ORIGINAL RESEARCH

Using Primary Health Care Electronic Medical Records to Predict Hospitalizations, Emergency Department Visits, and Mortality: A Systematic Review

Rebecca Johnson, BSc, Thomas Chang, BHSc, Rahim Moineddin, PhD, Tara Upshaw, BSc, MHSc, Noah Crampton, MD, CCFP, MSc, Emma Wallace, MB, BAO, BcH, PhD, MICGP, and Andrew D. Pinto, MD, CCFP, FRCPC, MSc

Introduction: **High-quality primary care can reduce avoidable emergency department visits and emergency hospitalizations. The availability of electronic medical record (EMR) data and capacities for data storage and processing have created opportunities for predictive analytics. This systematic review examines studies which predict emergency department visits, hospitalizations, and mortality using EMR data from primary care.**

Methods: **Six databases (Ovid MEDLINE, PubMed, Embase, EBM Reviews (Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials, Cochrane Methodology Register, Health Technology Assessment, NHS Economic Evaluation Database), Scopus, CINAHL) were searched to identify primary peer-reviewed studies in English from inception to February 5, 2020. The search was initially conducted on January 18, 2019, and updated on February 5, 2020.**

Results: **A total of 9456 citations were double-reviewed, and 31 studies met the inclusion criteria. The predictive ability measured by C-statistics (ROC) of the best performing models from each study ranged from 0.57 to 0.95. Less than half of the included studies used artificial intelligence methods and only 7 (23%) were externally validated. Age, medical diagnoses, sex, medication use, and prior health service use were the most common predictor variables. Few studies discussed or examined the clinical utility of models.**

Conclusions: **This review helps address critical gaps in the literature regarding the potential of primary care EMR data. Despite further work required to address bias and improve the quality and reporting of prediction models, the use of primary care EMR data for predictive analytics holds promise. (J Am Board Fam Med 2024;37:583–606.)**

Keywords: **Artificial Intelligence, Electronic Health Records, Emergency Room Visits, Hospitalization, Primary Health Care, Systematic Review**

Introduction

Primary health care is the foundation of health systems. High-quality primary care reduces the need for more expensive acute health services and is associated with improved population health outcomes.[1,2] Strengthening primary health care has a

This article was externally peer reviewed.

Submitted 20 October 2023; revised 11 February 2024; accepted 12 February 2024.

From the Upstream Lab, MAP Centre for Urban Health Solutions, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Ontario, Canada (RJ, ADP); Institute for Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada (RJ, RM, ADP); Undergraduate Medical Education, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada (TC); Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada (TU); Department of Family and Community Medicine, Toronto Western Hospital, University Health Network, Toronto, Ontario, Canada (NC); Department of Family and Community Medicine, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada (NC, ADP); Department of General Practice, University College Cork, Cork, Ireland (EW); Department of Family and Community Medicine, St. Michael's Hospital, Toronto, Ontario, Canada (ADP).

Funding: This project was supported in part by an Early Researcher Award from the Government of Ontario. Andrew Pinto is supported as a Clinician-Scientist by the Department of Family and Community Medicine, Faculty of Medicine at the University of Toronto and St. Michael's Hospital, and the Li Ka Shing Knowledge Institute, Unity Health Toronto. He is also supported by a CIHR Applied Public Health Chair. The opinions, results and conclusions reported in this article are those of the authors and are independent of any funding sources.

Conflict of interest: The authors have no conflicts of interest to declare.

Corresponding author: Andrew D. Pinto, MD, CCFP, FRCPC, MSc, Upstream Lab, 30 Bond Street, Toronto,

direct impact on system performance and resiliency during public health emergencies.[3,4] Rising health care costs[5–7], increases in health services utilization,[8] and the limits to the capacities of acute care services[9–11], drive the need for more proactive and preventative interventions in primary health care.[12] Past research has identified that some hospitalizations and a significant proportion of emergency department visits are preventable and amenable to primary care intervention.[13]

Primary health care offers a unique and effective setting to intervene to reduce acute health service use, the need for costly interventions, and to reduce premature mortality.[14] Greater engagement by patients with primary care has been associated with decreased risk of emergency hospitalization,[15] and emergency department visits,[16] and early contact after hospital discharge has been found to reduce readmissions by as much as 50%.[17,18] Primary health care plays a significant role in patient coordination of care and the redistribution of health system burden and resource use.[19]

We now have an opportunity to implement data-driven approaches to support clinical decision making and to reduce acute care service use through proactive care within primary care settings.[9,20] Over the past decade, the adoption of electronic medical records (EMR) within primary health care has gained momentum. According to the Commonwealth Fund International Health Policy survey, the number of family physicians who report using EMRs in practice has grown steadily in recent years.[21] This is particularly true in the United States and Canada, where rates of EMR use have doubled over 10 years (46% to 92% and 37% to 86%, respectively). As of 2019, an average of 93% of primary care physicians report using EMRs in practice internationally.[22,23] The longitudinal nature and population-based health approach of primary care means that primary care EMRs offer a rich source of data that holds the potential for use in predictive analytics.[24,25] Furthermore, the growth in primary care EMR data availability, coupled with advances in data storage and processing capabilities, have paved the way for new technologies, such as artificial intelligence, to improve medical care.[26,27]

Ontario, Canada, M5B 1W8 (E-mail: andrew.pinto@utoronto.ca).

Expanding on prior systematic reviews that have explored the use of prediction models for identifying hospitalizations, ED visits, or mortality, there are notable gaps in the literature regarding the use of primary care EMR data. Although prior reviews have made significant contributions in predicting these outcomes, they predominantly feature studies reliant on hospital data, or administrative health databases such as physician billing claims databases, rather than solely focusing on EMR data.[28–31] Moreover, even within reviews focused on EMR data, substantial gaps remain, particularly emphasizing the use of primary care EMR data.[32–34] Furthermore, many of these studies have found variable performance and modest discriminative ability. Given the longitudinal nature and richness of primary care EMR data, using such data in prediction models has the potential to significantly enhance model performance.

Despite the need and opportunity for data-driven, proactive primary care interventions, little is known about the prevalence, rigor, and clinical suitability of prediction models that process primary care data to predict of emergency department visits, hospitalizations, and mortality. Currently, no previous review has thoroughly examined the use of primary care EMR data in these outcomes. This review aims to address these intersecting gaps, providing a nuanced understanding of prediction models which use EMR data within primary care settings. Considering the gaps in literature, this systematic review aims to address the following primary research question: what are the published studies on the use of primary health care EMR data to predict emergency department visits, hospitalizations, and mortality? Our objectives were threefold: 1) to examine candidate predictors which contribute to high-performance prediction models; 2) describe model performance; and 3) identify and report on existing model's contributions to clinical care and decision making.

Methods

A systematic review was conducted to identify all relevant studies on the use of primary care electronic medical record data to predict emergency department visits, hospitalizations, and mortality. The study protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO, registration: CRD42020136625), and results are presented in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

Databases and Search Strategy

Six electronic databases (Ovid MEDLINE, PubMed, Embase, EBM Reviews (Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials, Cochrane Methodology Register, Health Technology Assessment, NHS Economic Evaluation Database), Scopus, CINAHL) were used to search the peer-reviewed literature from inception. The search was initially conducted on January 18, 2019, and updated on February 5, 2020. The Journal of Medical Internet Research (JMIR), Journal of Medical Informatics (JMI) and the Journal of the American Medical Informatics Association (JAMIA), as well as the reference lists of selected studies, were hand-searched for additional citations. In consultation with an information specialist, specific search strategies were developed for each database (sample search strategy can be found in Appendix A).

Eligibility Criteria

Studies were included if they were: (1) primary quantitative studies; (2) evaluated the performance of a new statistical or mathematical model, algorithm, or other forms of model, or external validation of an existing model; (3) predicted a single endpoint outcome of either emergency department visits, hospitalization, or mortality; (4) used electronic medical record data from outpatient primary health care. We included studies related to people of any age. Exclusion criteria include data originating from article records or sources outside of primary care, such as emergency departments, census data, or surveys. Commentaries, editorials, thesis dissertations, and reviews were also excluded. Included studies were restricted to those published in English; however, no restrictions were made to the search by country.

Study Identification

After removing duplications, all relevant citations were imported into DistillerSR (Evidence Partners, Ottawa, ON) to support citation management, screening, and conflict resolution. A number of study volunteers were engaged and trained to assist with reviewing citations. Title and abstracts were screened by 2 independent reviewers to assess for study inclusion. The full texts that were considered suitable for inclusion were reviewed by 2 independent members of the study team to ensure eligibility and then proceeded to data extraction. Any disagreements between reviewers were addressed by the study team and resolved with the principal investigator (ADP) as the arbitrator.

Data Extraction

A standardized data extraction form was developed in accordance with guidelines established by the Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies (CHARMS)[35] and was prepiloted by study team members (Appendix B). Relevant data were extracted in duplicate (R.J., T.C.), and disagreements were resolved through study team meetings. The extracted data included information on the study setting and participants, data sources, outcomes, predictor variables, sample size, missing data, model development, and results. If information was not available from an article it was noted during data extraction.

