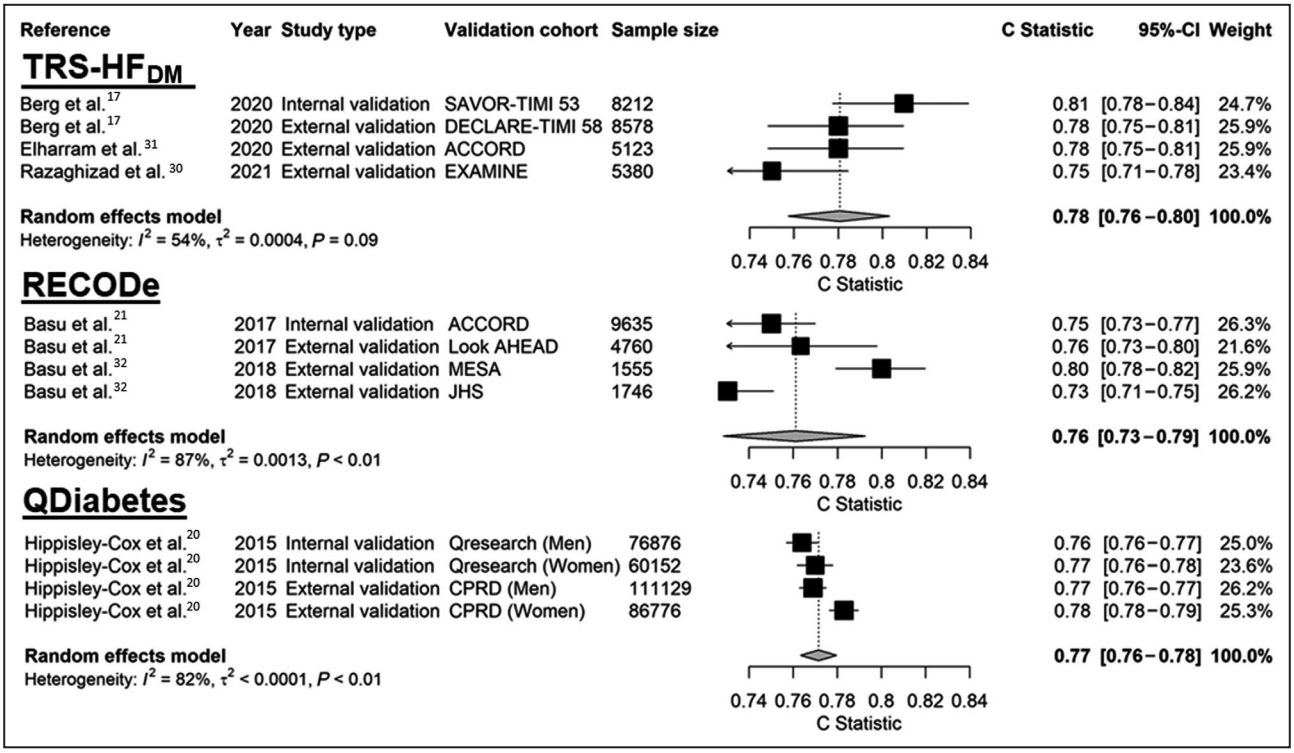

Figure 3. Meta-analysis of externally validated clinical prediction models' discrimination.
ACCORD indicates Action to Control Cardiovascular Risk in Diabetes trial; CPRD, Clinical Research Practice Datalink database; DECLARE-TIMI 58, The Dapagliflozin Effect on Cardiovascular Events–Thrombolysis in Myocardial Infarction 58 trial; EXAMINE, Examination of Cardiovascular Outcomes With Alogliptin versus Standard of Care trial; JHS, Jackson Heart Study; Look AHEAD, Action for Health in Diabetes trial; MESA, The Multi-Ethnic Study of Atherosclerosis trial; RECODe, Risk Equations for Complications of Type 2 Diabetes; SAVOR-TIMI 53, Saxagliptin Assessment of Vascular Outcomes Recorded in Patients with Diabetes Mellitus–Thrombolysis in Myocardial Infarction 53 trial; and TRS-HF$_{DM}$, Thrombolysis in Myocardial Infarction (TIMI) Risk Score for Heart Failure in Diabetes.

complications, making it potentially apt for comprehensive metabolic, cardiovascular, and renal risk prediction.[21] The TRS-HF$_{DM}$, in contrast, was one of the simplest model for HHF as it was an integer-based risk score that only required 5 common variables. Berg and colleagues, who developed the model, demonstrated it could identify a 20-fold risk gradient and patients who derive greatest absolute benefit from sodium-glucose co-transporter 2 inhibition.[17,35] Despite these strengths, the TRS-HF$_{DM}$ has yet to be externally validated in a low-risk population-based cohort. Future research may seek to simplify the RECODe risk equations or externally validate the TRS-HF$_{DM}$ in a low-risk non-clinical trial population.

Defining HF Hospitalization

At present, it is unclear if research should focus on developing new models for incident HHF or validating and using models that already predict new-onset HHF alongside recurrent events (eg, RECODe, TRS-HF$_{DM}$). On one hand, future prediction modeling studies may need to focus on incident HHF as the therapeutics options (eg, quadruple therapy, sodium-glucose co-transporter 2 inhibitors) for individuals with diabetes who have prevalent HF are already well established.[36–39] The prediction of new-onset HHF, in addition, may be particularly important as the use of sodium-glucose co-transporter 2 inhibitors to prevent the development HF has traditionally been overlooked compared with the prevention of traditional major adverse cardiovascular events.[40] Therefore, the prediction of new-onset HF could offer more avenues for clinician-patient discussions to facilitate the implementation of therapies that prevent HF development. On the other hand, predicting new-onset or recurrent HHF provides generalizable models that are applicable for a larger segment of individuals with diabetes who are at risk of HF. Ultimately, both avenues present a good path to increase the implementation of guideline-directed medical therapy. However, at present there is no consensus on this issue.

Overcoming Barriers to Adoption

Despite the proliferation of prediction models for HHF in individuals with T2D, few models have been successfully incorporated into routine care. A previous

systematic review of mortality models for patients with HF highlighted deficiencies that have limited the adoption of clinical prediction models. The review, by Alba and colleagues,[41] prompted a former FDA commissioner to critique the fact that most models have insufficient validation and the fact that those with the greatest levels of validation are derived from populations with limited generalizability (eg, clinical trials).[42]

Our review adds to this literature as it shows that after almost 10 years the same problems have been left unresolved. Furthermore, we have identified that few models are applicable for point-of-care use because of either the large number of variables required, or the sole reporting of regression coefficients. As a result, in the absence of electronic medical records that can integrate complex algorithms for decision support,[42] we encourage investigators to develop models that can readily be implemented at the point-of-care.

Our data demonstrated that 80% of clinical predictions models for HHF in T2D included at least one measure of renal function (eg, albuminuria, estimated glomerular filtration rate, serum creatinine, blood urea nitrogen), underscoring its importance in HHF prediction. However, our data also demonstrated that cardiac biomarkers (eg, natriuretic peptides, troponins) may currently be underutilized. As a result, future modeling studies may consider further leveraging cardiac biomarkers to improve predictive performance and thus the impetus for applying predictive models in practice.[43] However, it may be important to develop biomarker- and non-biomarkers versions to facilitate adoption in resource-limited settings.

Methodological and Reporting Issues

As part of the systematic review, a detailed critical appraisal of the risk of bias and applicability of the included studies was performed. This was done to identify potential methodological issues in the conduct of included studies. Our data demonstrated that most studies developing models for HHF did not or were not able to blind outcome adjudication to candidate baseline variables as most studies used data already collected from clinical trials. This was not a major concern in a few studies (ie, those that accounted for model overfitting). However, models which included prior HF as a risk predictor were at high risk of biased predictive accuracy as knowledge of predictors can influence outcome determination.[13,44,45] We acknowledge blinding adjudication to patients' baseline clinical information would likely present forbidding practical challenges. Therefore, in response, we recommend investigators planning to use data from established clinical trials to include statistical methods to address model overfitting.

In the included studies, information relevant to the handling of missing data were also often excluded. Few studies implemented statistical imputation for dealing with missing data and even fewer reported the characteristics of patients with missing data. This can severely affect model validity as patients who are lost to follow-up can differ significantly from the target population.[46,47] Similarly, few studies considered competing risks (eg, death) in model development, which can lead to informative censoring and risk overestimation.[48,49] As the complications of HF and diabetes can disproportionality affect older patients, it is important to use techniques (eg, Fine and Gray regression, cause specific hazard models)[50] to account for competing risk and maximize model performance. To ensure model quality, investigators should reference reporting standards (eg, the Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis) and risk of bias tools (eg, PROBAST) during model development.[13,51]

Strengths and Limitations

This review adhered to numerous best practices for systematic reviews. To ensure transparent reporting and analysis, the protocol was publicly registered in Open Science Framework.[9] A highly sensitive search strategy was also developed to provide a comprehensive overview of existing models. The search included several resources including bibliographic databases, conference proceedings, content experts, and a clinical prediction model registry. Likewise, no restrictions were placed on the date, status, or language of publications. All aspects of the review were also done in duplicate including the data extraction, which followed Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies guidance,[11] and the risk of bias and applicability assessment, which was done with PROBAST.[13] Consequently, to our knowledge, this systematic review constitutes an evidence map of the highest available evidence on clinical prediction models for HHF in adults with T2D to date. The review may thus facilitate evidence-based decisions at the point of care, which may have been hampered in the past by a lack of clarity in the literature.

Despite the strengths, there were limitations with the review. First, most models that we included were judged to have potential concerns with risk of bias or applicability in their development, stemming largely from their methods of analysis and predictor selection. However, this fact was one of the most important findings of the review as it motivated our proposed guidance for future model development. In addition, it is important to highlight that *potential for bias* does not mean that the included models were significantly flawed. Likewise, nor does it mean that studies with fewer potential concerns are more valid than studies with multiple potential flaws.

Second, relevant studies that did not mention or did not include index terms for T2D may have been missed. Third, as noticed in this review, the clinical utility or net benefit of the identified models was not evaluated in any of the studies. Therefore, the effect of these models on real patient outcomes remains unclear and warrants evaluation in external validation and impact studies. Finally, because only 1 study[21] was judged at low risk of bias in development, no association between predicted outcomes and methodological quality could be inferred.

Future Directions

Clinical prediction models may improve health outcomes and resource usage only to the degree that they affect individual patients' or health care providers' behaviors. As mentioned, no studies evaluated the effect of an eligible model on behavioral change or patient events. As a result, cluster randomized controlled trials evaluating the impact of the identified models may be warranted. However, because clinical trials can be expensive, other study designs may be leveraged to conduct initial feasibility assessments. For instance, studies assessing health care providers' judgments before-and-after being presented a model could be a cost-effective study design for such means.[52] Likewise, decision curve analysis, a method for evaluating prediction models, may be used to evaluate potential net benefit (eg, number of unnecessary treatments avoided) because of attributable to risk prediction.[53,54] Although a detailed explanation of the methodology is outside the scope of this review, positive findings compared with standard-of-care could inform the implementation of risk tools into guideline-directed T2D management.