Data Synthesis and Analysis

Meta-analysis was considered to examine pooled outcomes of model performance, however, heterogeneity in terms of study settings, populations, data, and outcomes precluded this approach, and it was deemed not feasible for this study. Therefore, results were examined using narrative synthesis. To bring together key features of model development and performance, results were presented according to the prediction outcome of interest when appropriate and organized as follows: study setting and population, predictor variables considered and included in final models, prediction outcomes, and model development and performance. Discrimination, the model's ability to accurately distinguish between individuals who experience the event of interest and those who do not, was used to assess predictive performance and reported using concordance (c) statistics, or area under the curve (AUC) with 95% confidence intervals when available.[36] A c statistic of 0.5 indicates no discriminative ability, akin to random chance. Optimizing predictive accuracy and discriminative ability is crucial for real-world application, instilling confidence in the model's ability to inform clinical decision making. The performance of predictive models carries significant clinical implications for patients, care clinicians, and health system resources.[36] Instead of *c*-statistics or AUC measures, sensitivity, specificity, negative predictive value, and positive predictive value were reported when necessary.

Quality Appraisal and Risk of Bias Assessment

The quality and risk of bias (ROB) of individual studies were assessed using the Prediction Model Risk of Bias Assessment Tool (PROBAST).[37] Each study was independently evaluated (R.J., T.C.) using the provided signaling questions and subsequently 7afforded a score of "high," "low," or "unclear" risk of bias in accordance with PROBAST scoring guidelines.

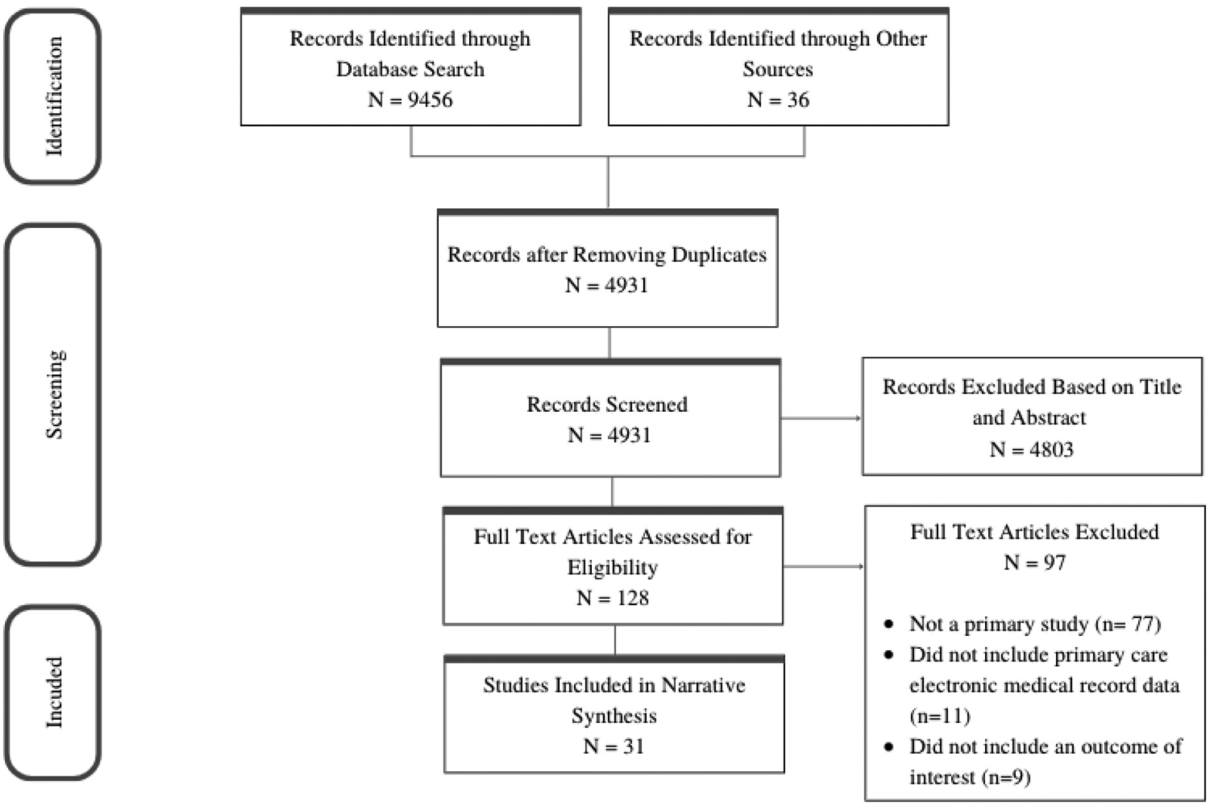
Results

The electronic databases search strategy retrieved 9456 studies. Of these, 4967 studies were screened by title and abstract, resulting in 164 being selected for full-text review. The final review included 31 studies which met inclusion criteria (Figure 1). Overall, of the analyzed risk prediction models, 16 (44.4%) aimed to identify individuals at risk of hospitalization, 7 (19.4%) focused on the risk of ED visits, and 13 (36.1%) on the risk of mortality. Although a subset of these models focused on specific patient populations, such as those with chronic kidney disease or HIV-positive patients admitted to hospital, the majority were developed in an adult general practice or community hospital population.

Study Characteristics

Within the 31 included studies, most (n = 18, 58.1%) reported on models developed using data from the United States.[38–55] The remainder of the studies were from the United Kingdom (n = 7, 22.6%)[56–62], Canada,[63] Ireland,[64] and Australia[65] (n = 1, 3.2% each), and Israel (n = 2, 6.5%)[66,67], whereas 1 study utilized data from health centers across multiple countries.[68] Among the 31 included studies, 5 studies utilized unique models to address

Figure 1. PRISMA flow diagram. ***Abbreviation:*** **PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.**



more than 1 outcome, creating a total of 36 prediction models. Sixteen models (44.4%) focused on the prediction of hospitalization; of these, 5 focused on the risk of hospital readmission.[44,46,47,66,67] Thirteen models (36.1%) were developed to predict mortality, including 3 which specifically examined the outcome of patient suicide mortality,[49,51] and 1 risk of opioid overdose.[50] Furthermore, 7 models (19.4%) assessed the risk of emergency department visits.[38–42,63,65] A majority of studies (n = 30, 96.7%) implemented a retrospective cohort design, 2 of which coupled this with prospective cohort validation, and 1 study exclusively utilized a prospective cohort design.[28]

Of the total studies included in this review, 23 (74.2%) utilized electronic medical record data in combination with other large sources of data, including health care administrative data (n = 7) death repositories (n = 7) and population demographics data (such as census data) (n = 8). In addition to primary care medical record data, 6 (19.4%) studies included in this review utilized data linkages to capture hospital and health service use, but more commonly, this data were often stored in tandem with primary care data (n = 25, 80.6%). Of the 25 studies which developed models using a single source of EMR data, more than half (n = 13) used data housed within institutional or national health care repositories. The total sample size in each study ranged from 607 patients to >4.6 million patients. A complete overview of study characteristics is presented in Table 1.

Model Characteristics

Several of the studies evaluated the predictive performance of models across a number of outcomes (n = 8), populations (n = 4), prediction windows (n = 4), or across a number of prediction methods (n = 3) and variable subsets (n = 8). The most frequently used statistical analysis method was logistic regression, which was used in 17 studies (54.8%). Of the total, 13 studies (41.9%) developed predictive models using artificial intelligence methods. Most validated models internally (n = 20, 64.5%) or used a combination of both internal and external validation (n = 7, 22.6%). Most commonly, methods of internal validation used a split sample approach (n = 16, 51.6%) or the more comprehensive technique of cross-validation (n = 6, 19.4%), which assesses model performance across multiple subsets of data.

Four studies[28,39,45,56] (12.9%) discussed the clinical implementation of prediction models or developed models for real-time utilization in clinical practice and to inform clinical decision making. More specifically, Morawski, Dvorkis, & Monsen (2020) developed a model that utilizes a data warehouse, which couples clinical EMR data with 12-month administrative claims and updates data on a weekly basis to predict real-time risk of hospitalization for primary care patients. Similarly, Hu et al., (2015), developed a model to assess patient risk of ED visits and integrated this tool into an accessible, web-based dashboard used to inform care. Although prospective cohort studies explore prediction performance in real-time settings, Wallace et al., (2016) specifically note model characteristics which have advantages and lend to clinical utility, such as high-risk stratification. Lastly, the Devon Predictive Model identifies clinicians of patients within the top 5% highest risk of hospital admission to inform case management care.[69]

Inclusion and Use of Prediction Variables

Table 2 presents the 20 most common variables included in final models across all outcomes. Within the 31 studies, 51 various sets of predictive features were examined. Prediction features were classified into 9 broad categories: sociodemographic variables, patient health profile, medical history, medication use, clinical findings, procedure history, health service utilization, social supports, and others.