Lastly, 2 studies by Segar and colleagues developed machine learning models according to sex, race and ethnicity, and HF subtype.[25,33] Their data showed risk prediction models may yield significantly different results between these groups. Therefore, future model development and validation studies should aim to validate model performance according to sex, race and ethnicity, and HF subtype as the risk factors and approaches to mitigate risk may differ.[25] As sex was under-represented in the prediction models included in this review, it will be particularly important for future studies to underscore the role of sex as T2D confers an excess of risk for HF in women than in men.[55]

CONCLUSIONS

Evaluating cardiovascular risk is critical for guiding the selection of preventive therapies in patients with T2D. While there has been a proliferation of cardiovascular prediction models for HHF in these patients, there is a lack of external validation studies to ensure their performance and generalizability. Moreover, most models for HHF have potential concerns with risk of bias and applicability. In terms of the best available models, the TRS-HF_{DM} was identified as particularly apt for routine point-of-care use, while RECODe may be helpful in instances where holistic cardiovascular risk assessment is required. Nevertheless, the actual effect of even the best models remains unclear because of an absence of clinical impact studies. As a result, studies evaluating model-based judgments in comparison with existing clinical practice may be warranted to evaluate clinical prediction model utility in real-world practice before implementation into guideline-directed medical care.

ARTICLE INFORMATION

Received November 24, 2021; accepted March 3, 2022.

Affiliations

Centre for Outcomes Research and Evaluation, Research Institute of the McGill University Health Centre, Montreal, QC, Canada (A.R., J.M.B., A.S.); Division of Cardiology (A.R., J.M.B., N.G., A.S.); and DREAM-CV Laboratory (A.R., E.O., J.G., A.S.), McGill University Health Centre, Montreal, Quebec, Canada; Department of Cardiovascular Medicine, Kaufman Center for Heart Failure and Recovery, Heart, Vascular and Thoracic Institute, Cleveland Clinic, Cleveland, OH (V.K.R.); (J.P.F.); Department of Surgery and Physiology, Cardiovascular Research and Development Center, Faculty of Medicine of the University of Porto (J.P.F.); Division of Cardiology (S.J.G., G.M.F., M.F.); Duke Clinical Research Institute, Durham, NC (S.J.G., G.M.F., M.F.); Division of Endocrinology, Department of Medicine, McGill University, Montreal, QC, Canada (M.T., T.M.P.); Centre for Clinical Epidemiology, Lady Davis Institute for Medical Research, Montreal, QC, Canada (T.M.P.); Division of Nephrology, Department of Medicine, McGill University Health Centre and Research Institute, Montreal, Canada (T.A.M.); and Division of Cardiology, University of Alberta, Edmonton, Alberta, Canada (J.E.).

Sources of Funding

None.

Disclosures

Mr Razaghizad has received a Graduate Excellence Award in Medicine and a Ethelwyn and Ernie Nyman Fellowship from McGill University. Dr Greene has received research support from the Duke University Department of Medicine Chair's Research Award, American Heart Association, National Heart Lung and Blood Institute, Amgen, AztraZeneca, Bristol-Myers Squibb, Cytokinetics, Merck, Novartis, and Pfizer; has served on advisory boards for Amgen, AstraZeneca, and Cytokinetics; and has served as a consultant for Amgen, Bayer, Bristol-Myers Squibb, Merck, and Vifor. Dr Fudim was supported by the National Heart, Lung, and Blood Institute (K23HL151744), the American Heart Association (20IPA35310955), Mario Family Award, Duke Chair's Award, Translating Duke Health Award, Bayer, Bodyport, and BTG Specialty Pharmaceuticals. He receives consulting fees from AxonTherapies, Bodyport, Boston Scientific, CVRx, Daxor, Edwards LifeSciences, Fire1, NXT Biomedical, Viscardia, Zoll. Dr Brophy reports receiving support from the Fonds de Recherche Santé Quebec. Dr Sharma reports receiving support from the Canadian Institutes of Health Research (175095), Fonds de Recherche Santé Quebec, Alberta Innovates Health Solution, European Society of Cardiology young investigator grant, Roche Diagnostics, Boeringer-Ingelheim, Novartis, and Takeda. Dr Fudim reports consulting fees from AxonTherapies and Daxor. Dr Tsoukas reports no disclosures in relation to this work but has received speaker honoraria from NovoNordisk, Eli Lilly, Boehringer-Ingelheim, and AstraZeneca. Dr Mavrakanas has received honoraria from Daiichi Sankyo, BMS Canada, Janssen, and Pfizer and has served on advisory boards for Boehringer Ingelheim outside the submitted work. He is supported from the Fonds de Recherche Santé Quebec (FRSQ) Junior 1 clinician scholar program and from the McGill University Department of Medicine salary support program. Dr Felker has received research grants

from NHLBI, American Heart Association, Amgen, Bayer, BMS, Merck, Cytokinetics, and CSL-Behring; he has acted as a consultant to Novartis, Amgen, BMS, Cytokinetics, Medtronic, Cardionomic, Boehringer-Ingelheim, American Regent, Abbott, Astra-Zeneca, Reprieve, Myovant, Sequana, Windtree Therapuetics, and Whiteswell, and has served on clinical endpoint committees/data safety monitoring boards for Amgen, Merck, Medtronic, EBR Systems, V-Wave, LivaNova, Siemens, and Rocket Pharma. Dr Peters has received support from the Fonds de Recherche Santé Quebec. Dr Ezekowitz reports research grants or consulting work outside this work from American Regent, Amgen, AstraZeneca, Bayer, Boehringer-Ingelheim, Bristol-Myers Squibb/Pfizer, eko.ai, US2.ai, Merck, Novartis, Otsuka, Sanofi, and Servier. The remaining authors have no disclosures to report.

Supplemental Material

Data S1
Tables S1–S2
Figures S1–S4
References 56–61

REFERENCES

1. World Health Organization. The top 10 causes of death.
2. Seferović PM, Petrie MC, Filippatos GS, Anker SD, Rosano G, Bauersachs J, Paulus WJ, Komajda M, Cosentino F, de Boer RA, et al. Type 2 diabetes mellitus and heart failure: a position statement from the Heart Failure Association of the European Society of Cardiology. *Eur J Heart Fail*. 2018;20:853–872. doi: 10.1002/ejhf.1170
3. McMurray JJV, Gerstein HC, Holman RR, Pfeffer MA. Heart failure: a cardiovascular outcome in diabetes that can no longer be ignored. *Lancet Diabetes Endocrinol*. 2014;2:843–851. doi: 10.1016/S2213-8587(14)70031-2
4. Sharma A, Cooper LB, Fiuzat M, Mentz RJ, Ferreira JP, Butler J, Fitchett D, Moses AC, O'Connor C, Zannad F. Antihyperglycemic therapies to treat patients with heart failure and diabetes mellitus. *JACC Heart Failure*. 2018;6:813–822. doi: 10.1016/j.jchf.2018.05.020
5. Varela-Roman A, Shamagian LG, Caballero EB, Ramos PM, Veloso PR, Gonzalez-Juanatey JR. Influence of diabetes on the survival of patients hospitalized with heart failure: a 12-year study. *Eur J Heart Fail*. 2005;7:859–864. doi: 10.1016/j.ejheart.2005.01.017
6. Sharma A, Pagidipati NJ, Califf RM, McGuire DK, Green JB, Demets D, George JT, Gerstein HC, Hobbs T, Holman RR, et al. Impact of regulatory guidance on evaluating cardiovascular risk of new glucose-lowering therapies to treat Type 2 diabetes mellitus: lessons learned and future directions. *Circulation*. 2020;141:843–862. doi: 10.1161/CIRCULATIONAHA.119.041022
7. Kalyani RR. Glucose-lowering drugs to reduce cardiovascular risk in Type 2 diabetes. *N Engl J Med*. 2021;384:1248–1260. doi: 10.1056/NEJMcp2000280
8. Steyerberg EW. *Clinical prediction models*. Cham: Springer International Publishing; 2019.
9. Foster ED, Deardorff A. Open science framework (OSF). *J Med Libr Assoc: JMLA*. 2017;105:203.
10. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*. 2009;6:e1000097. doi: 10.1371/journal.pmed.1000097
11. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, Reitsma JB, Collins GS. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Medicine*. 2014;11:e1001744. doi: 10.1371/journal.pmed.1001744
12. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol*. 2016;75:40–46. doi: 10.1016/j.jclinepi.2016.01.021
13. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med*. 2019;170:W1. doi: 10.7326/M18-1377
14. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925–1931. doi: 10.1093/eurheartj/ehu207
15. Willis M, Asseburg C, Slee A, Nilsson A, Neslusan C. Macrovascular risk equations based on the CANVAS program. *Pharmacoeconomics*. 2021;39:447–461. doi: 10.1007/s40273-021-01001-0
16. Sharma A, Vaduganathan M, Ferreira JP, Liu Y, Bakris GL, Cannon CP, White WB, Zannad F. Clinical and biomarker predictors of expanded heart failure outcomes in patients with Type 2 diabetes mellitus after a recent acute coronary syndrome: insights from the EXAMINE trial. *J Am Heart Assoc*. 2020;9. doi: 10.1161/JAHA.119.012797
17. Berg DD, Wiviott SD, Scirica BM, Gurmu Y, Mosenzon O, Murphy SA, Bhatt DL, Leiter LA, McGuire DK, Wilding JPH, et al. Heart failure risk stratification and efficacy of sodium-glucose cotransporter-2 inhibitors in patients with Type 2 diabetes mellitus. *Circulation*. 2019;140:1569–1577. doi: 10.1161/CIRCULATIONAHA.119.042685
18. Kim E, Caraballo PJ, Castro MR, Pieczkiewicz DS, Simon GJ. Towards more accessible precision medicine: building a more transferable machine learning model to support prognostic decisions for micro- and macrovascular complications of Type 2 diabetes mellitus. *J Med Syst*. 2019;43. doi: 10.1007/s10916-019-1321-6
19. Fraty M, Velho G, Gand E, Fumeron F, Ragot S, Sosner P, Mohammedi K, Gellen B, Saulnier P-J, Halimi J-M, et al. Prognostic value of plasma MR-proADM vs NT-proBNP for heart failure in people with Type 2 diabetes: the SURDIAGENE prospective study. *Diabetologia*. 2018;61:2643–2653. doi: 10.1007/s00125-018-4727-7
20. Shao H, Fonseca V, Stoecker C, Liu S, Shi L. Novel risk engine for diabetes progression and mortality in USA: Building, Relating, Assessing, and Validating Outcomes (BRAVO). *Pharmacoeconomics*. 2018;36:1125–1134. doi: 10.1007/s40273-018-0662-1
21. Basu S, Sussman JB, Berkowitz SA, Hayward RA, Yudkin JS. Development and validation of Risk Equations for Complications Of Type 2 Diabetes (RECODe) using individual participant data from randomised trials. *Lancet Diabetes Endocrinol*. 2017;5:788–798. doi: 10.1016/S2213-8587(17)30221-8
22. Wolsk E, Claggett B, Pfeffer MA, Diaz R, Dickstein K, Gerstein HC, Lawson FC, Lewis EF, Maggioni AP, McMurray JJV, et al. Role of B-type natriuretic peptide and N-terminal prohormone BNP as predictors of cardiovascular morbidity and mortality in patients with a recent coronary event and Type 2 diabetes mellitus. *J Am Heart Assoc*. 2017;6:e004743. doi: 10.1161/JAHA.116.004743
23. Kiadaliri AA, Gerdtham UG, Nilsson P, Eliasson B, Gudbjornsdottir S, Carlsson KS. Towards renewed health economic simulation of Type 2 diabetes: risk equations for first and second cardiovascular events from Swedish register data. *PLoS One*. 2013;8:e62650. doi: 10.1371/journal.pone.0062650
24. Williams BA, Geba D, Cordova JM, Shetty SS. A risk prediction model for heart failure hospitalization in Type 2 diabetes mellitus. *Clin Cardiol*. 2020;43:275–283. doi: 10.1002/clc.23298
25. Segar MW, Vaduganathan M, Patel KV, McGuire DK, Butler J, Fonarow GC, Basit M, Kannan V, Grodin JL, Everett B, et al. Machine learning to predict the risk of incident heart failure hospitalization among patients with diabetes: the WATCH-DM risk score. *Diabetes Care*. 2019;42:2298–2306. doi: 10.2337/dc19-0587
26. Halon DA, Ayman J, Rubinshtein R, Zafrir B, Azencot M, Lewis BS. Cardiac computed tomography angiographic findings as predictors of late heart failure in an asymptomatic diabetic cohort: an 8-year prospective follow-up study. *Cardiology*. 2017;138:218–227. doi: 10.1159/000478995
27. Hippisley-Cox J, Coupland C. Development and validation of risk prediction equations to estimate future risk of heart failure in patients with diabetes: a prospective cohort study. *BMJ OPEN*. 2015;5. doi: 10.1136/bmjopen-2015-008503
28. Pfister R, Cairns R, Erdmann E, Schneider CA. A clinical risk score for heart failure in patients with Type 2 diabetes and macrovascular disease: an analysis of the PROactive study. *Int J Cardiol*. 2013;162:112–116. doi: 10.1016/j.ijcard.2011.05.056
29. Yang X, Ma RC, So W-Y, Kong AP, Ko GT, Ho C-S, Lam CW, Cockram CS, Tong PC, Chan JC. Development and validation of a risk score for hospitalization for heart failure in patients with Type 2 diabetes mellitus. *Cardiovasc Diabetol*. 2008;7:1–8. doi: 10.1186/1475-2840-7-9
30. Razaghizad A, Pedro Ferreira J, Ni J, Zannad F, Sharma A. Validation of the thrombolysis in myocardial infarction risk score for heart failure in diabetes (TRS-HFDM) in patients with recent acute coronary syndrome: an analysis of the EXAMINE trial. *McGill J Med*. 2021;19.
31. Elharram M, Ferreira JP, Huynh T, Ni J, Giannetti N, Verma S, Zannad F, Sharma A. Prediction of heart failure outcomes in patients with type 2

diabetes mellitus: validation of the thrombolysis in myocardial infarction risk score for heart failure in diabetes (TRS-HFDM) in patients in the ACCORD trial. *Diabetes Obes Metab*. 2021;23:782–790. doi: 10.1111/dom.14283
32. Basu S, Sussman JB, Berkowitz SA, Hayward RA, Bertoni AG, Correa A, Mwasongwe S, Yudkin JS. Validation of risk equations for complications of Type 2 diabetes (RECODe) using individual participant data from diverse longitudinal cohorts in the U.S. *Diabetes Care*. 2018;41:586–595. doi: 10.2337/dc17-2002
33. Segar MW, Jaeger BC, Patel KV, Nambi V, Ndumele CE, Correa A, Butler J, Chandra A, Ayers C, Rao S, et al. Development and validation of machine learning-based race-specific models to predict 10-year risk of heart failure: a multicohort analysis. *Circulation*. 2021;143:2370–2383. doi: 10.1161/CIRCULATIONAHA.120.053134
34. Prieto-Merino D, Pocock SJ. The science of risk models. *Eur J Prev Cardiol*. 2012;19:7–13. doi: 10.1177/2047487312448995
35. Verma S, Leiter LA, Zinman B, Sharma A, Mattheus M, Fitchett D, George J, Ofstad AP, Kosiborod MN, Wanner C, et al. Time to cardiovascular benefits of empagliflozin: a post hoc observation from the EMPA-REG OUTCOME trial. *ESC Heart Failure*. 2021;8:2603–2607. doi: 10.1002/ehf2.13374
36. McMurray JJV, Solomon SD, Inzucchi SE, Køber L, Kosiborod MN, Martinez FA, Ponikowski P, Sabatine MS, Anand IS, Bělohlávek J, et al. Dapagliflozin in patients with heart failure and reduced ejection fraction. *N Engl J Med*. 2019;381:1995–2008. doi: 10.1056/NEJMoa1911303
37. Anker SD, Butler J, Filippatos G, Ferreira JP, Bocchi E, Böhm M, Brunner–La Rocca H-P, Choi D-J, Chopra V, Chuquiure-Valenzuela E, et al. Empagliflozin in heart failure with a preserved ejection fraction. *N Engl J Med*. 2021;385:1451–1461. doi: 10.1056/NEJMoa2107038
38. Packer M, Anker SD, Butler J, Filippatos G, Pocock SJ, Carson P, Januzzi J, Verma S, Tsutsui H, Brueckmann M, et al. Cardiovascular and renal outcomes with Empagliflozin in heart failure. *N Engl J Med*. 2020;383:1413–1424. doi: 10.1056/NEJMoa2022190
39. Greene SJ, Khan MS. Quadruple medical therapy for heart failure. *J Am Coll Cardiol*. 2021;77:1408–1411. doi: 10.1016/j.jacc.2021.02.006
40. Sharma A, Bhatt DL, Calvo G, Brown NJ, Zannad F, Mentz RJ. Heart failure event definitions in drug trials in patients with Type 2 diabetes. *Lancet Diabetes Endocrinol*. 2016;4:294–296.
41. Alba AC, Agoritsas T, Jankowski M, Courvoisier D, Walter SD, Guyatt GH, Ross HJ. Risk prediction models for mortality in ambulatory patients with heart failure: a systematic review. *Circ Heart Fail*. 2013;6:881–889. doi: 10.1161/CIRCHEARTFAILURE.112.000043
42. Califf RM, Pencina MJ. Predictive models in heart failure. *Circ Heart Fail*. 2013;6:877–878. doi: 10.1161/CIRCHEARTFAILURE.113.000659
43. Scirica BM, Bhatt DL, Braunwald E, Raz I, Cavender MA, Im K, Mosenzon O, Udell JA, Hirshberg B, Pollack PS, et al. Prognostic implications of biomarker assessments in patients with Type 2 diabetes at high cardiovascular risk a secondary analysis of a randomized clinical trial. *JAMA Cardiol*. 2016;1:989–998. doi: 10.1001/jamacardio.2016.3030
44. Moons KG, Grobbee DE. When should we remain blind and when should our eyes remain open in diagnostic studies? *J Clin Epidemiol*. 2002;55:633–636. doi: 10.1016/S0895-4356(02)00408-0
45. Sackett DL, Haynes RB, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine*. Brown and Company: Little; 1985.
46. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;339:157–160. doi: 10.1136/bmj.b2393
47. Janssen KJM, Donders ART, Harrell FE, Vergouwe Y, Chen Q, Grobbee DE, Moons KGM. Missing covariate data in medical research: to impute is better than to ignore. *J Clin Epidemiol*. 2010;63:721–727. doi: 10.1016/j.jclinepi.2009.12.008
48. Schuster NA, Hoogendijk EO, Kok AAL, Twisk JWR, Heymans MW. Ignoring competing events in the analysis of survival data may lead to biased results: a nonmathematical illustration of competing risk analysis. *J Clin Epidemiol*. 2020;122:42–48. doi: 10.1016/j.jclinepi.2020.03.004
49. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. *Circulation*. 2016;133:601–609. doi: 10.1161/CIRCULATIONAHA.115.017719
50. Wolbers M, Koller MT, Witteman JCM, Steyerberg EW. Prognostic models with competing risks: methods and application to coronary risk prediction. *Epidemiol*. 2009;20:555–561. doi: 10.1097/EDE.0b013e3181a39056
51. Collins GS, Reitsma JB, Altman DG, Moons Karel GM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD). *Circulation*. 2015;131:211–219. doi: 10.1161/CIRCULATIONAHA.114.014508
52. Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, Woodward M. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart*. 2012;98:691–698. doi: 10.1136/heartjnl-2011-301247
53. Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. *JAMA*. 2015;313:409–410. doi: 10.1001/jama.2015.37
54. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26:565–574. doi: 10.1177/0272989X06295361
55. Ohkuma T, Komorita Y, Peters SAE, Woodward M. Diabetes as a risk factor for heart failure in women and men: a systematic review and meta-analysis of 47 cohorts including 12 million individuals. *Diabetologia*. 2019;62:1550–1560. doi: 10.1007/s00125-019-4926-x
56. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. *BMJ*. 2009;338. doi: 10.1136/bmj.b605
57. Nam B-H, D'Agostino RB. Discrimination Index, the Area Under the ROC Curve. In: Huber-Carol C, Balakrishnan N, Nikulin MS, Mesbah M eds. *Goodness-of-fit tests and model validity*. Statistics for Industry and Technology. Boston, MA: Birkhäuser; 2002:267–279.
58. Royston P. Tools for checking calibration of a Cox model in external validation: approach based on individual event probabilities. *STATA J*. 2014;14:738–755. doi: 10.1177/1536867X1401400403
59. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol*. 2013;13:33. doi: 10.1186/1471-2288-13-33
60. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness of fit in the survival setting. *Stat Med*. 2015;34:1659–1680. doi: 10.1002/sim.6428
61. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology*. 2010;21:128–138. doi: 10.1097/EDE.0b013e3181c30fb2

SUPPLEMENTAL MATERIAL

Data S1. Detailed Description of Internal- and External Validation, and Measures Used to Evaluate These Metrics.