Across all prediction outcomes, age (n = 45, 88%), specific medical diagnoses (n = 43, 84%), and sex (n = 40, 78%) were the top 3 variables most frequently included in final models. Twenty-nine (57%) models included medication use, a variable often captured by a count of individual prescriptions, medication classifications, or drug classes. Prior health service utilization, often examined within the past year, is a key variable commonly included in final prediction models. This included both prior hospital admissions and emergency department visits. Less frequently, the number of inpatient bed days was included as a final prediction variable. Across all outcomes, medication use was frequently included in final prediction models (ED visits, n = 4, 50%; hospitalizations, n = 13, 62%; mortality, n = 12, 75%). Clinical laboratory results, such as levels of bilirubin and creatinine, were included in 22 final models (43%), and

Table 1. Characteristics of Studies and Model Development

Study, Country	Model Name (If Applicable)	Study Population, Setting, Time Period	Outcome(s)	Sample Size	Outcome Rate	Modelling Methods	Artificial Intelligence Methods Used†	Validation Approach	Validated Model*c-statistic** (95% CI)
Model Outcome: Emergency Department Visits									
Frost DW, et al., 2017 Canada		Aged ≥50, ≥1-year General Practice, 2011 to 2012	Frequent ED Use (≥3 ED visits) (12-month)	Derivation: 21,680 Validation: 895	Training Cohort = 5.7%	Logistic Regression	Yes	Internal: Split sample	Validation *c* = 0.71 (no CIs)
Howell P, & Elkin PL., 2019 United States		Not Reported, 2017 to 2018	Emergency Department Visit (12-month)	2991	Not Reported	Random Forest	Yes	External validation only	Validation *c* = 0.83 (no CIs)
Pearce et al., 2019 Australia	POLAR Diversion	All patients, General Practice, 2010 to 2015	Emergency Department Visit (12-month)	37,665	≥1 ED Visit = 23%	Support Vector Machine	Yes	Internal: Split sample 10-Fold Cross-validation	0 to 30 days: Sensitivity = 68%; PPV = 73.7% 31 to 365 days: Sensitivity = 10%; PPV = 36.8%
Hu Z, et al., 2015 United States		All patients, 2012 to 2013	Emergency Department Visit (6-month, all-cause)	Retrospective cohort: 829,641 Prospective cohort: 875,979	Retrospective = 11.48%, Prospective = 11.37%	Survival Forest Decision Trees	Yes	External validation: Prospective	Retrospective *c* = 0.74 (no CIs) Prospective *c* = 0.73 (no CIs)
Hao et al., 2014 United States		All patients, 2012 to 2013	Emergency Department Revisit (30-days)	(ED encounters) Derivation: 293,461 Prospective validation: 193,886	Retrospective = 19.4% Prospective = 20.5%	Decision Tree	Yes	Split sample External validation: Prospective	Retrospective *c* = 0.71 (no CIs) Prospective *c* = 0.70 (no CIs)
Bhavsar et al., 2018 United States	(1) Electronic Health Record (HER) (2) EHR + Neighborhood Socioeconomic Status (SES)	Aged ≥18, ≥1 health care encounter in previous year, Durham County resident, 2009 to 2015	Emergency Department Visit	Derivation: 90,097 Validation: 122,812	Not Reported	Random Survival Forest	Yes	Temporal Split	(1) *c* = 0.75 (no CIs) (2) *c* = 0.75 (no CIs)
Crane et al., 2010 United States		Aged ≥60, inpatient, primary care community dwelling, assisted living patients, 2005 to 2006	Emergency Room Visit (2-year)	12,650	Not Reported	Logistic Regression	No	Bootstrapping (450 samples)	AUC = 0.64 (no CIs)

Continued

Table 1. Continued

Study, Country	Model Name (If Applicable)	Study Population, Setting, Time Period	Outcome(s)	Sample Size	Outcome Rate	Modelling Methods	Artificial Intelligence Methods Used†	Validation Approach	Validated Model *c-statistic** (95% CI)
Model Outcome: Hospitalizations									
Rahimian et al., 2018 United Kingdom	Predictor Sets (1) QAdmissions (QA) (2) QAdmissions+ (QA+) (3) Temporal (T)	Aged 18 to 100, ≥1-year at General Practice Clinic, 1985 to 2015	Emergency hospital admission (12-, 24-, 36-, 48-, 60-months)	Total = 4637,297 Derivation= 3749,932 Validation= 887,365	Derivation Avg =7.8% Validation Avg =10.4%	(1) Cox proportional hazards (CPH) (2) Gradient boosting classifier (GBC) (3) Random forest (RF)	Yes	Split Sample (80/20), 5-fold Cross validation External validation	(1) GBC AUC = 0.80 (no CIs) (2) GBC AUC = 0.81 (no CIs) (3) GBC AUC = 0.83 (no CIs)
Gao et al., 2014 United States	Predictor Sets: (1) Hospital characteristics + patient demographic, socioeconomic variables (2) Model 2 + prior year utilization, cost (3) Model 3: + 394 HCCs	Veterans Affairs (VA) Patients Treated for Ambulatory Care Sensitive Conditions (ACSCs), 2011 to 2012	ACSC Hospital Admission (90-days, 1-year)	2987,052	90-day admission =0.73% 1-year admission =2.39%	Logistic regression (hierarchical)	No	Split sample (50/50), 2-fold cross-validation	90-day: (2) $c = 0.72$ (0.72-0.73) (3) $c = 0.83$ (0.82-0.83) (4) $c = 0.86$ (0.85-0.86) 1-year: (3) $c = 0.83$ (0.83-0.84)
Perkins et at., 2013 United States		Aged 18 to 88, General practice patients with chronic kidney disease stage ≥3, hospitalized with heart failure diagnosis, 2004 to 2010	Hospital readmission (30-day)	607	19.10%	Logistic regression (multivariate)	No	Bootstrap resampling with 1000 samples	$c = 0.74$ (no CIs)
Morawski et al., 2020 United States	Predictor Sets: (1) Electronic Health Record (EHR) (2) EHR + Claims Data	Aged ≥18, Medicare/ Medicaid insured patients, 2013 to 2015	Hospital admission (6-month)	185,388	5%	Logistic regression	No	Split sample (80/20)	(2) AUC = 0.84 (0.83-0.85) (3) AUC = 0.84 (0.84-0.85)

Continued

Table 1. Continued

Study, Country	Model Name (If Applicable)	Study Population, Setting, Time Period	Outcome(s)	Sample Size	Outcome Rate	Modelling Methods	Artificial Intelligence Methods Used†	Validation Approach	Validated Model *c-statistic** (95% CI)
Shadmi et al., 2015 Israel	Preadmission Readmission Detection Model (PREADM)	Aged ≥18, general medicine patients admitted to hospital, 2009 to 2010	Hospital readmission (30-day, all-cause)	Total Admissions = 33,639 Derivation = 22,406 Validation = 11, 233	16.80%	Logistic regression (multivariate)	Yes	Split sample (33/66) External validation: temporal	External validation $c = 0.69$ (no CIs)
Brisimi et al., 2018 United States		All ages, patients with ≥1 heart-related diagnosis or procedure, or diabetes mellitus diagnosis, 2001 to 2010	Hospitalization (heart-disease related, 12-month) Hospitalization (diabetes related, 12-month)	Heart disease = 45,579 Diabetes = 33,122	Heart-disease dataset = 6.7% Diabetes dataset = 17.0%	(1) SVM (2) Random Forest (RF) (3) Sparse Logistic Regression (LR) (4) K-Likelihood Ratio Test (k-LRT) (5) Joint Clustering and Classification (JCC)	Yes	Split sample (60/40) RF: bootstrapping with replacement, cross-validation ACC: out of sample classification	Heart Data: Random Forest AUC = 81.62% (no CIs) Diabetes Data: Random Forest AUC = 84.53% (no CIs)
Wallace et al., 2016 Ireland	(1) Original Pra (2) Modified Pra	Aged ≥ 70, general medical services, 2010 to 2012	Emergency hospital admission (12-month, all-cause)	862	18%	Logistic regression	No	External validation: Prospective	(1) $c = 0.65$ (0.61-0.7) (2) $c = 0.67$ (0.62-0.72)
Chenore et al., 2013 United Kingdom		All patients, general practice in Devon UK, 2008 to 2010	Emergency hospital admission (12-month)	722,383	Emergency admissions = 5.6%	Binary logistic regression	No	Split sample (80/20)	$c = 0.78$ (0.78-0.78)
Donnan et al., 2008 United Kingdom	Predicting Emergency Admissions Over the Next Year (PEONY)	Aged ≥40, general practice, residing in Tayside, Scotland, 1996 to 2004	Emergency hospital admission (12-month)	Total = 186,523 Derivation = 90,522 Validation = 96,001	Derivation =7.5%	Logistic regression	No	Split sample (50/50)	$c = 0.79$ (no CIs)
Hippisley-Cox et al., 2013 United Kingdom	(1) QAdmissions Score (2) QAdmissions + Hospital Episode Statistics Linked Data	Aged 18 to 100, ≥1-year general practice, 2010 to 2011	Emergency hospital admission (1- and 2-year)	Derivation = 2849,381 Internal Validation = 1340,622 External Validation = 2475,360	Derivation = 9.3% Internal validation = 9.9% External validation = 9.5%	Cox proportional hazards (CPH)	No	Internal validation (75% of practices) External validation	(1) Women $c = 0.76$ (0.76-0.77), Men $c = 0.767$ (0.76-0.77) (2) Women $c = 0.77$ (0.77-0.77), Men $c = 0.77$ (0.77-0.77)