A clinical prediction models' performance may be evaluated in internal- or external validation. Internal validation reflects a models' reproducibility, and it includes apparent validation, where the model is validated in the derivation cohort; split-sample validation, where the data is randomly split into a training and validation set; bootstrapping, where multiple training and validation datasets are created by random draw; and cross-validation, where training is done in a random segment of the cohort and tested in the remaining part.[8] External validation conversely reflects a models' generalizability, and it includes geographic validation, where validation is done in another country or center; independent validation, where validation is done by other researchers; and temporal validation, where validation is done using data from a different period.[8,56]

Model performance is normally evaluated through discrimination and calibration. Discrimination reflects a model's ability to distinguish between patients who do and do not experience an outcome of interest.[14] Discrimination is frequently assessed with measures of concordance (e.g., c-statistic, AUC) and it can range between 0.5 for a model no better than the play of chance to 1.0 for a perfect model.[57] Concordance estimates the probability that a randomly selected patient who experienced an outcome had a higher predicted risk than a patient who did not. Calibration reflects a model's predictive accuracy: i.e., the agreement between the predicted probability of events and the actual proportion of events observed.[14] Calibration is frequently assessed with statistical tests for goodness-of-fit (e.g., Hosmer-Lemeshow, $p<0.05$ signifies poor calibration), or graphical plots for visual assessment (e.g., calibration plot slope <0.7 signifies poor calibration).[58–60] Less common performance measures are described elsewhere and include R^2, Brier score, sensitivity, specificity, accuracy, and net reclassification.[61]

Table S1. Systematic Search Algorithms.

Database: Ovid MEDLINE(R) ALL <1946 to February 24, 2021> 1,054 records

1. *diabetes mellitus, type 2/
2. (diabet* and ("type 2" or "type ii" or non-insulin* or noninsulin*)).mp.
3. (T2DM or DMT2 or TIIDM or DMTII or NIDDM).mp.
4. exp *heart failure/
5. ((heart or cardiac or myocardial) adj2 failure*).mp.
6. ((prognos* or predict* or risk* or strati*) and (model* or tool* or scor* or index or nomogram* or formula* or staging or calculat* or equation* or strati* or chart* or function* or engine* or algorithm*)).ti,ab,kw.
7. *risk assessment/ or exp *risk factors/ or *multivariate analysis/ or exp regression analysis/ or exp survival analysis/ or disease-free survival/ or kaplan-meier estimate/ or progression-free survival/ or proportional hazards models/ or logistic models/ or nomograms/ or area under curve/ or exp models, statistical/
8. ("disease free survival" or "proportional hazard* model*" or (survival adj2 anal*) or "kaplan-meier estimate*" or "progression-free survival" or develop* or (cox adj3 (model* or anal*)) or (random adj2 forest*) or regress* or (logistic* adj2 model*) or multivari* or (likelihood adj2 function) or (area under adj2 curve) or (statistical adj3 model*) or discrimin* or calibrat* or valid* or "integer-based" or "support vector*" or (machine adj2 learning*) or mathematic* or concordance* or c-statistic* or c-ind* or hosmer-lemeshow* or hazard* or wald* or "survival rate*" or "survival time*" or "survival funct*").mp.

(1 or 2 or 3) and (4 or 5) and 6 and (7 or 8)

Database: Embase Classic+Embase <1947 to February 24, 2021> 3,473 records

1. exp *non insulin dependent diabetes mellitus/
2. (diabet* and ("type 2" or "type ii" or non-insulin* or noninsulin*)).mp.
3. (T2DM or DMT2 or TIIDM or DMTII or NIDDM).mp.
4. exp *heart failure/
5. ((heart or cardiac or myocardial) adj2 failure*).mp.
6. (prognosis/) and (model/)
7. ((prognos* or predict* or risk* or strati*) and (model* or tool* or scor* or index or nomogram* or formula* or staging or calculat* or equation* or strati* or chart* or function* or engine* or algorithm*)).ti,ab,kw.
8. exp risk assessment/ or exp risk factor/ or exp multivariate analysis/ or exp regression analysis/ or *disease free survival/ or exp proportional hazards model/ or *statistical model/ or exp nomograms/ or *area under the curve/ or exp mathematical phenomena/
9. ("disease free survival" or "proportional hazard* model*" or (survival adj2 anal*) or "kaplan-meier estimate*" or "progression-free survival" or develop* or (cox adj3 (model* or anal*)) or (random adj2 forest*) or regress* or (logistic* adj2 model*) or multivari* or (likelihood adj2 function) or (area under adj2 curve) or (statistical adj3 model*) or discrimin* or calibrat* or valid* or "integer-based" or "support vector*" or (machine adj2 learning*) or mathematic* or concordance* or c-statistic* or c-ind* or hosmer-lemeshow* or hazard* or wald* or "survival rate*" or "survival time*" or "survival funct*").mp.

(1 or 2 or 3) and (4 or 5) and (6 or 7) and (8 or 9)

Database: Web of Science Core Collection <database inception to February 24, 2021> 1,426 records

1. ts= (diabet* and ("type 2" or "type ii" or noninsulin or “non-insulin”))
2. ts=((heart or cardiac or myocardial) near/2 failure*)
3. ts=((prognos* or predict* or risk* or strati*) and (model* or tool* or scor* or index or nomogram* or formula* or staging or calculat* or equation* or strati* or chart* or function* or engine* or algorithm*))
4. ts=("disease free survival" or "proportional hazard* model*" or (survival near/2 anal*) or "kaplan-meier estimate*" or "progression-free survival" or develop* or (cox near/3 (model* or anal*)) or (random near/2 forest*) or regress* or (logistic* near/2 model*) or multivari* or (likelihood near/2 function) or (area under near/2 curve) or (statistical near/3 model*) or discrimin* or calibrat* or valid* or "integer-based" or "support vector*" or (machine near/2 learning*) or mathematic* or concordance* or c-statistic* or c-ind* or hosmer-lemeshow* or hazard* or wald* or "survival rate*" or "survival time*" or "survival funct*")

#1 AND #2 AND #3 AND #4

Database: Google Scholar <database inception to February 24, 2021> first 200 records
1. "type 2" 2. "diabetes" 3. "heart failure" 4. ("risk" or "prediction" or "stratification" or “model”) 1 and 2 and 3 and 4
Database: Tufts Predictive Analytics and Comparative Effectiveness Clinical Prediction Model Registry <database inception to February 24, 2021> 37 records
1. Keyword contains: diabetes 2. Outcome contains: heart failure 3. Outcome contains: composite 4. Outcome contains: hospitalization 1 and (2 or 3 or 4)

Table S2. List of Excluded Studies Cataloged by Reason for Exclusion.

EXCLUDED: NOT A MODEL DEVLEOPMENT, UPDATE, OR VALIDATION STUDY
1. Akter, S., et al. “Predictors of Incident Heart Failure in Community-Dwelling Older Adults with Diabetes Mellitus.” *Diabetologia*, vol. 53, Sept. , p. S85.
2. Altrabsheh, E., et al. “Pdb64 Association between Cardiovascular Disease Risk Factors and Hypoglycaemic Events in Type 2 Diabetes Using the Iqvia Core Diabetes Model.” *Value in Health*, vol. 22, May , p. S151.
3. Ang, Donald SC, et al. “A Comparison between B-Type Natriuretic Peptide, Global Registry of Acute Coronary Events (GRACE) Score and Their Combination in ACS Risk Stratification.” *Heart*, vol. 95, no. 22, 2009, pp. 1836–42.
4. Blin, P., et al. “Real World Risk of Major Outcomes for Type 2 Diabetes with Stable Coronary Artery Disease without Prior MI or Stroke and THEMIS-like Patients Using the SNDS French Nationwide Claims Database.” *European Heart Journal*, vol. 41, Nov. , p. 1314.
5. Breunig, I. M., et al. “Development of Heart Failure in Medicaid Patients with Type 2 Diabetes Treated with Pioglitazone, Rosiglitazone, or Metformin.” *Journal of Managed Care & Specialty Pharmacy*, vol. 20, no. 9, Sept. , pp. 895–903.
6. Bucher, S., et al. “Predictive factors of hospitalization in non institutionalized elderly diabetic patients. Data from the S.AGES cohort.” *Exercer-La Revue Francophone De Medecine Generale*, no. 146, Oct. , pp. 340–47.
7. Davis, W. A., et al. “Contemporary Cardiovascular Risk Assessment for Type 2 Diabetes Including Heart Failure as an Outcome: The Fremantle Diabetes Study Phase II.” *Journal of Clinical Medicine*, vol. 9, no. 5, May .
8. Fadini, G. P., et al. “Risk of Hospitalization for Heart Failure in Patients with Type 2 Diabetes Newly Treated with DPP-4 Inhibitors or Other Oral Glucose-Lowering Medications: A Retrospective Registry Study on 127,555 Patients from the Nationwide OsMed Health-DB Database.” *European Heart Journal*, vol. 36, no. 36, Sept. , pp. 2454–62.
9. Foos, V., et al. “Validation and Evaluation of the Risk-to-Benefit Ratio of Glucose Lowering Therapies in High Cardiovascular Risk Type 2 Diabetes Patients; Projections Using the IMS CORE Diabetes Model.” *Diabetes*, vol. 61, June , p. A36.
10. Garcia Carretero, R., et al. “Cardiovascular Prognostic Factors in Prediabetic Patients within a Hypertensive Population.” *Journal of Hypertension*, vol. 36, June , p. e25.
11. Gokhale, M., et al. “Calendar Time as an Instrumental Variable in Assessing the Risk of Heart Failure with Antihyperglycemic Drugs.” *Pharmacoepidemiology and Drug Safety*, vol. 25, Aug. , pp. 45–46.
12. Halon, D. A., et al. “Prediction of Heart Failure in Asymptomatic Type 2 Diabetics: An 8 Year Prospective Study Following Cardiac CT Angiography.” *European Journal of Heart Failure*, vol. 19, May , p. 178.
13. Hayes, A. J., et al. “An Improved Model to Estimate Lifetime Health Outcomes of Patients with Type 2 Diabetes Using 30-Year Follow-up Data from the United Kingdom Prospective Diabetes Study.” *Diabetologia*, vol. 54, Sept. , p. S8.
14. Jhund, P., et al. “NT-ProBNP and HsTnT Improve Cardiovascular Risk Prediction in Patients with Type 2 Diabetes Mellitus, Chronic Kidney or Cardiovascular Disease or Both.” *Journal of the American College of Cardiology*, vol. 63, no. 12, 1, p. A1279.
15. Kempf, Tibor, et al. “Prognostic Utility of Growth Differentiation Factor-15 in Patients with Chronic Heart Failure.” *Journal of the American College of Cardiology*, vol. 50, no. 11, 2007, pp. 1054–60.
16. Leal, J., et al. “Temporal Validation of the UKPDS Outcomes Model Using 10-Year Posttrial Monitoring Data.” *Diabetes Care*, vol. 36, no. 6, June , pp. 1541–46.
17. Maxion-Bergemann, S., et al. “Diabetes Mellitus Model (DMM): Internal Validation of a Computer Simulation Model for Type 1 and Type 2 Diabetes.” *Journal of Medical Economics*, vol. 9, no. 69, 2006, pp. 69–82.
18. McAlister, F. A., et al. “Association between Glycated Haemoglobin Levels and Cardiovascular Outcomes in Patients with Type 2 Diabetes and Cardiovascular Disease: A Secondary Analysis of TheTECOSrandomized Clinical Trial.” *European Journal of Heart Failure*, vol. 22, no. 11, Nov. , pp. 2026–34.
19. McEwan, P., V. Foos, et al. “Approaches to Standardising Cardiovascular Risk Equation End-Points in Order to Facilitate Their Inclusion within a Type 2 Diabetes Model.” *Value in Health*, vol. 20, May , p. A323.
20. McEwan, P., H. Bennett, et al. “Refitting of the UKPDS 68 Risk Equations to Contemporary Routine Clinical Practice Data in the UK.” *Pharmacoeconomics*, vol. 33, no. 2, Feb. , pp. 149–61.

21. McEwan, P. “Validation of the Ukpds Outcomes Equations Using the Cardiff Type-2 Diabetes Model.” *Value in Health*, vol. 16, May , p. A165.
22. Monteiro, S., et al. “Pdb35 Contrasting Framingham Risk Equations in Type 2 Diabetes Using the Iqvia Core Diabetes Model.” *Value in Health*, vol. 22, May , p. S145.
23. Neuwahl, S. J., et al. “Patient Health Utility Equations for a Type 2 Diabetes Model.” *Diabetes Care*, vol. 44, no. 2, Feb. , pp. 381–89.
24. Nichols, G. A., et al. “Congestive Heart Failure in Type 2 Diabetes - Prevalence, Incidence, and Risk Factors.” *Diabetes Care*, vol. 24, no. 9, Sept. , pp. 1614–19.
25. Ono, Y., et al. “Validity of Claims Diagnosis Codes for Cardiovascular Diseases in Diabetes Patients in Japanese Administrative Database.” *Clinical Epidemiology*, vol. 12, 2020, pp. 367–75.
26. Pagano, E., A. Gray, et al. “Prediction of Mortality and Macrovascular Complications in Type 2 Diabetes: Validation of the UKPDS Outcomes Model in the Casale Monferrato Survey, Italy.” *Diabetologia*, vol. 56, no. 8, Aug. , pp. 1726–34.
27. Pagano, E., S. R. A. Konings, et al. “Prediction of Mortality and Major Cardiovascular Complications in Type 2 Diabetes: External Validation of UK Prospective Diabetes Study Outcomes Model Version 2 in Two European Observational Cohorts.” *Diabetes Obesity & Metabolism*.
28. Presley, C. A., et al. “Validation of an Algorithm to Identify Heart Failure Hospitalisations in Patients with Diabetes within the Veterans Health Administration.” *BMJ Open*, vol. 8, no. 3, 2018, http://bmjopen.bmj.com/content/early/by/section. rayyan-653723784.
29. Rajagopalan, R., et al. “Association between Congestive Heart Failure and Hospitalization in Patients with Type 2 Diabetes Mellitus Receiving Treatment with Insulin or Pioglitazone: A Retrospective Data Analysis.” *Clinical Therapeutics*, vol. 26, no. 9, Sept. , pp. 1400–10.
30. Reynoso-Noveron, N., et al. “Estimated Incidence of Cardiovascular Complications Related to Type 2 Diabetes in Mexico Using the UKPDS Outcome Model and a Population-Based Survey.” *Cardiovascular Diabetology*, vol. 10, Jan. .
31. Rorth, R., et al. “Risk of Incident Heart Failure in Patients With Diabetes and Asymptomatic Left Ventricular Systolic Dysfunction.” *Diabetes Care*, vol. 41, no. 6, 6, pp. 1285–91.
32. Schievink, B., et al. “Prediction of the Effect of Atrasentan on Renal and Heart Failure Outcomes Based on Short-Term Changes in Multiple Risk Markers.” *European Journal of Preventive Cardiology*, vol. 23, no. 7, May , pp. 758–68.
33. Shao, H., et al. “Updating Risk Engine for Diabetes Progression and Mortality in the United States: Internal Validation.” *Value in Health*, vol. 20, May , p. A162.
34. Su, Z. T., et al. “The Use of Computer Simulation Modelling to Estimate Complications in Patients with Type 2 Diabetes: Validation of the Cornerstone Diabetes Simulation Model.” *Diabetologia*, vol. 61, Oct. , p. S138.
35. Tomlin, A. M., et al. “Risk Factors for Hospitalization Due to Diabetes Complications.” *Diabetes Research and Clinical Practice*, vol. 80, no. 2, May , pp. 244–52.
36. Yang, H., et al. “What Is the Best Model to Predict Incident Heart Failure?” *Journal of the American College of Cardiology*, vol. 67, no. 13, 5, p. 1337.
37. Ye, W., et al. “The Michigan Model for Coronary Heart Disease in Type 2 Diabetes: Development and Validation.” *Diabetes Technology & Therapeutics*, vol. 17, no. 10, Oct. , pp. 701–11.
38. Zhuo, X., et al. “External Validation of the Uk Prospective Diabetes Study Risk Equations in 14,740 Israel Type 2 Diabetes Patients.” *Diabetes*, vol. 65, 2016, p. A92.

EXCLUDED: INCORRECT OUTCOME

1. Aminian, A., et al. “Predicting 10-Year Risk of End-Organ Complications of Type 2 Diabetes With and Without Metabolic Surgery: A Machine Learning Approach.” *Diabetes Care*, vol. 43, no. 4, Apr. , pp. 852–59.
2. Basu, S., et al. “Benefit and Harm of Intensive Blood Pressure Treatment: Derivation and Validation of Risk Models Using Data from the SPRINT and ACCORD Trials.” *Plos Medicine*, vol. 14, no. 10, Oct. .
3. Bergemann, R., et al. “Prediction of Acute and Chronic Complications by a New Computer Simulation Model for Type 1 and Type 2 Diabetes: The Diabetes Mellitus Model (DMM).” *Journal of Medical Economics*, vol. 9, no. 83, 2006, pp. 83–99.
4. Bhattacharyya, S., et al. “Carotid Intima-Media Thickness, Echocardiography and Incident Cardiovascular Disease: Comparing the Predictive Performance of Ultrasound Imaging in Patients with Type 2 Diabetes (T2DM).” *Heart Lung and Circulation*, vol. 2, 2010, p. S169.
5. Clarke, P. M., et al. “Using the EQ-5D Index Score as a Predictor of Outcomes in Patients With Type 2 Diabetes.” *Medical Care*, vol. 47, no. 1, Jan. , pp. 61–68.

6. Cosson, E., et al. “Cardiovascular Risk Prediction Is Improved by Adding Asymptomatic Coronary Status to Routine Risk Assessment in Type 2 Diabetic Patients.” *Diabetes*, vol. 60, July , pp. A141–42.
7. Cox, A. J., et al. “Prediction of Mortality Using a Multi-Bed Vascular Calcification Score in the Diabetes Heart Study.” *Cardiovascular Diabetology*, vol. 13, no. 1, 2014, http://www.cardiab.com/home/.
8. De Magalhaes, L. S., et al. “Ukpds: Stratification of Macrovascular Risk of People with Type 2 Diabetes Mellitus at Nucleo de Atencao Em Diabetes of Blumenau-Santa Catarina.” *Diabetology and Metabolic Syndrome. Conference: 21st Brazilian Diabetes Society Congress. Brazil.*, vol. 10, 2018.
9. Enaa, J., et al. “Derivation and validation of a predictive model for the readmission of patients with diabetes mellitus treated in internal medicine departments.” *Revista Clinica Espanola*, vol. 218, no. 6, Aug. , pp. 271–78.
10. Funabashi, N., et al. “Risk Stratification of Type 2 Diabetes Mellitus Patients with Coronary Calcification for Occurrence of Major Adverse Cardiac and Cerebrovascular Events Using Coronary Agatston Score on 320 Slice CT and Quantitative Average HBA 1C Levels during Follow-Up.” *Circulation. Conference*, vol. 138, 2018.
11. Geba, D., et al. “A Risk Prediction Model for Cardiovascular Disease-Related Death in Patients with Type 2 Diabetes Mellitus.” *Circulation. Conference*, vol. 138, 2018.
12. Hayden, J. D., et al. “Cardiovascular Risk Scoring and Stratification in Patients with Type 2 Diabetes Enrolled in a Medicare Advantage Plan.” *Diabetes. Conference: 80th Scientific Sessions of the American Diabetes Association, ADA*, vol. 69, 2020, https://diabetes.diabetesjournals.org/content/69/Supplement_1/404-P.
13. Hillis, G. S., et al. “The Relative and Combined Ability of High-Sensitivity Cardiac Troponin T and N-Terminal Pro-B-Type Natriuretic Peptide to Predict Cardiovascular Events and Death in Patients With Type 2 Diabetes.” *Diabetes Care*, vol. 37, no. 1, Jan. , pp. 295–303.
14. Hoerger, T. J., et al. “Developing New Risk Equations to Predict Diabetes-Related Complications and Mortality in u.s. Adults with Type 2 Diabetes.” *Diabetes. Conference: 80th Scientific Sessions of the American Diabetes Association, ADA*, vol. 69, 2020, https://diabetes.diabetesjournals.org/content/69/Supplement_1/1520-P.
15. J, Cederholm, et al. “Risk Prediction of Cardiovascular Disease in Type 2 Diabetes: A Risk Equation from the Swedish National Diabetes Register.” *Diabetes Care*, vol. 31, no. 10, 2008, pp. 2038–43.
16. Lau, E., et al. “A Modified Risk Equation for Development of Coronary Heart Disease in Hong Kong Chinese with Type 2 Diabetes.” *Diabetes Research and Clinical Practice*, vol. 106, Nov. , p. S17.
17. Lin, M. Y., et al. “Pdb57 Validation of Ukpds Outcomes and Taiwan Diabetes Models on Taiwan Type 2 Dm Population.” *Value in Health*, vol. 23, May , pp. S118–19.
18. M, Woodward, et al. “Adding Social Deprivation and Family History to Cardiovascular Risk Assessment: The ASSIGN Score from the Scottish Heart Health Extended Cohort (SHHEC).” *Heart (British Cardiac Society)*, vol. 93, no. 2, 2007, pp. 172–76.
19. McMurray, J. J. V., et al. “Predictors of Fatal and Nonfatal Cardiovascular Events in Patients with Type 2 Diabetes Mellitus, Chronic Kidney Disease, and Anemia: An Analysis of the Trial to Reduce Cardiovascular Events with Aranesp (Darbepoetin-Alfa) Therapy (TREAT).” *American Heart Journal*, vol. 162, no. 4, Oct. , pp. 748-U208.
20. Mudrikova, T., et al. “Cardiovascular Risk Factors as Predictors of Mortality in Type II Diabetic Patients.” *Wiener Klinische Wochenschrift*, vol. 111, no. 2, Jan. , pp. 66–69.
21. Nakamura, M., et al. “Brachial-Ankle Pulse Wave Velocity as a Risk Stratification Index for the Short-Term Prognosis of Type 2 Diabetic Patients with Coronary Artery Disease.” *Hypertension Research*, vol. 33, no. 10, Oct. , pp. 1018–24.
22. Nishimura, Tsunehiko, et al. “Prognostic Study of Risk Stratification among Japanese Patients with Ischemic Heart Disease Using Gated Myocardial Perfusion SPECT: J-ACCESS Study.” *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 35, no. 2, 2008, pp. 319–28.
23. Parrinello, C. M., et al. “Risk Prediction of Major Complications in Individuals with Diabetes: The Atherosclerosis Risk in Communities Study.” *Diabetes Obesity & Metabolism*, vol. 18, no. 9, Sept. , pp. 899–906.
24. Porrini, E., et al. “10-Year Cardiovascular Risk in Type 2 Diabetes Is Predicted by Measurable Urinary Albumin Even in the Normoalbuminuric Range.” *Nephrology Dialysis Transplantation*, vol. 27, May , pp. ii55–56.
25. Prausmüller, Suriya, Michael Resl, Henrike Arfsten, Georg Spinka, Raphael Wurm, Stephanie Neuhold, Philipp Bartko, et al. *Performance of NT-ProBNP as a Single Biomarker in Comparison to SCORE and the*

Recommended ESC/EASD Cardiovascular Risk Stratification Model for Risk Prediction in Type 2 Diabetes Mellitus. 2020.