Continued

Table 1. Continued

Study, Country	Model Name (If Applicable)	Study Population, Setting, Time Period	Outcome(s)	Sample Size	Outcome Rate	Modelling Methods	Artificial Intelligence Methods Used†	Validation Approach	Validated Model *c-statistic** (95% CI)
Watson et al., 2011 United States		Patients hospitalized and discharged with heart failure diagnosis, 2007 to 2008	Hospital readmission (30-day, all-cause)	790	12.8%	Logistic regression (multivariate)	No	None	*c* = 0.67 (no CIs)
Zeltzer, 2019 Israel	(1) Claims and EMR full (2) Claims, EMR and Admission full	Patients admitted to hospital overnight, 2016 to 2017	Readmission to hospital (30-day)	144,966 index hospital admissions (118,510 patients)	14.7%	Extreme gradient boosting	Yes	Split sample (83/17) 10-fold cross validation, bootstrapping	(1) AUC = 0.70 (0.69-0.71) (2) AUC = 0.71 (0.70-0.72)
Bhavsar et al., 2018 United States	(1) Electronic Health Record (HER) (2) EHR + Neighborhood Socioeconomic Status (SES)	Aged ≥18, ≥1 health care encounter in previous year, Durham County resident, 2009 to 2015	Inpatient encounters; Hospitalizations due to accidents, asthma, influenza, myocardial infarction, stroke	Derivation= 90,097 Validation= 122,812	Not Reported	Random Survival Forest	Yes	Temporal split	Hospitalizations: Myocardial infarction (1) *c* = 0.892 (2) *c* = 0.892 Stroke: (1) *c* = 0.854 (2) *c* = 0.855 Asthma: (1) *c* = 0.752 (2) *c* = 0.756 Accident: (1) *c* = 0.747 (2) *c* = 0.755 Influenza: (1) *c* = 0.562 (2) *c* = 0.565 Inpatient visit: (1) *c* = 0.740 (2) *c* = 0.742
Crane et al., 2010 United States		Aged ≥60, inpatient, primary care community dwelling, assisted living patients, 2005 to 2006	Hospitalizations (total number) (2-year)	12,650	Not Reported	Logistic Regression	No	Bootstrapping (450 samples)	AUC = 0.705 (no CIs)

Continued

Table 1. Continued

Study, Country	Model Name (If Applicable)	Study Population, Setting, Time Period	Outcome(s)	Sample Size	Outcome Rate	Modelling Methods	Artificial Intelligence Methods Used†	Validation Approach	Validated Model *c-statistic** (95% CI)
Wang et al., 2013 United States		Aged ≥18, Veterans Health Administration (VHA) enrolled patient, 2009 to 2010	(1) First hospitalization (90-day, all-cause) (2) First hospitalization (1-year, all-cause)	4598,408	90-day hospitalization = 2.7% 1-year hospitalization = 8.2%	Logistic regression (multinomial)	No	Split sample (60/40)	(1) *c* = 0.83 (0.83-0.83) (2) *c* = 0.81 (0.81-0.81)
Model Outcome: Mortality									
Simon et al., 2018 United States		Aged ≥13, mental health diagnosis at time of outpatient visit at primary care clinic, 2009 to 2015	(1) Suicide Mortality (90-day after mental health specialty visit) (2) Suicide Mortality (90-day after primary care visit)	19,961,059 visits (2960,929 patients)	Suicide mortality = 0.04%	Logistic regression	No	Split sample (65/35)	(1) *c* = 0.86 (0.85, 0.88), (2) *c* = 0.83 (0.81, 0.85)
Glanz et al., 2018 United States		Aged ≥18, ≥3 opioid prescription dates within 90 days, 2006 to 2014	Fatal pharmaceutical opioid and heroin overdoses	Derivation = 42,828 Validation = 10,708	Derivation = 0.03% Validation = 0.50%	Cox proportional hazards regression	No	Internal validation: Harrell bootstrap resampling External validation: Geographic validation	*c* = 0.75 (0.70-0.80)
DelPozo-Banos et al., 2018 United Kingdom		Aged ≥10, general practice ≥80% EMR data 5 years prior, residents of Wales, 2001 to 2015	Suicide mortality	54, 684	4.76%	Artificial neural networks	Yes	10 × 10 fold cross-validation	Mean Error Rate 26.78% (S.D. 1.46) Sensitivity = 64.6 Specificity = 81.9
Hippisley-Cox, 2017 United Kingdom		Aged 15 to 99, general practice patients with colorectal cancer, 1998 to 2014	(1) Death (all-cause) (2) Death (colorectal cancer)	Derivation= 44,145 Internal validation= 15,214 External validation= 437,821	(1) Death (all-cause) Derivation = 60.9% Validation = 30.6% (2) Death (colorectal cancer) Derivation = 61.8% Validation = 31.3%	Cox-hazard models	No	Sample split, random (75/25) External validation	(1) Women *c* = 0.778 (0.77-0.78), Men *c* = 0.76 (0.76-0.76) (2) Women *c* = 0.80 (0.78-0.81) Men *c* = 0.80 (0.79-0.81)

Continued

Table 1. Continued

Study, Country	Model Name (If Applicable)	Study Population, Setting, Time Period	Outcome(s)	Sample Size	Outcome Rate	Modelling Methods	Artificial Intelligence Methods Used†	Validation Approach	Validated Model c-statistic* (95% CI)
Barak-Corren et al., 2017 United States		Aged 10 to 90, inpatient and outpatient care, 1998 to 2012	Suicide Mortality	1724,885	Not reported	Bayesian classifier models	No	Simulated prospective approach	Female AUC = 0.77 (0.77-0.78) Male AUC = 0.76 (0.75-0.77)
Bloom et al., 2019 United Kingdom		Patients with a diagnosis of chronic, obstructive pulmonary disease (COPD), 2010 to 2015	Mortality (COPD, 12-month)	Derivation = 54,990 Validation = 4931	Derivation = 21% Validation = 29%	Cox regression models	No	Split sample (50/50), External validation	c = 0.67 (0.65-0.70)
Jung et al., 2019 United States		Aged 65 to 89, general practice, 2011 to 2014	Mortality (all-cause, 1-year)	349,667	2.1%	(1) Logistic regression (LR) (2) Gradient boosted trees (GBT)	Yes	Split sample (70/30) LR: 10-fold cross validation	LR AUC = 80.7% (No CIs) GBT AUC = 84.8% (no CIs)
Mathias et al., 2013 United States		Patients aged ≥50, 2003 to 2008	Mortality (5-year, all-cause)	7463	11%	Rotation forest	Yes	10-fold cross-validation	c = 0.86 (0.85-0.87)
Tierney et al., 1997 United States		Aged ≥14, general practice patients with reactive airway disease, 1992 to 1995	Mortality (3-year, all-cause)	1536 Derivation = 752 Validation = 784	12%	Logistic regression (multivariable)	No	Split sample (50/50)	c = 0.76 (no CIs)
O'Mahony et al., 2014 United Kingdom, Spain, Greece, Italy		Aged ≥16, European centers, patients with hypertrophic cardiomyopathy	Mortality (sudden cardiac death, 5-year)	Derivation = 3675 External validation = 1593	5%	Cox regression models (multivariable)	No	Bootstrapping 200 samples External validation: 1 health center	c = 0.67 (0.64-0.70)
Nijhawan et al., 2012 United States		HIV-positive patients admitted to hospital, 2006 to 2008	Mortality (30-day from index admission discharge)	1509 index hospital admissions (2476 patients)	3%	Logistic regression (multivariate)	No	Split sample (50/50), cross-validation	c = 0.79 (0.74-0.84)
Zeltzer, 2019 Israel	(1) Claims and EMR full (2) Claims, EMR and Admission full	Patients admitted to hospital overnight, 2016 to 2017	(2) Inpatient mortality (3) Mortality (12-month, all-cause)	144,966 index hospital admissions (118,510 patients)	Inpatient mortality = 2.6% Mortality (12-month) = 12.5%	Extreme gradient boosting	Yes	Split sample (83/17), 10-fold cross validation, bootstrapping	Inpatient mortality: (1) AUC = 0.91 (no CIs) (2) AUC = 0.95 (0.94-0.96) 1-year all-cause mortality: (1) AUC = 0.91 (0.92-0.93) (2) AUC = 0.92 (no CIs)

Continued

Table 1. Continued

Study, Country	Model Name (If Applicable)	Study Population, Setting, Time Period	Outcome(s)	Sample Size	Outcome Rate	Modelling Methods	Artificial Intelligence Methods Used†	Validation Approach	Validated Model*c-statistic** (95% CI)
Wang et al., 2013 United States		Aged ≥18, Veterans Health Administration (VHA) enrolled patient, 2009 to 2010	Death (without hospitalization, 90-day) Death (without hospitalization, 1-year)	4598,408	Death (90-day) = 0.7% Death (1-year) = 2.8%	Logistic regression (multinomial)	No	Split sample (60/40)	Death within 90-days c = 0.87 (0.86-0.87) Death within 1 year c = 0.85 (0.85-0.85)

†Artificial intelligence methods were defined as the use of a computer to simulate intelligent behavior and more specifically, virtual applications in which algorithms improve learning through experience, with minimal human intervention.[83]
Abbreviations: CI, confidence interval; AUC, area under the receiver operating characteristic curve.

several models considered and included sociodemographic variables such as race/ethnicity, socioeconomic status, access to care, marital status, and insurance payer.

Of the 31 included studies, 6 studies (19.4%) did not report the candidate predictor variables evaluated, and 7 studies (22.6%) did not report the relative predictive strength of individual variables. Of the studies which presented variable importance, notable overlap was found between variables frequently included in final models and those which held significant predictive value.

More specifically, age and specific medical diagnoses were identified as strong predictive contributors across all predictive outcomes, and less commonly, prior hospitalizations and laboratory results also demonstrated significant predictive power.

Two studies included free-text medical record data; 1 study used both structured and unstructured data fields for prediction,[46] whereas another solely processed free-text data and fit models using individual and strings of text.[63]

Predictive Accuracy of Included Studies

For studies with more than 1 prediction model, those with the highest c-statistic were considered the preferred/selected model. Table 1 presents an overview of model development and performance for each predictive outcome. Overall, almost all studies (n = 30, 97%) reported discrimination using c-statistics, of which the selected models ranged from 0.57 to 0.95 after validation. Of the best-performing models, the average c-statistics for ED visits, hospitalizations, and mortality were 0.73, 0.77, and 0.81, respectively. Twenty-four (77%) studies presented at least 1 measure of sensitivity, specificity, positive predictive value, or negative predictive value, often at various levels of predicted risk or probability thresholds.

Methodological Quality of Included Studies

Overall, the methodological quality of the included studies was poor. Although most studies included sufficient information on participants, many did not provide sufficient details on their specific analyses (Table 3). Calibration assesses the alignment between the probabilities predicted by the model and the actual observed probabilities of outcomes. In a well-calibrated model, predicted probabilities match the true probabilities of events, indicating that the model's predictions are not systematically

Table 2. Top 20 Predictor Variables Included and Considered in Models Predicting Emergency Department Visits, Hospitalizations, and Mortality

Study Outcomes	Emergency Department Visits (Model $n = 8$)		Hospital Admissions (Model $n = 27$)		Mortality (Model $n = 16$)	
	Number of models which included or excluded the respective variables in final models					
Predictor Variables	Included in Final Model	Excluded after Evaluation	Included in Final Model	Excluded after Evaluation	Included in Final Model	Excluded after Evaluation
Sociodemographic						
Age	7 (87.5%)	0 (0.0%)	23 (85.2%)	3 (11.1%)	15 (93.8%)	1 (6.25%)
Sex	5 (62.5%)	2 (25.0%)	22 (81.5%)	4 (14.8%)	13 (81.3%)	2 (12.5%)
Race/Ethnicity	3 (37.5%)	1 (12.5%)	15 (55.6%)	3 (11.1%)	6 (37.5%)	5 (31.3%)
Socioeconomic Status*	2 (25.0%)	1 (12.5%)	11 (40.7%)	5 (18.5%)	2 (12.5%)	6 (37.5%)
Marital Status	1 (12.5%)	0 (0.0%)	9 (33.3%)	1 (3.7%)	3 (18.8%)	2 (12.5%)
Insurance Payer	3 (37.5%)	0 (0.0%)	11 (40.7%)	1 (3.7%)	4 (25.0%)	2 (12.5%)
Access to Care*	0 (0.0%)	0 (0.0%)	9 (33.3%)	1 (3.7%)	0 (0.0%)	2 (12.5%)
Health Profile						
Smoking Status	1 (12.5%)	0 (0.0%)	7 (25.9%)	2 (7.4%)	3 (18.8%)	1 (6.3%)
BMI, Weight	2 (25.0%)	0 (0.0%)	8 (29.6%)	2 (7.4%)	5 (31.3%)	1 (6.3%)
Medical History						
Medical Diagnoses	8 (100%)	0 (0.0%)	21 (77.8%)	0 (0.0%)	14 (87.5%)	1 (6.3%)
Mental Illness	1 (12.5%)	0 (0.0%)	11 (40.7%)	2 (7.4%)	8 (50.0%)	2 (12.5%)
Substance Use	0 (0.0%)	0 (0.0%)	4 (14.8%)	3 (11.1%)	5 (31.3%)	4 (25.0%)
Medication Use	4 (50.0%)	0 (0.0%)	13 (48.1%)	1 (3.7%)	12 (75.0%)	2 (12.5%)
Clinical Findings						
Laboratory Tests	4 (50.0%)	0 (0.0%)	11 (40.7%)	0 (0.0%)	5 (31.3%)	0 (0.0%)
Laboratory Results	3 (37.5%)	0 (0.0%)	10 (37.0%)	0 (0.0%)	9 (56.3%)	0 (0.0%)
Vital Signs	1 (12.5%)	0 (0.0%)	8 (29.6%)	1 (3.7%)	4 (25.0%)	2 (12.5%)
Procedure History*	2 (25.0%)	1 (12.5%)	6 (22.2%)	1 (3.7%)	7 (43.8%)	1 (6.3%)
Health Care Utilization						
Prior Emergency Department Visits	4 (50.0%)	0 (0.0%)	18 (66.7%)	1 (3.7%)	7 (43.8%)	3 (18.8%)
Prior Inpatient Admissions	4 (50.0%)	0 (0.0%)	20 (74.1%)	1 (3.7%)	11 (68.8%)	1 (6.3%)
Emergency Admissions	0 (0.0%)	0 (0.0%)	12 (44.4%)	0 (0.0%)	4 (25.0%)	0 (0.0%)
Non-Urgent Admissions	0 (0.0%)	0 (0.0%)	4 (14.8%)	2 (7.4%)	4 (25.0%)	1 (6.3%)
No. of Inpatient Bed Days	4 (50.0%)	1 (12.5%)	8 (29.6%)	1 (3.7%)	2 (12.5%)	2 (12.5%)
Primary Care Visits	2 (25.0%)	0 (0.0%)	11 (40.7%)	4 (14.8%)	3 (18.8%)	6 (37.5%)
Outpatient Visits	4 (50.0%)	0 (0.0%)	6 (22.2%)	0 (0.0%)	2 (12.5%)	0 (0.0%)

*Variable inclusion examples: Socioeconomic Status: neighborhood income, individual income, deprivation index, zip code proxy measure; Access to Care: Health Region, proximity to health center, access to family doctor; Procedure History: surgical procedures, cardiovascular procedures.
Abbreviation: BMI, Body mass index.

too high or too low. Less than half of the studies reported calibration; the most frequently used methods of assessing calibration were calibration curve ($n = 7$), Hosmer-Lemeshow test ($n = 4$), less frequently, reported slope or raw predicted and observed values. Of these, 12 studies reported that models were well calibrated or stated that results indicated good calibration.

Discussion

The aims of this systematic review were threefold: 1) to examine candidate predictors which contribute

Table 3. Methodological Quality Assessment of Included Prediction Models following the Prediction Model Risk of Bias Assessment Tool (PROBAST) Guidelines

Study	Participants	Predictors	Outcome	Analysis	Overall
Frost DW, et al., 2017	−	−	−	?	−
Howell P, & Elkin PL., 2019	−	?	?	−	?
Pearce et al., 2019	−	+	−	+	+
Hu Z, et al., 2015	−	−	−	−	−
Hao et al., 2014	+	−	−	+	+
Bhavsar et al., 2018	−	−	?	+	+
Crane et al., 2010	−	−	+	+	+
Rahimian et al., 2018	−	−	−	−	−
Gao et al., 2014	−	?	+	?	+
Perkins et at., 2013	−	−	+	+	+
Morawski et al., 2020	−	+	−	+	+
Shadmi et al., 2015	−	?	+	+	+
Brisimi et al., 2018	−	+	+	?	+
Wallace et al., 2016	+	−	+	+	+
Chenore et al., 2013	−	−	−	+	+
Donnan et al., 2008	−	−	−	+	+
Watson et al., 2011	−	−	−	+	−
Hippisley-Cox et al., 2013	−	−	−	−	−
Nijhawan et al., 2012	−	+	−	+	+
Zeltzer et al., 2019	−	?	−	+	+
Wang et al., 2013	−	−	−	+	+
Simon et al., 2018	−	+	−	+	+
Glanz et al., 2018	−	−	−	?	+
DelPozo-Banos et al., 2018	−	−	−	+	+
Hippisley-Cox, 2017	−	−	−	−	−
Barak-Corren et al., 2017	−	−	−	+	+
Bloom et al., 2019	−	−	−	−	−
Jung et al., 2019	−	−	−	+	+
Mathias et al., 2013	+	−	−	?	+
Tierney et al., 1997	−	?	−	+	+
O'Mahony et al., 2014	−	+	−	+	+

Note: ROB, risk of bias.
− indicates low ROB; + indicates high ROB; and ? indicates unclear ROB.

to high-performing prediction models; 2) describe model performance; and 3) assess model suitability to contribute to clinical care and decision making. This systematic review included 31 studies with 36 unique prediction models which process primary care electronic medical record data. Sixteen (44.4%) of the models aimed to identify individuals at risk of hospitalization, 7 (19.4%) focused on the risk of ED visits, and 13 (36.1%) on the risk of mortality. Although a number of studies focused on specific populations such as patients with chronic kidney disease or HIV-positive patients admitted to hospital, most risk prediction models were developed in an adult general practice or community hospital population. Of the best-performing models within each study, more than 85% of models demonstrated adequate or good prediction accuracy. Across all prediction outcomes, models which showed poor predictive accuracy were often developed and validated using smaller sample sizes.