26. Prausmüller, Suriya, Michael Resl, Henrike Arfsten, Georg Spinka, Raphael Wurm, Stephanie Neuhold, Philipp E. Bartko, et al. "Performance of the Recommended ESC/EASD Cardiovascular Risk Stratification Model in Comparison to SCORE and NT-ProBNP as a Single Biomarker for Risk Prediction in Type 2 Diabetes Mellitus." *Cardiovascular Diabetology*, vol. 20, no. 1, 2021, pp. 1–12.
27. Resl, M., et al. "Targeted Multiple Biomarker Approach in Predicting Cardiovascular Events in Patients with Diabetes." *Heart*, vol. 102, no. 24, Dec. , pp. 1963–68.
28. Tang, Olive, et al. "Performance of High-Sensitivity Cardiac Troponin Assays to Reflect Comorbidity Burden and Improve Mortality Risk Stratification in Older Adults with Diabetes." *Diabetes Care*, vol. 43, no. 6, 2020, pp. 1200–08.
29. "The Detection of Ischemia in Asymptomatic Diabetics (DIAD) Study Risk Score." *Diabetes*, vol. 61, June , p. A103.
30. Torremocha, F., et al. "Prediction of Major Coronary Events by Coronary Risk Profile and Silent Myocardial Ischaemia: Prospective Follow-up Study of Primary Prevention in 72 Diabetic Patients." *Diabetes & Metabolism*, vol. 27, no. 1, Feb. , pp. 49–57.
31. Van Der Leeuw, J., et al. "The Validation of Cardiovascular Risk Scores for Patients with Type Diabetes Mellitus." *Heart*, vol. 101, no. 3, 1, pp. 222–29.
32. Wang, Y., et al. "Comparison of the Heart Failure Risk Stratification Performance of the CKD–EPI Equation and the MDRD Equation for Estimated Glomerular Filtration Rate in Patients with Type 2 Diabetes." *Diabetic Medicine*, vol. 33, no. 5, 2016, pp. 609–20.
33. Williams, B., et al. "A Prediction Model for Identifying Type 2 Diabetics at Highest Risk of Cardiorenal Outcomes." *Journal of Cardiac Failure*, vol. 25, Aug. , p. S98.
34. Young, J. B., et al. "Development of Predictive Risk Models for Major Adverse Cardiovascular Events among Patients with Type 2 Diabetes Mellitus Using Health Insurance Claims Data." *Cardiovascular Diabetology*, vol. 17, Aug. .

EXCLUDED: INCORRECT PATIENT POPULATION

1. Arzilli, Chiara, et al. "N-Terminal Fraction of pro-B-Type Natriuretic Peptide versus Clinical Risk Scores for Prognostic Stratification in Chronic Systolic Heart Failure." *European Journal of Preventive Cardiology*, vol. 25, no. 8, 2018, pp. 889–95.
2. Atlantis, E., et al. "Predictive Value of Serum Testosterone for Type 2 Diabetes Risk Assessment in Men." *Bmc Endocrine Disorders*, vol. 16, May .
3. Barthel, Petra, et al. "Reflex and Tonic Autonomic Markers for Risk Stratification in Patients with Type 2 Diabetes Surviving Acute Myocardial Infarction." *Diabetes Care*, vol. 34, no. 8, 2011, pp. 1833–37.
4. Bavishi, Aakash, et al. "Systematic Examination of a Heart Failure Risk Prediction Tool: The Pooled Cohort Equations to Prevent Heart Failure." *PloS One*, vol. 15, no. 11, 2020, p. e0240567.
5. Berezin, Alex, et al. "The Utility of Biomarker Risk Prediction Score in Patients with Chronic Heart Failure." *Clinical Hypertension*, vol. 22, no. 1, 2015, pp. 1–11.
6. Brouwers, Frank P., et al. "Clinical Risk Stratification Optimizes Value of Biomarkers to Predict New-Onset Heart Failure in a Community-Based Cohort." *Circulation: Heart Failure*, vol. 7, no. 5, 2014, pp. 723–31.
7. Gary, T., et al. "CHA2DS2-VASc Score and Risk for Reobstruction after Endovascular Treatment of the Superficial Femoral Artery: Differences between Balloon Angioplasty and Stenting." *Journal of Thrombosis and Haemostasis*, vol. 11, July , p. 387.
8. Iacovoni, A., et al. "Natriuretic Peptides and the Framingham Risk Score for Screening of Asymptomatic Left Ventricular Systolic Dysfunction in High-Risk Patients in Primary Care. The DAVID-BERG Study." *International Journal of Cardiology*, vol. 168, no. 5, 12, pp. 5093–95.
9. JV, Wylie, et al. "Validated Risk Score Predicts the Development of Congestive Heart Failure after Presentation with Unstable Angina or Non-ST-Elevation Myocardial Infarction: Results from OPUS-TIMI 16 and TACTICS-TIMI 18." *American Heart Journal*, vol. 148, no. 1, 2004, pp. 173–80.
10. Khan, Sadiya S., et al. "10-Year Risk Equations for Incident Heart Failure in the General Population." *Journal of the American College of Cardiology*, vol. 73, no. 19, 2019, pp. 2388–97.
11. L, Wang, et al. "Predicting Risk of Hospitalization or Death among Patients with Heart Failure in the Veterans Health Administration." *The American Journal of Cardiology*, vol. 110, no. 9, 2012, pp. 1342–49.
12. Montero-Perez-Barquero, Manuel, et al. "Utility of the SENIORS Elderly Heart Failure Risk Model Applied to the RICA Registry of Acute Heart Failure." *International Journal of Cardiology*, vol. 182, 2015, pp. 449–53.

13. Naccarelli, G. V., M. P. Panaccio, et al. “CHADS2 and CHA2DS2-VASc Risk Factors to Predict First Cardiovascular Hospitalization among Atrial Fibrillation/Atrial Flutter Patients.” *American Journal of Cardiology*, vol. 109, no. 10, May , pp. 1526–33.
14. Naccarelli, G. V., G. Cummins, et al. “CHADS2 Risk Factors Predict Cardiovascular Hospitalization among Atrial Fibrillation/Atrial Flutter Patients.” *Heart Rhythm*, vol. 8, no. 5, May , p. S337.
15. Nozaki, Toshimitsu, et al. “Significance of a Multiple Biomarkers Strategy Including Endothelial Dysfunction to Improve Risk Stratification for Cardiovascular Events in Patients at High Risk for Coronary Heart Disease.” *Journal of the American College of Cardiology*, vol. 54, no. 7, 2009, pp. 601–08.
16. P, et al. “Biomarker-Based Risk Prediction of Incident Heart Failure in Pre-Diabetes and Diabetes.” *JACC: Heart Failure*, 2021.
17. “Plasma Concentration of Amino-Terminal pro-Brain Natriuretic Peptide in Chronic Heart Failure: Prediction of Cardiovascular Events and Interaction with the Effects of Rosuvastatin: A Report from CORONA (Controlled Rosuvastatin Multinational Trial in Heart Failure).” *Journal of the American College of Cardiology*, vol. 54, no. 20, 2009, pp. 1850–59.
18. Protasov, V. N., et al. “Multimarker Approach in Risk Stratification of Patients with Decompensated Heart Failure.” *Kardiologiia*, vol. 59, no. 1, 2019, pp. 53–64.
19. Sabatine, Marc S., et al. “Multimarker Approach to Risk Stratification in Non-ST Elevation Acute Coronary Syndromes: Simultaneous Assessment of Troponin I, C-Reactive Protein, and B-Type Natriuretic Peptide.” *Circulation*, vol. 105, no. 15, 2002, pp. 1760–63.
20. Schildcrout, J. S., et al. “A Prognostic Model Based on Readily Available Clinical Data Enriched a Pre-Emptive Pharmacogenetic Testing Program.” *Journal of Clinical Epidemiology*, vol. 72, 1, pp. 107–15.
21. Scirica, B. M., et al. “A Diabetes Risk Score for Cardiovascular Events in Primary and Secondary Prevention-Observations from the SAVOR-TIMI 53 Study.” *Diabetes*, vol. 64, June , p. A113.
22. Spinar, Jindrich, et al. “AHEAD Score—Long-Term Risk Classification in Acute Heart Failure.” *International Journal of Cardiology*, vol. 202, 2016, pp. 21–26.
23. Tsang, Teresa SM, et al. “Prediction of Risk for First Age-Related Cardiovascular Events in an Elderly Population: The Incremental Value of Echocardiography.” *Journal of the American College of Cardiology*, vol. 42, no. 7, 2003, pp. 1199–205.
24. Tse, G., et al. “The CHADS2 and CHA2DS2-VASc Scores for Predicting Healthcare Utilization and Outcomes: Observations on the Appropriate Use and Misuse of Risk Scores.” *International Journal of Cardiology*, vol. 245, 15, pp. 181–82.
25. Tse, Gary, et al. “Multi-Modality Machine Learning Approach for Risk Stratification in Heart Failure with Left Ventricular Ejection Fraction≤ 45%.” *ESC Heart Failure*, vol. 7, no. 6, 2020, pp. 3716–25.
26. Wamil, M., et al. “Predicting Heart Failure Events in Patients with Coronary Heart Disease and Impaired Glucose Tolerance: Insights from the Acarbose Cardiovascular Evaluation (ACE) Trial.” *Diabetes Research and Clinical Practice*, vol. 170, Dec. .
27. Y, Wu, et al. “Estimation of 10-Year Risk of Fatal and Nonfatal Ischemic Cardiovascular Diseases in Chinese Adults.” *Circulation*, vol. 114, no. 21, 2006, pp. 2217–25.