This review identified numerous variables widely found to be important predictors of emergency department visits, hospitalizations, and mortality, including age, prior health care utilization, medical diagnoses, and sex. In contrast to the current study, functional status, and activities of daily living,[70] as well as measures of multimorbidity and disease severity,[71] were common and significant predictors identified in other reviews of prediction models. This finding is noteworthy given the research which highlights the association between these variables and the outcomes of interest. For example, disease severity was found to have a significant association with emergency department use,[72] and the development of the well-known LACE index found Charlson Comorbidity score to be 1 of 4 variables independently associated with mortality or 30-day hospital readmission.[73] This could point toward a paucity of such patient information routinely collected and/or systematically stored within primary care EMR data.

Further, a growing amount of literature has supported the feasibility and benefits of processing EMR free-text and clinical notes for predictive modeling.[74] However, despite the amount of rich medical record data available in unstructured fields, only 2 studies[46,63] included in the current review utilized free-text data for prediction.[75,76] The absence of unstructured EMR data in prediction models may indicate a missed opportunity to process all available data in an effort to expand and improve predictive models within PHC.

Consistent with previous research,[77,8] more than half of the studies included at least 1 measure of patient sociodemographic data in final prediction models, including those which demonstrated high predictive performance. As primary care EMR data facilitates the use of routinely collected patient information valuable for prediction, this finding may reflect the increasing recognition of sociodemographic variables, such as socioeconomic status, as important patient data to be systematically collected across health systems.[78]

Previous research has explored the prediction of emergency department visits, hospitalization, or mortality, which process broad sources of data, including administrative, and self-report and survey data in addition to EMR data. Overall, the distribution of predictive ability reported in previous literature is consistent with that found across outcomes in the current review. A number of reviews have focused on the prediction of hospitalization and have reported *c*-statistics ranging from 0.60 to 0.83.[71] Furthermore, 2 reviews exploring the accuracy of mortality prediction tools reported similar discriminative ability, with *c*-statistics ranging from 0.56 to 0.85.[79] Although a portion of these models demonstrate adequate predictive accuracy, research highlights the importance of harnessing and building on robust model development and evaluation features, particularly due to the clinical implications of deficits in model performance.[80]

Many of the studies included in this review demonstrated poor reporting of predictive methodology and analyses. This is a common critique across multiple systematic reviews, which have also emphasized the need for improved reporting, external validation, and approaches to mitigate bias and the appropriate handling of missing data.[79,81] Thorough reporting of methods and results is particularly important for reproducibility, given the increasing amount of new models emerging in the field and the increasing potential for uptake of these models in clinical practice. Future research should adhere to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD)[82] statement to report the development and validation of predictive models. Only 4 studies discussed or considered clinical implementation in the study/model design. This is particularly important as there are many barriers to translation from a model on article to an effective predictive tool in clinical practice.

Strengths and Limitations

There are several strengths of the current review. There is a lack of systematic reviews which examine the use of primary care electronic medical record data in the prediction of emergency department visits, hospitalizations, and mortality. This review helped address critical gaps in the literature regarding the potential of primary care EMR data. Further, by bringing together a number of important prediction outcomes, this review provided connections between outcomes, and a comparison of insights. In addition, the use of a comprehensive search strategy has contributed to a robust, and extensive review spanning multiple countries and populations. However, there are limitations. First, the exclusion of models that predict multiple endpoint outcomes, such as models which predict hospitalization or mortality, may inadvertently overlook valuable insights that could enrich the existing literature. Second, the current search included only English studies, which may limit the scope of the available evidence. Lastly, although the findings in this review provide a comprehensive overview of the available research, inconsistent reporting of prediction model development and analysis may limit the ability to draw conclusions from findings presented here.

Conclusion

This review identified 31 studies which apply prediction methods to primary health care electronic medical record data to predict unplanned ED visits, hospital admissions, and mortality. Although demonstrating variable predictive performance, many models share commonalities in predictive variables of importance. Future research can build on these findings to develop and evaluate algorithms used to predict health service use, using EMR primary care EMR data. Such algorithms could be integrated into the workflow of primary care clinicians and clinics. For example, teams could set aside time at the beginning of the week to review a list of patients identified at high-risk, and identify actions (ie, outreach, booking a visit, organizing tests). Over time, such a system could be trained on data on the action taken (or lack of action) and perform better over time. Despite more work required to improve the quality and reporting of such models, the use of primary care EMR data for use in predictive analytics holds promise.

We are grateful to Melissa Perri, Rachelle Perron and Aakriti Pyakurel for assistance with reviewing the literature, and Robert Smith for input on this study. We appreciate helpful comments from Tyler Williamson and Jeremy Petch.

To see this article online, please go to: http://jabfm.org/content/37/4/583.full.

References

1. Haque M, Islam T, Rahman NAA, McKimm J, Abdullah A, Dhingra S. Strengthening primary health-care services to help prevent and control long-term (chronic) non-communicable diseases in low- and middle-income countries. RMHP 2020; Volume 13:409–26.
2. Canadian Health Services Research Foundation. What are the critical attributes of a high-quality primary healthcare system? Quality 2009;60.
3. Van Weel C, Kidd MR. Why strengthening primary health care is essential to achieving universal health coverage. Cmaj 2018;190:E463–E466.
4. WHO. *Primary health care measurement framework and indicators*. 2022.
5. Hernández-Peña P. *Global spending on health: weathering the storm 2020*. 2020.
6. Canadian Institute for Health Information (CIHI). Cost of a standard hospital stay. Your Health System.
7. Canadian Institute for Health Information. Hospital spending focus on the emergency department. 2020; 1–9.
8. Marchildon GD, Matteo L. *Bending the Cost Curve in Health Care: Canada's Provinces in International Perspective*. University of Toronto Press. 2015.
9. Devlin R, Brown A, Clerici C, et al. Hallway health care: a system under strain. Premier's council on improving healthcare and ending hallway medicine. Ontario. 2019;35.
10. Health Quality Ontario. *Under pressure: emergency department performance in Ontario*. 2016;28. Available at: https://www.hqontario.ca/portals/0/Documents/system-performance/under-pressure-report-en.pdf.
11. Pines JM, Hilton JA, Weber EJ, et al. International perspectives on emergency department crowding. Academic Emergency Medicine 2011;18:1358–70.
12. Wise A, MacIntosh E, Rajakulendran N, Khayat Z. *Transforming health: shifting from reactive to proactive and predictive care*. Available at: https://www.marsdd.com/news/transforming-health-shifting-from-reactive-to-proactive-and-predictive-care/. 2016.
13. de Oliveira C, Cheng J, Kurdyak P. Determining preventable acute care spending among high-cost patients in a single-payer public health care system. Eur J Health Econ 2019;20:869–78.
14. Rasmussen SR, Kilsmark J, Hvenegaard A, et al. Preventive health screenings and health consultations in primary care increase life expectancy without increasing costs. Scand J Public Health 2007;35:365–72.
15. Gruneir A, Bronskill SE, Maxwell CJ, et al. The association between multimorbidity and hospitalization is modified by individual demographics and physician continuity of care: A retrospective cohort study. BMC Health Serv Res 2016;16:1–9.
16. Yoon J, Cordasco KM, Chow A, Rubenstein LV. The relationship between same-day access and continuity in primary care and emergency department visits. PLoS One 2015;10:e0135274–12.
17. Bricard D, Or Z. Impact of early primary care follow-up after discharge on hospital readmissions. Eur J Health Econ 2019;611–623.
18. Jackson C, Shahsahebi M, Wedlake T, Dubard CA. Timeliness of outpatient follow-up: an evidence-based approach for planning after hospital discharge. Ann Fam Med 2015;13:115–22.
19. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. Milbank Q 1997;93:65–8.
20. WHO. *Operational framework for primary health Care*. Vol 2009. 2020.
21. Canadian Institute for Health Information. *How Canada compares: results from the Commonwealth Fund's 2019 International Health Policy Survey of Primary Care Physicians – Accessible Report*. 2020.