EXCLUDED: NOT ORIGINAL DATA

1. Al-Lawati, J. A., et al. “Cardiovascular Risk Assessment in Diabetes Mellitus: Comparison of the General Framingham Risk Profile versus the World Health Organization/ International Society of Hypertension Risk Prediction Charts in Arabs - Clinical Implications.” *Angiology*, vol. 64, no. 5, July , pp. 336–42.
2. Almeda-Valdes, P., et al. “UKPDS Risk Engine, DECODE and Diabetes PHD Models for the Estimation of Cardiovascular Risk in Patients with Diabetes.” *Current Diabetes Reviews*, vol. 6, no. 1, Jan. , pp. 1–8.
3. Angelidi, A., et al. “Association and Predictability of the Novel ASCVD Risk Score in the Establishment of Heart Failure in Patients with Type 2 Diabetes: Data from a 10-Year Prospective Study.” *European Journal of Heart Failure*, vol. 16, May , p. 345.
4. Cannistraci, R., et al. “Risk Stratification Tools for Heart Failure in the Diabetes Clinic.” *Nutrition Metabolism and Cardiovascular Diseases*, vol. 30, no. 7, June , pp. 1070–79.
5. Coleman, R. L., et al. “Estimating Cardiovascular Risk in People with Type 2 Diabetes and Cardiovascular Disease: UKPDS Outcomes Model.” *Diabetes. Conference: 79th Scientific Sessions of the American Diabetes Association, ADA*, vol. 68, 2019, https://diabetes.diabetesjournals.org/content/68/Supplement_1/1466-P.
6. Collins, Sean P., and Alan B. Storrow. “Moving toward Comprehensive Acute Heart Failure Risk Assessment in the Emergency Department: The Importance of Self-Care and Shared Decision Making.” *JACC: Heart Failure*, vol. 1, no. 4, 2013, pp. 273–80.

7. Davis, T., and W. A. Davis. “Predictors and Outcome of Heart Failure Complicating Type 2 Diabetes: The Fremantle Diabetes Study.” *Diabetes*, vol. 64, June , p. A387.
8. Davis, W. A., et al. “A Contemporary Australian Cardiovascular Risk Equation for Type 2 Diabetes: The Fremantle Diabetes Study Phase II.” *Diabetologia*, vol. 61, Oct. , pp. S575–76.
9. Fadini, G. P., et al. “Intraclass Differences in the Risk of Hospitalization for Heart Failure among Patients with Type 2 Diabetes Initiating a Dipeptidyl Peptidase-4 Inhibitor or a Sulphonylurea: Results from the OsMed Health-DB Registry.” *Diabetes Obesity & Metabolism*, vol. 19, no. 10, Oct. , pp. 1416–24.
10. Jarolim, P., et al. “Prognostic Implications of Simultaneous Biomarker Assessments in Patients with Type 2 Diabetes Mellitus - Observations from the SAVOR-TIMI 53 Trial.” *Clinical Chemistry*, vol. 1, 2014, p. S232.
11. Lau, Y. C., and G. Y. Lip. “Post Myocardial Infarction and Atrial Fibrillation: Thromboprophylaxis and Risk Stratification Using the CHA2DS2-VASc Score.” *Cardiology Journal*, vol. 21, no. 5, 2014, pp. 451–53.
12. McEwan, P., et al. “Modelling Cardiovascular Outcomes in Type 2 Diabetes in the Era of Cardiovascular Outcomes Trials.” *Value in Health*, vol. 20, Oct. , p. A747.
13. Petretta, M., et al. “Cardiovascular Risk Stratification in Diabetic Patients.” *Clinical and Translational Imaging*, vol. 1, no. 5, 1, pp. 325–39.
14. Scirica, B. M. “Use of Biomarkers in Predicting the Onset, Monitoring the Progression, and Risk Stratification for Patients with Type 2 Diabetes Mellitus.” *Clinical Chemistry*, vol. 63, no. 1, Jan. , pp. 186–95.
15. Targher, G., and C. D. Byrne. “Treatment Algorithm in Patients with Type 2 Diabetes and Atherosclerotic Cardiovascular Disease or High/Very High Cardiovascular Risk.” *European Heart Journal*, vol. 41, no. 2, 1, p. 331.
16. Verma, Subodh, et al. “Predictors of Heart Failure Development in Type 2 Diabetes: A Practical Approach.” *Current Opinion in Cardiology*, vol. 34, no. 5, 2019, p. 578.
17. Wang, Juan, et al. “Novel Biomarkers for Cardiovascular Risk Prediction.” *Journal of Geriatric Cardiology: JGC*, vol. 14, no. 2, 2017, p. 135.

EXCLUDED: INSUFFICIENT INFORMATION FOR ACTUAL USE

1. Dworzynski, P., et al. “Nationwide Prediction of Type 2 Diabetes Comorbidities.” *Scientific Reports*, vol. 10, no. 1, Feb. .
2. Ferket, B. S., et al. “Lifetime Predictions of Coronary Heart Disease, Congestive Heart Failure, and Stroke to Enable Preventive Treatment Decisions in Type 2 Diabetes Patients.” *Circulation. Conference: American Heart Association’s*, vol. 134, 2016.
3. Ferreira, J. P., et al. “Multi-Proteomic Approach to Predict Specific Cardiovascular Events in Patients with Diabetes and Myocardial Infarction: Findings from the EXAMINE Trial.” *Clinical Research in Cardiology*, vol. 13, Aug. , p. 13.
4. Gori, Mauro, et al. “Natriuretic Peptide and High-Sensitivity Troponin for Cardiovascular Risk Prediction in Diabetes: The Atherosclerosis Risk in Communities (ARIC) Study.” *Diabetes Care*, vol. 39, no. 5, 2016, pp. 677–85.
5. Hong, D., et al. “Pdb130 Risk Equations for Complications of Type2 Diabetes Using Real-World Electronic Medical Records: Localized Vs. Generalizable.” *Value in Health*, vol. 22, May , p. S162.
6. Ohkuma, T., et al. “Cardiac Stress and Inflammatory Markers as Predictors of Heart Failure in Patients With Type 2 Diabetes: The ADVANCE Trial.” *Diabetes Care*, vol. 40, no. 9, Sept. , pp. 1203–09.
7. Patel, T., et al. “Predictors of Hospitalization or Death Due to Heart Failure in Diabetic Patients by Gender in the Accord Trial Using Random Survival Forests.” *Circulation. Conference: Resuscitation Science Symposium, ReSS*, vol. 136, 2017.
8. Scirica, B., et al. “Prognostic Implications of Simultaneous Biomarker Assessments in Patients with Type 2 Diabetes Mellitus: Observations from the Savor-TIMI 53 Trial.” *Journal of the American College of Cardiology*, vol. 63, no. 12, 1, p. A1550.
9. Shao, H., et al. “Globalization Module for a Diabetes Progression Prediction Model: The Building, Relating, Acting, and Validating Outcomes (BRAVO) Model.” *Value in Health*, vol. 21, May , p. S5.
10. Sharma, A., et al. “Prognostic Utility of GDF-15 in Chronic Heart Failure: Insights from the HF-ACTION Study.” *European Heart Journal*, vol. 37, Aug. , p. 441.
11. Ward, T., et al. “Investigating the Validity of the UKPDS Outcomes Equations in Current Clinical Practice.” *Diabetologia*, vol. 56, Sept. , pp. S13–14.
12. Wells, B. J., et al. “Prediction of Morbidity and Mortality in Patients with Type 2 Diabetes.” *Peerj*, vol. 1, June.

EXCLUDED: DUPLICATES NOT REMOVED AT TITLE / ABSTRACT SCREENING

1. Berg, D., et al. “Heart Failure Risk Stratification and Efficacy of Dapagliflozin in Patients with Type 2 Diabetes Mellitus.” *European Heart Journal*, vol. 40, Oct. , p. 153.
2. Elharram, M., et al. “Predicting Heart Failure Hospitalization in Type 2 Diabetes Mellitus: Validation of the TRS-HFDM Score in the ACCORD Trial.” *Circulation*, vol. 142, Nov. .
3. Hayden, Jennifer D., et al. *404-P: Cardiovascular Risk Scoring and Stratification in Patients with Type 2 Diabetes Enrolled in a Medicare Advantage Plan*. Am Diabetes Assoc, 2020.
4. Idzerda, N. M. A., et al. “Prediction of the Effect of Dapagliflozin on Kidney and Heart Failure Outcomes Based on Short-Term Changes in Multiple Risk Markers.” *Nephrology Dialysis Transplantation*, vol. 35, no. 9, Sept. , pp. 1570–76.
5. Naccarelli, G. V., et al. “CHADS2 and CHA2DS2-VASc Risk Factors to Predict First Cardiovascular Hospitalization among Atrial Fibrillation/Atrial Flutter Patients.” *American Journal of Cardiology*, vol. 109, no. 10, 15, pp. 1526–33.
6. R, Pfister, et al. “A Clinical Risk Score for Heart Failure in Patients with Type 2 Diabetes and Macrovascular Disease: An Analysis of the PROactive Study.” *International Journal of Cardiology*, vol. 162, no. 2, 2013, pp. 112–16.
7. Sharma, A., et al. “Evaluating Expanded Heart Failure Outcomes in Patients with Type 2 Diabetes after an Acute Coronary Syndrome: Insights from the Examine Trial.” *Circulation. Conference: Resuscitation Science Symposium, ReSS*, vol. 136, 2017.
8. Williams, B. A., et al. “A Risk Prediction Model for Heart Failure Hospitalization in Type 2 Diabetes Mellitus.” *Clinical Cardiology*, vol. 43, no. 3, Mar. , pp. 275–83.
9. X, Yang, et al. “Development and Validation of a Risk Score for Hospitalization for Heart Failure in Patients with Type 2 Diabetes Mellitus.” *Cardiovascular Diabetology*, vol. 7, 2008, p. 9.

EXCLUDED: MISSING PERFORMANCE MEASURES

1. Idzerda, N., et al. “Prediction of the Effect of Dapagliflozin on Renal and Heart Failure Outcomes Based on Short-Term Changes in Multiple Risk Markers.” *Journal of the American Society of Nephrology*, vol. 29, 2018, p. 10.
2. Scott, C. A. B., et al. “Predictors of Incident Hospitalization for Heart Failure in the ACE Trial Population.” *Diabetes. Conference: 79th Scientific Sessions of the American Diabetes Association, ADA*, vol. 68, 2019, https://diabetes.diabetesjournals.org/content/68/Supplement_1/1479-P.
3. Verma, S., et al. “Application of the Timi Heart Failure Risk Score to the Empa-Reg Outcome Population.” *Journal of the American College of Cardiology*, vol. 75, 24, p. 1851.
4. Verma, S., et al. “Empagliflozin Reduces the Risk of Mortality and Hospitalization for Heart Failure across Thrombolysis In Myocardial Infarction Risk Score for Heart Failure in Diabetes Categories: Post Hoc Analysis of the EMPA-REG OUTCOME Trial.” *Diabetes Obesity & Metabolism*, vol. 22, no. 7, July , pp. 1141–50.
5. Zethelius, B., et al. “Impact of Socioeconomic Status and Ethnicity on Risk of Stroke, Hospitalisation for Heart Failure and Death in 371,092 Individuals with Type 2 Diabetes.” *Diabetologia*, vol. 58, no. 1, Sept. , p. S433.
6. Scirica, B. M., et al. “Prognostic Implications of Biomarker Assessments in Patients With Type 2 Diabetes at High Cardiovascular Risk A Secondary Analysis of a Randomized Clinical Trial.” *Jama Cardiology*, vol. 1, no. 9, Dec. , pp. 989–98.
7. Shao, H., et al. “Using the BRAVO Risk Engine to Predict Cardiovascular Outcomes in Clinical Trials With Sodium-Glucose Transporter 2 Inhibitors.” *Diabetes Care*, vol. 43, no. 7, July , pp. 1530–36.

Figure S1. Risk of Bias and Applicability of Included Clinical Prediction Model Development Studies.

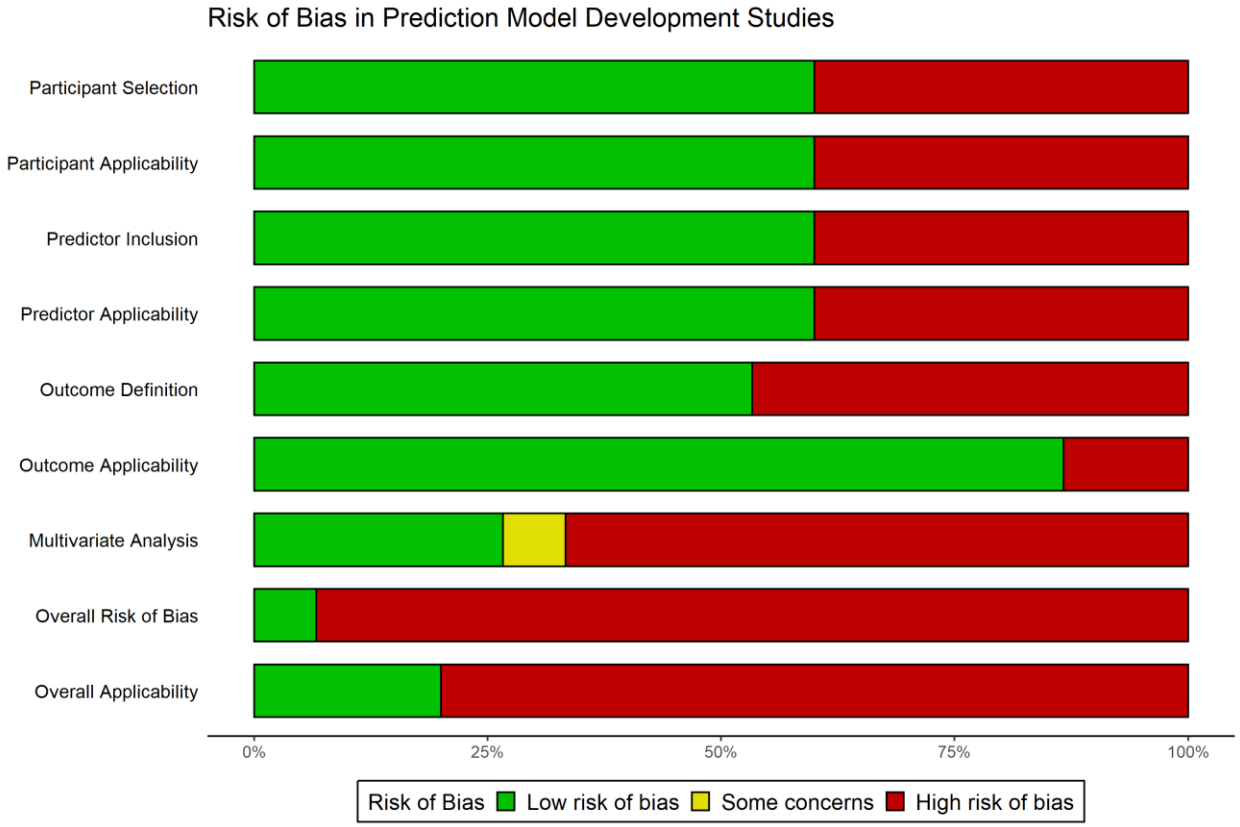

Figure S2. Risk of Bias and Applicability of Included Clinical Prediction Model Validation Studies.

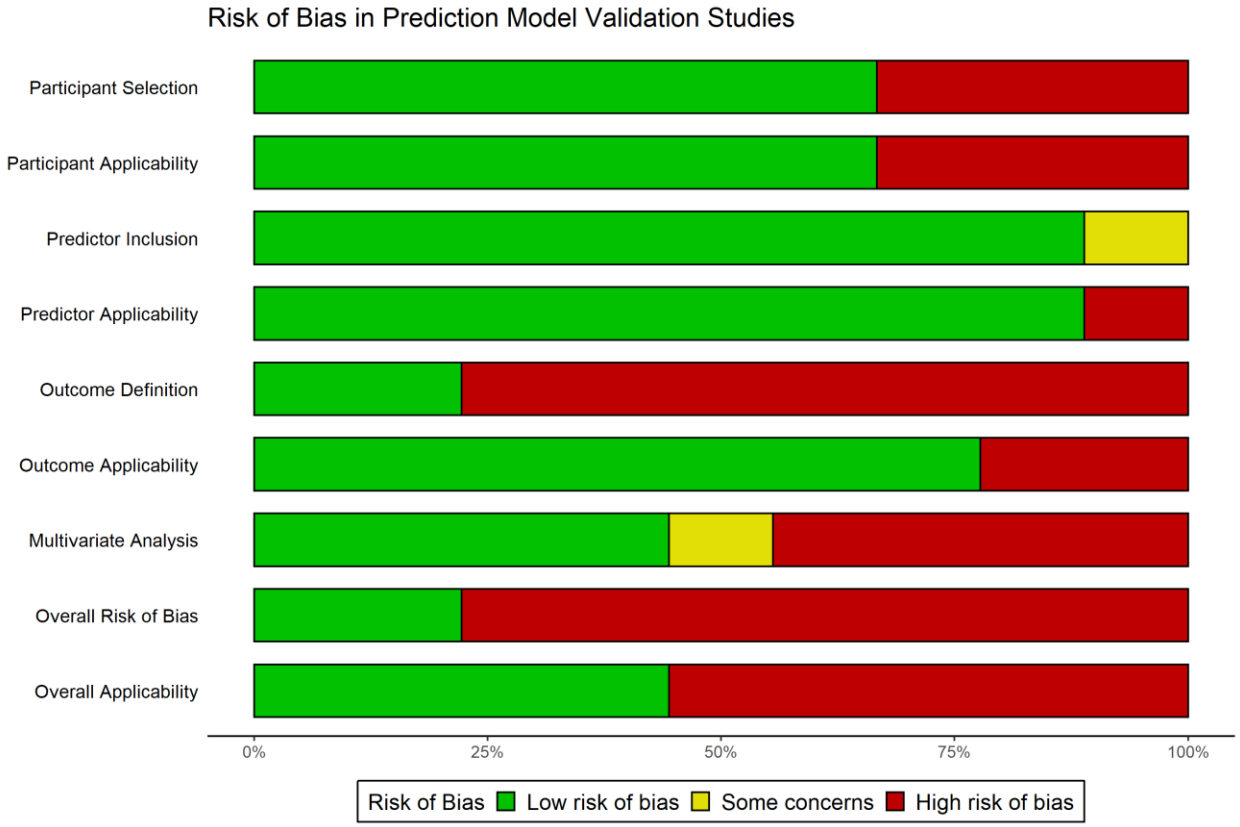

Figure S3. Risk of Bias and Applicability of Included Clinical Prediction Model Development Studies.

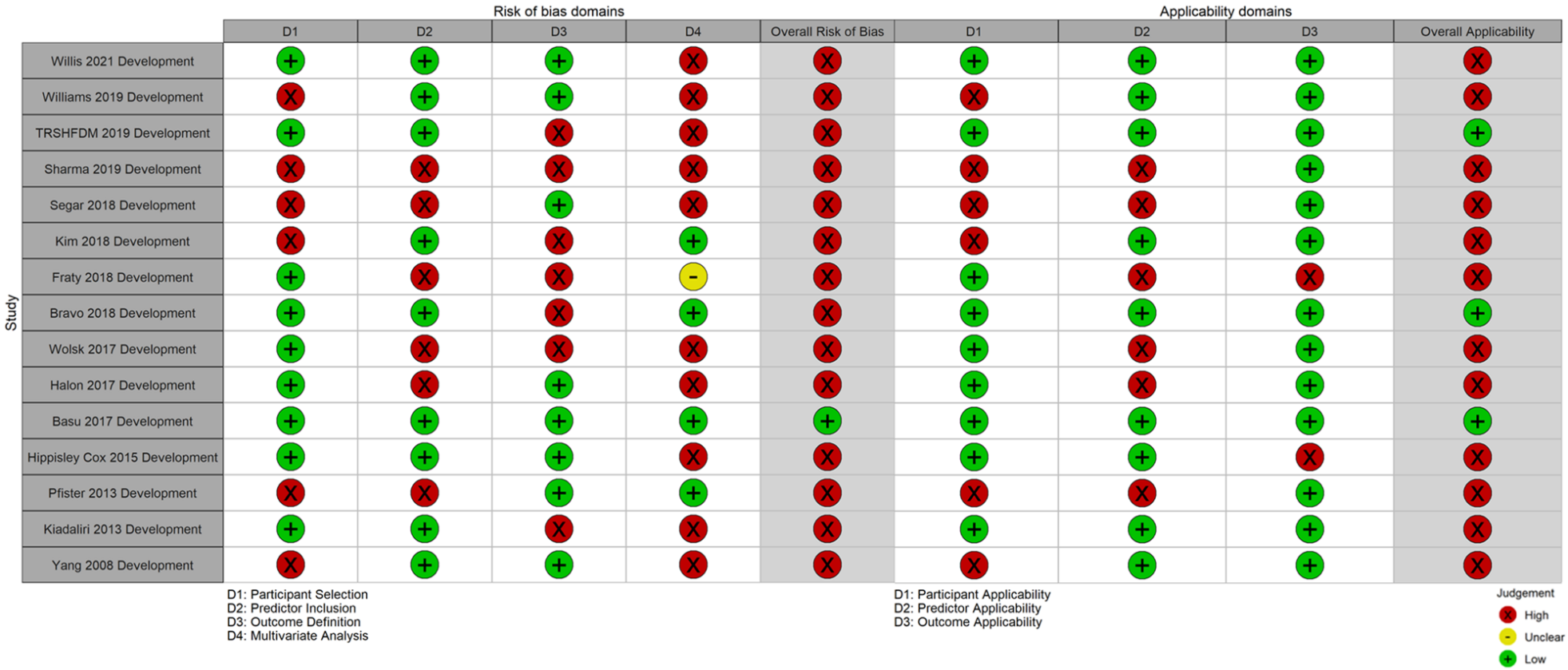

Figure S4. Risk of Bias and Applicability of Included Clinical Prediction Model Validation Studies.