22. Canadian Institute for Health Information. *How Canada compares: results from the Commonwealth Fund's 2019 International Health Policy Survey of Primary Care Physicians.* 2020.
23. CMWF. *Commonwealth Fund International Health Policy Survey of Primary Care Physicians.* 2009.
24. Birtwhistle R, Williamson T. Primary care electronic medical records: a new data source for research in Canada. Cmaj 2015;187:239–40.
25. Brule S, Mcdonald H. Health fact sheets - primary health care providers, 2019. Statistics Canada 2019; Catalogue.
26. Naylor CD. On the prospects for a (deep) learning health care system. JAMA 2018;320:1099–100.
27. Hinton G. Deep learning—a technology with the potential to transform health care. JAMA 2018;320: 1101–2.
28. Wallace E, McDowell R, Bennett K, Fahey T, Smith SM. External validation of the probability of repeated admission (Pra) risk prediction tool in older community-dwelling people attending general practice: a prospective cohort study. BMJ Open 2016;6:e012336–8.
29. Klunder JH, Panneman SL, Wallace E, et al. Prediction models for the prediction of unplanned hospital admissions in community-dwelling older adults: a systematic review. PLoS One 2022;17: e0275116.
30. Kim P, Daly JM, Berry-Stoelzle MA, et al. Prognostic indices for advance care planning in primary care: a scoping review. J Am Board Fam Med 2020;33:322–38.
31. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. JAMA 2011;306:1688–98.
32. Mahmoudi E, Kamdar N, Kim N, Gonzales G, Singh K, Waljee AK. Use of electronic medical records in development and validation of risk prediction models of hospital readmission: systematic review. The BMJ 2020;369:1–10.
33. Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: a systematic review. Journal of the American Medical Informatics Association 2020;27:1764–73.
34. O'Caoimh R, Cornally N, Weathers E, et al. Risk prediction in the community: a systematic review of case-finding instruments that predict adverse healthcare outcomes in community-dwelling older adults. Maturitas 2015;82:3–21.
35. Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014;11:e1001744
36. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–38.
37. Wolff RF, Moons KGM, Riley RD, for the PROBAST Group†, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;170:51–8.
38. Howell P, Elkin PL. Can solo practitioners survive in value-based healthcare? Validating a predicative model for ED utilization. Stud Health Technol Inform 2019;264:1682–3.
39. Hu Z, Jin B, Shin AY, et al. Real-time web-based assessment of total population risk of future emergency department utilization: statewide prospective active case finding study. Interact J Med Res 2015;4:e2.
40. Hao S, Jin B, Shin AY, et al. Risk prediction of emergency department revisit 30 days post discharge: a prospective study. PLoS One 2014;9: e112944.
41. Bhavsar NA, Gao A, Phelan M, Pagidipati NJ, Goldstein BA. Value of neighborhood socioeconomic status in predicting risk of outcomes in studies that use electronic health record data. JAMA Netw Open 2018;1:e182716.
42. Crane SJ, Tung EE, Hanson GJ, Cha S, Chaudhry R, Takahashi PY. Use of an electronic administrative database to identify older community dwelling adults at high-risk for hospitalization or emergency department visits: the elders risk assessment index. BMC Health Serv Res 2010;10:338.
43. Gao J, Moran E, Li YF, Almenoff PL. Predicting potentially avoidable hospitalizations. Med Care 2014;52:164–71.
44. Perkins RM, Rahman A, Bucaloiu ID, et al. Readmission after hospitalization for heart failure among patients with chronic kidney disease: a prediction model. CN 2013;80:433–40.
45. Morawski K, Dvorkis Y, Monsen CB. Predicting hospitalizations from electronic health record data. American Journal of Managed Care 2020;26:1–7.
46. Watson AJ, O'Rourke J, Jethwani K, et al. Linking electronic health record-extracted psychosocial data in real-time to risk of readmission for heart failure. Psychosomatics 2011;52:319–27.
47. Nijhawan AE, Clark C, Kaplan R, Moore B, Halm EA, Amarasingham R. An electronic medical record-based model to predict 30-day risk of readmission and death among HIV-infected inpatients. J Acquir Immune Defic Syndr (1988)2012;61:349–58.
48. Wang L, Porter B, Maynard C, et al. Predicting risk of hospitalization or death among patients receiving primary care in the Veterans Health Administration. Med Care 2013;51:368–73.
49. Simon GE, Johnson E, Lawrence JM, et al. Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records. American Journal of Psychiatry 2018;175:951–60.
50. Glanz JM, Narwaney KJ, Mueller SR, et al. Prediction model for two-year risk of opioid overdose

among patients prescribed chronic opioid therapy. J Gen Intern Med 2018;33:1646–53.

51. Barak-Corren Y, Castro VM, Javitt S, et al. Predicting suicidal behavior from longitudinal electronic health records. American Journal of Psychiatry 2017;174:154–62.
52. Jung K, Sudat SEK, Kwon N, Stewart WF, Shah NH. Predicting need for advanced illness or palliative care in a primary care population using electronic health record data. J Biomed Inform 2019; 92:103115.
53. Tierney WM, Murray MD, Gaskins DL, Zhou XH. Using computer-based medical records to predict mortality risk for inner-city patients with reactive airways disease. Journal of the American Medical Informatics Association 1997;4:313–21.
54. Mathias JS, Agrawal A, Feinglass J, Cooper AJ, Baker DW, Choudhary A. Development of a 5 year life expectancy index in older adults using predictive mining of electronic health record data. Journal of the American Medical Informatics Association 2013;20:e118–e124.
55. Brisimi TS, Xu T, Wang T, Dai W, Adams WG, Paschalidis IC. Predicting chronic disease hospitalizations from electronic health records: an interpretable classification approach. Proceedings of the IEEE 2018;106:690–707.
56. Chenore T, Pereira Gray DJ, Forrer J, Wright C, Evans PH. Emergency hospital admissions for the elderly: insights from the Devon Predictive Model. J Public Health (Bangkok) 2013;35:616–23.
57. Donnan PT, Dorward DWT, Mutch B, Morris AD. Development and validation of a model for predicting emergency admissions over the next year (PEONY): a UK historical cohort study. Arch Intern Med 2008;168:1416–22.
58. Bloom CI, Ricciardi F, Smeeth L, Stone P, Quint JK. Predicting COPD 1-year mortality using prognostic predictors routinely measured in primary care. BMC Med 2019;17:1–10.
59. Hippisley-Cox J, Coupland C. Development and validation of risk prediction equations to estimate survival in patients with colorectal cancer: cohort study. BMJ (Online) 2017;357:1–25.
60. DelPozo-Banos M, John A, Petkov N, et al. Using neural networks with routine health records to identify suicide risk: feasibility study. JMIR Ment Health 2018;5:e10144.
61. Hippisley-Cox J, Coupland C. Predicting risk of emergency admission to hospital using primary care data: derivation and validation of QAdmissions score. BMJ Open 2013;3:e003482–16.
62. Rahimian F, Salimi-Khorshidi G, Payberah AH, et al. Predicting the risk of emergency admission with machine learning: development and validation using linked electronic health records. PLoS Med 2018;15:e1002695.
63. Frost DW, Vembu S, Wang J, Tu K, Morris Q, Abrams HB. Using the electronic medical record to identify patients at high risk for frequent emergency department visits and high system costs. American Journal of Medicine 2017;130:601.e17-601–e22.
64. Wallace E, McDowell R, Bennett K, Fahey T, Smith SM. External validation of the Vulnerable Elder's Survey for predicting mortality and emergency admission in older community-dwelling people: a prospective cohort study. BMC Geriatr 2017;17:1–8.
65. Pearce C, McLeod A, Rinehart N, et al. POLAR diversion: using general practice data to calculate risk of emergency department presentation at the time of consultation. Appl Clin Inform 2019;10:151–7.
66. Shadmi E, Flaks-Manov N, Hoshen M, Goldman O, Bitterman H, Balicer RD. Predicting 30-day readmissions with preadmission. Med Care 2015; 53:283–9.
67. Zeltzer D, Balicer RD, Tzvi S, Flaks-Manov N, Einav L, Shadmi E. Prediction accuracy with electronic medical records versus administrative claims. 2019;57:551–9.
68. O'Mahony C, Jichi F, Pavlou M, for the Hypertrophic Cardiomyopathy Outcomes Investigators, et al. A novel clinical risk prediction model for sudden cardiac death in hypertrophic cardiomyopathy (HCM Risk-SCD). Eur Heart J. 2014;35:2010–20.
69. Chenore T, Pereira Gray DJ, Forrer J, Wright C, Evans PH. Emergency hospital admissions for the elderly: insights from the Devon Predictive Model. Journal of Public Health (United Kingdom)2013;35: 616–23.
70. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. JAMA - JAMA 2011;306:1688–98.
71. Wallace E, Stuart E, Vaughan N, Bennett K, Fahey T, Smith SM. Risk prediction models to predict emergency hospital admission in community-dwelling adults: a systematic review. Med Care 2014;52: 751–65.
72. Sun BC, Burstin HR, Brennan TA. Predictors and Outcomes of frequent emergency department users. Academic Emergency Medicine 2003;10:320–8.
73. Van Walraven C, Dhalla IA, Bell C, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. CMAJ Canadian Medical Association Journal 2010;182:551–7.
74. Seinen TM, Fridgeirsson EA, Ioannou S, et al. Use of unstructured text in prognostic clinical prediction models: a systematic review. Journal of the American Medical Informatics Association 2022;29:1292–302.
75. Chrusciel J, Girardon F, Roquette L, Laplanche D, Duclos A, Sanchez S. The prediction of hospital

length of stay using unstructured data. BMC Med Inform Deci's Mak 2021;21:1–9.

76. Li I, Pan J, Goldwasser J, et al. Neural Natural Language Processing for Unstructured Data in Electronic Health Records: A Review. Vol 37. Association for Computing Machinery; 2021. arXiv:2107.02975.
77. Zhang Y, Zhang Y, Shelle E, et al. Assessing the impact of social determinants of health on predictive models for potentially avoidable 30-day readmission or death. PLoS One 2020;15:e0235064–15.
78. Pinto AD, Glattstein-Young G, Mohamed A, Bloch G, Leung FH, Glazier RH. Building a foundation to reduce health inequities: routine collection of sociodemographic data in primary care. J Am Board Fam Med 2016;29:348–55.
79. Kim P, Daly JM, Berry-Stoelzle MA, et al. Prognostic indices for advance care planning in primary care: a scoping review. J Am Board Fam Med 2020;33: 322–38.
80. Rajkumar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. NPJ Digit Med 2018;1:1–10.
81. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. Journal of the American Medical Informatics Association 2017; 24:198–208.
82. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable Prediction model for Individual Prognosis Or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1–W73.
83. Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism 2017;69:S36–S40.

Appendix A

Sample Search Strategy

Database: Ovid MEDLINE: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE® Daily and Ovid MEDLINE® <1946-Present>

Search Strategy:

----

1 Primary Health Care/ (70433)
2 Comprehensive Health Care/ (6432)
3 exp General Practice/ (72688)
4 Community Health Centers/ (6765)
5 Community Health Services/ (30311)
6 Child Health Services/ (19792)
7 Community Mental Health Services/ (17984)
8 exp Maternal Health Services/ (44875)
9 exp Community Mental Health Centers/ (3143)
10 Maternal-Child Health Centers/ (2292)
11 Physicians, Family/ (15929)
12 General Practitioners/ (6632)
13 Physicians, Primary Care/ (2796)
14 Pediatricians/ (499)
15 pediatrics/ or neonatology/ or perinatology/ (54422)
16 Ambulatory Care Facilities/ (17443)
17 Outpatient Clinics, Hospital/ (15342)
18 Outpatients/ (14149)
19 Preventive Medicine/ (11355)
20 primary care.tw,kf. (103761)
21 primary healthcare.tw,kf. (4884)
22 primary health care.tw,kf. (24402)
23 Family practice*.tw,kf. (8313)
24 general practice*.tw,kf. (40180)
25 family medicine.tw,kf. (9656)
26 general practitioner*.tw,kf. (46993)
27 family physician*.tw,kf. (13841)
28 family doctor*.tw,kf. (4415)
29 Community Health Centre*.tw,kf. (809)
30 Community Health Center*.tw,kf. (2751)
31 Community Healthcare.tw,kf. (615)
32 Community Health Care.tw,kf. (1067)
33 Community Health service*.tw,kf. (1600)
34 ((outpatient* or ambulatory or community) adj4 (clinic or clinics or healthcare or health care or centre* or centers)).tw,kf. (60834)
35 patient-centered medical home*.tw,kf. (1545)
36 patient-centred medical home*.tw,kf. (34)

37 Family Health Team*.tw,kf. (325)
38 or/1-37 (510304)
39 exp Medical Records Systems, Computerized/ (36267)
40 medical record linkage/ (4476)
41 Databases, Factual/ (69069)
42 Datasets as Topic/ (2472)
43 exp Medical Informatics/ (436809)
44 electronic medical record*.tw,kf. (13195)
45 electronic health record*.tw,kf. (12183)
46 computerized medical record*.tw,kf. (671)
47 computerised medical record*.tw,kf. (114)
48 computerized health record*.tw,kf. (21)
49 computerised health record*.tw,kf. (6)
50 automated medical record*.tw,kf. (90)
51 automated health record*.tw,kf. (2)
52 health information technolog*.tw,kf. (2993)
53 administrative data.tw,kf. (7354)
54 medical informatics.tw,kf. (2750)
55 or/39-54 (469117)
56 Data Interpretation, Statistical/ (54490)
57 exp Artificial Intelligence/ (80268)
58 exp Data Mining/ (7134)
59 big data/ (64)
60 deep learning/ (78)
61 Data Science/ (10)
62 Data Analysis/ (89360)
63 Pattern Recognition, Automated/ (24171)
64 artificial intelligence.tw,kf. (3790)
65 Machine learning.tw,kf. (17867)
66 neural network*.tw,kf. (37993)
67 data mining.tw,kf. (8120)
68 data science.tw,kf. (535)
69 data sciences.tw,kf. (77)
70 Data Analytic.tw,kf. (432)
71 Data Analytics.tw,kf. (688)
72 text mining.tw,kf. (2092)
73 deep learning.tw,kf. (3324)
74 supervised learning.tw,kf. (1966)
75 unsupervised learning.tw,kf. (990)
76 big data.tw,kf. (4736)
77 deep architecture*.tw,kf. (101)
78 pattern recognition.tw,kf. (15417)
79 exp risk/ (1104210)
80 Forecasting/ (81083)

81 Prognosis/ (462954)
82 Probability/ (54316)
83 algorithms/ (232345)
84 models, statistical/ (85771)
85 exp Regression Analysis/ (394966)
86 predict*.tw,kf. (1400959)
87 prognostic*.tw,kf. (268647)
88 Probabili*.tw,kf. (200461)
89 forecast*.tw,kf. (15240)
90 (risk or risks).tw,kf. (1963244)
91 algorithm*.tw,kf. (218625)
92 statistical model*.tw,kf. (14542)
93 regression.tw,kf. (639902)
94 or/56-93 (4717078)
95 hospitalization/ or patient admission/ or patient readmission/ (128838)
96 admission*.tw,kf. (194751)
97 readmission*.tw,kf. (22237)
98 re-admission*.tw,kf. (1774)
99 hospital visits.tw,kf. (1455)
100 hospitalization*.tw,kf. (125145)
101 hospitalisation*.tw,kf. (15565)
102 Emergency Service, Hospital/ (60724)
103 (emergency adj4 (visit or visits or department* or admission* or service*)).tw,kf. (102397)
104 Emergency Treatment/ (10170)
105 death/ or death, sudden/ (28231)
106 mortality/ or hospital mortality/ or survival rate/ (228682)
107 Survival/ (4550)
108 survival analysis/ (124690)
109 (death or deaths or died or survival or mortality).tw,kf. (2021652)
110 or/95-109 (2443144)
111 38 and 55 and 94 and 110 (2056)

Appendix B

Systematic Review Data Extraction Fields, Guided by the CHARMS Checklist

Study Information	Study Reference Number
Setting and participants (see CHARMS p. 4-5)	Country
	Primary care setting (ex. outpatient cardiology or general practice)
	Participant description (ex. patients with cystic fibrosis or all patients of included general practices)
	Intervention, if any (please enter 'NA' if no action/intervention)
	Success of intervention (please enter 'NA' if no action/intervention)
Data sources (see CHARMS p. 3-5)	Study design (i.e., Retrospective or prospective, nested case-control or cohort)
	Study objective
	Data start date (month/year)
	Data end date (month/year)
	EMR developer
	EMR Version
	Other sources of data (i.e., Insurance claims, administrative database,
Outcomes (see CHARMS p. 4, p.5-6)	Outcomes predicted
	Definition and method of measurement of outcomes (ex. single or combined endpoints; all-cause mortality vs. mortality due to myocardial infarction)
	Was the same outcome definition and measurement method used for all patients? (Y/N)
Candidate predictors (see CHARMS p. 4, p.6-7)	Number and type of predictors (ex. demographics, patient history, etc.)
	Definition and method for measurement of candidate predictors (ex. diabetes can be measured by oral glucose tolerance test, HbA1c measurement, fasting plasma glucose, or self-report)
	Were predictors assessed blinded for outcome, and for each other? (in the case where studies examined the added predictive utility of certain variables)
	Handling of predictors in model development (ex. continuous, linear, non-linear transformations or categorized)
Sample size (see CHARMS p. 4, p. 7)	Number of participants and number of outcomes/events
	Number of outcomes/events in relation to the number of candidate predictors (Events Per Variable)
	Number of participants with any missing value
	Number of participants with missing data for each predictor

Missing data (see CHARMS p. 4, p. 7)	Handling of missing data (ex. complete-case analysis, imputation, or other methods)
Model development (see CHARMS p. 4 and p.7-8)	Modelling methods used (ex. logistic regression, survival, neural networks, machine learning techniques)
	AI used? (Y/N)
	Modelling assumptions made
	Method for selection of predictors for inclusion in multivariable modelling
	Method for selection of predictors during multivariable modelling
	Shrinkage of predictor weights or regression coefficients
Results, including model performance and evaluation (see CHARMS p. 4 and p.8-9)	Variables included in best final models (ex. basic, extended, simplified) presented, including predictor weights or regression coefficients
	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) AND Discrimination (C-statistic, D-statistic, log-rank) measures with confidence intervals
	Sensitivity (please enter 'NR' if not reported)
	Specificity (please enter 'NR' if not reported)
	Predictive values (please enter 'NR' if not reported)
	F scores (please enter 'NR' if not reported)
	Receiver Operating Characteristic (ROC) (please enter 'NR' if not reported)
	Area Under ROC (AUC) (please enter 'NR' if not reported)
	Method used for testing model performance: development dataset only (random split of data, resampling methods, ex. bootstrap or cross-validation, none) OR separate external validation (ex. temporal, geographical, different setting, different investigators)
	Comparison of the distribution of predictors (including missing data) for development and validation datasets)
	In case of poor validation, whether model was adjusted or updated (e.g. intercept recalibrated, predictor effects adjusted, or new predictors added)
	Any alternative presentation of the final prediction models, e.g. sum score, nomogram, score chart, predictions for specific risk subgroups, with performance metrics)
Interpretation and discussion	Interpretation of presented models (confirmatory versus exploratory)
	Comparison with other studies, discussion of generalizability, strengths, and limitations

Abbreviations: CHARMS checklist, CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies.