Supplement

Machine Learning Models Predicting Inpatient Falls

Authors:
Hojjat Salehinejad, Ricky Rojas, Kingsley Iheasirim, Paul Takahashi, Bijan J. Borah, Mohammed Yousufuddin

Supplemental Content

- **Supplement Table 1.** Comparison of the study (or Minnesota) population with the U.S. national average for demographics, social determinants of health, and key comorbidities
- **Supplement Table 2.** ICD-10-CM codes used to define chronic disease conditions
- **Supplement Figure 1.** Workflow illustrating integration of data into model training and update processes
- **Supplement Figure 2.** Workflow for machine learning model training and evaluation
- **Supplement Figure 3.** Processing strategy for static and dynamic input variables
- **Supplement Figure 4.** System architecture for fall risk prediction
- **Supplement Table 3.** Hester-Davis Scale: Fall risk assessment tool and scoring criteria
- **Supplement Figure 5.** Calibration plots for machine learning models
- **Supplement Table 4.** Brier scores for evaluating fall risk machine learning model performance
- **Supplement Table 5.** TRIPOD checklist: Reporting standards for predictive modeling studies

Supplement Table 1. Comparison of the study (or Minnesota) population with the U.S. national average for demographics, social determinants of health, and key comorbidities

Characteristics	Study population/Minnesota	US National average
Age, years, median (IQR)	66	58[1]/57, [2]
Nonwhite %	8.2[3]/ 14[4]/6[5]	18.4[6]
Uninsured %	3.8[7]	8.0[8]
Median household income	$87,556[9]	$80,610[10]
Holders of bachelor's degree, %	40[11]	37.9[12]
Comorbidities, %		
Hypertension	52.7	40.2
Diabetes	22.5	16.6
Heart failure	15.3	3.5
Coronary artery disease	16.3	9.0
Cancer	15.7	10.1
Chronic kidney disease	18.5	5.5
Stroke	3.2	3.1
Liver disease		2.6

1. Rollings, K.A., Kunnath, N., Ryus, C.R., Janke, A.T. & Ibrahim, A.M. Association of Coded Housing Instability and Hospitalization in the US. *JAMA Network Open* **5**, e2241951-e2241951 (2022).
2. Hirode, G., Saab, S. & Wong, R.J. Trends in the Burden of Chronic Liver Disease Among Hospitalized US Adults. *JAMA Network Open* **3**, e201997-e201997 (2020).
3. Rocca, W.A.*, et al.* Data Resource Profile: Expansion of the Rochester Epidemiology Project medical records-linkage system (E-REP). *Int J Epidemiol* **47**, 368-368j (2018).
4. Yousufuddin, M.*, et al.* Rural-urban differences in Long-term Mortality and Readmission Following COVID-19 Hospitalization, 2020 to 2023. *Open Forum Infectious Diseases* (2024).
5. Yousufuddin, M.*, et al.* Systolic blood pressure, a predictor of mortality and life expectancy following heart failure hospitalization, 2010-2023. *Eur J Intern Med* **131**, 71-82 (2025).
6. Trivedi, A.N.*, et al.* Quality and equity of care in U.S. hospitals. *N Engl J Med* **371**, 2298-2308 (2014).
7. Health, M.D.o. Minnesota's uninsured rate hits all-time low despite coverage lapses, affordability concerns. (2024).
8. Bureau, U.s.c. Health Insurance Coverage in the United States: 2023. (2024).
9. zipcode, I.b. Minnesota income statistics for 2024. (2024).
10. Bureau, U.s.c. Income in the United states: 2023. (2023).
11. FRED. Bachelor's Degree or Higher for Minnesota (GCT1502MN). (2024).
12. Center, P.R. 10 facts about today's college graduates. (2022).

Supplement Table 2. ICD-10-CM codes for chronic disease conditions

Category	Description of individual chronic condition	ICD-10-CM codes
Cardiovascular	Hypertension	I10, I11.9, I15.0, I15.9
	Coronary artery disease	I20.0, I21.4, I21.AI, I25.42, I25.82, I25.110, I25.111, I125.118, I25.119, I25.710, Z95.1, Z95.5, Z98.61, Z98.61
	Atrial fibrillation or flutter	I48.0, I48.1, I48.11, I48.19, I48.2, I48.20, I48.3, I48.4, I48.91, I48.91
	Stroke	I61.11, I60.12, I60.31, I60.32, I60.4, I60.51, I60.6, I60.7, I60.8, I60.9, I61.1, I61.2, I61.5, I61.6, I61.8, I61.9, I62.00I63.00, I63.011, I63.012, I63.019, I63.02, I63.22, I63.233, I63.30, I63.031, I63.032, I63.10, I63.311, I63.20, I63.212, I63.219, I63.239, I63.321, I63.232, I63.312, I63.321, I63.322, I63.332, I63.331, I63.333, I63.39, I63.341, I63.342, I63.343, I63.413, I63.419, I63.429, I63.431, I63.432, I63.439, I63.40, I63.511, I63.512, I63.513, I63.521, I63.522, I63.523, I63.532, I63.541, I63.542, I63.543, I63.50, I63.59, I63.8, I63.81, I63.89, I63.9, I69.054, I69.122, I69.151, I69.154, I69.198, I69.222, I69.251, I69.254, I69.328, I69.331, I69.332, I69.331, I69.334, I69.352, I69.354, I69.359, I69.392, I69.398, I69.822, I69.851, I69.852, I69.854, I69.922, I69.928, I69.931, I69.934, I69.939, I69.951, I69.954, I69.959, I69.992, G81.90, G81.91, G81.92, G81.93, G81.94
Respiratory	COPD	J44.0, J44.1, J44.9
	OSA	G47.33
Neurologic	Dementia	F02.80, F02.81, F0.90, F0.91, F01.50, F0.51,G30.0, G30.1, G30.9, G31.09, G31.83,
Endocrine and Metabolism	Hyperlipidemia	E78.00, E78.2, E78.5
	Diabetes mellitus (Type 1 or Type 2)	E08.9, E08.10, E08.11, E08.21, E08.22, E08.40, E08.42, E08.52, E08.65, E08.649, E08.620, E08.3513, E08.3593, E10.9, E10.10, E10.11,

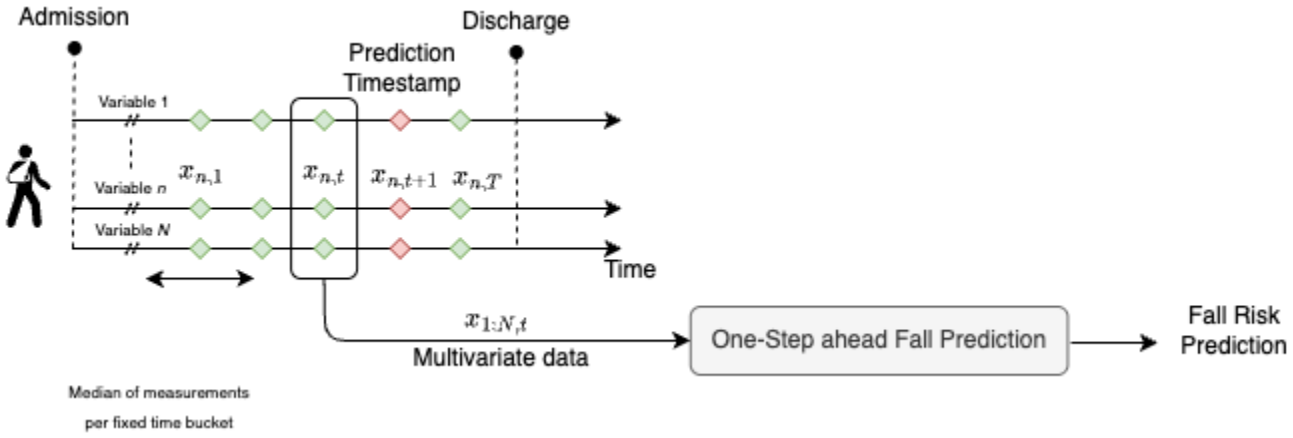
		E10.21, E10.22, E10.36, E10.40, E10.40, E10.43, E10.51, E10.52, E10.65, E10.69,E10.319, E10.621, E10.641, E10.649,E10.3513, E10.3593,
	Obesity	E66.01, E66.09, E66.09, E66.2, E66.8, E66.9, Z68.30, Z68.31, Z68.32, Z68.33, Z68.34, Z68.35, Z68.36, Z68.37, Z68.38, Z68.39, Z68.40, Z68.41, Z68.42, Z68.43, Z68.44, Z68.45
Kidney	Chronic kidney disease	N18.3, N18.31, N18.32, N18.4, N18.5, N18.6, Z99.2,
Hematology	Anemia	D50.0, D52.9, D63.0, D63.1, D63.8, D64.81, D64.9,
Oncology	All malignant conditions	C01, C02.1, C02.3, C02.9, C03.9, C06.1, C06.9, C07, C7A.00, C7A.010, C7A.012, C7A.019, C7A.020, C7A.025, C7A.090, C7A.092, C7A.098, C10.9, C11.9, C14.0, C15.4, C15.5. C15.8, C15.9, C16.0, C16.1, C16.2, C16.5, C16.6, C16.9, C17.2, C18.0, C18.1, C18.2, C18.3, C18.4, C18.5, C18.6, C18.7, C18.8, C18.9, C19, C20, C21.0, C21.1, C21.8, C22.0, C22.1, C22.9, C23, C24.0, C24.1, C24.9, C25.0, C25.1, C25.2, C25.4, C25.7, C25.9, C26.1, C26.9, C30.0, C32.1, C32.9, C34.01, C34.02, C34.10, C34.11, C34.12, C34.2, C34.30, C34.31, C34.32, C32.8, C34.80, C34.81, C34.82, C34.90, C34.91, C34.92, C37, C38.3, C41.1, C41.2, C41.4, C41.9, C47.9, C48.0, C48.2, C49.0, C49.21, C49.22, C49.3, C49.4, C49.8, C49.9, C50.012, C50.111, C50.112, C50.119,C50.211, C50.212, C50.312, C50.411, C50.412, C50.512, C50.811, C50.812, C50.819, C50.911, C50.912, C50.919, C51.0, C51.9, C53.0, C53.8, C53.9, C54.1, C55, C56.1, C56.2, C56.3, C56.9, C57.01, C57.7, C57.8, C57.9, C60.9, C61, C62.11, C64.1, C64.2, C64.9, C65.9, C66.1, C66.2, C66.9, C67.0, C67.1, C67.2, C67.4, C67.6, C67.8, C67.9, C68.8, C69.2, C69.21, C69.22, C69.31, C69.32, C70.1,C71.2, C71.3, C71.1, C71.6,

		C71.8, C71.9, C72.0, C73, C74.91, C74.92, C74.99, C75.9, C76.0, C76.1, C76.3, C76.42, C76.52, C76.8, C81.00, C81.02, C81.08, C81.10, C81.12, C81.18, C81.19, C81.90, C82.00, C82.02, C82.03, C82.04, C82.08, C82.09, C82.10, C82.11, C82.13, C82.15, C82.16, C82.17, C82.18, C82.19, C82.20, C82.21, C82.28, C82.29, C82.31, C82.38, C82.39, C82.40, C82.48, C82.49, C82.58, C82.80, C82.88, C82.89, C82.90, C82.93, C82.98, C82.99, C83.00, C83.01, C83.02, C83.03, C83.06, C83.07, C83.09, C83.10, C83.11, C83.12, C83.13, C83.18, C83.19, C83.30, C83.31, C83.32, C83.33, C83.34, C83.36, C83.37, C83.38, C83.39, C83.70, C83.73, C83.78, C83.79, C84.40, C84.48, C84.49, C84.60, C84.70, C84.71, C84.73, C84.74, C84.78, C84.79, C85.20, C85.22, C85.28, C85.29, C85.80, C85.82, C85.83, C85.88, C85.89, C85.90, C85.91, C85.92, C85.93, C85.94, C85.96, C85.98, C85.99, C86.0, C88.0, C88.4, C90.00, C90.02, C91.10, C91.11, C91.12, C91.40, C91.41, C91.Z0, C92.10, C92.11, C92.12, C92.50, C92.51, C92.90, C93.10, C93.12, C94.6, C95.00, C95.01, C95.02, C95.10, C95.90, D05.11, D45, D75.81, D47.1, R18.0
Psychiatric conditions	Depression (major depression and bipolar disorder)	F30.2, F30.3, F30.4, F30.9, F30.10, F30.12, F31.2, F31.4, F31.5, F31.9, F31.10, F31.12, F31.13, F31.30, F31.31, F31.32, F31.60, F31.62, F31.63, F31.64, F31.70, F31.71, F31.73, F31.74, F31.75 F31.76, F31.77, F31.81, F31.89, F32.1, F32.2, F33.0, F33.0, F33.1, F33.2, F33.3, F33.9, F33.40, F33.41, F33.42,

Supplementary Table 2. Hester-Davis Scale: Fall risk assessment and scoring system

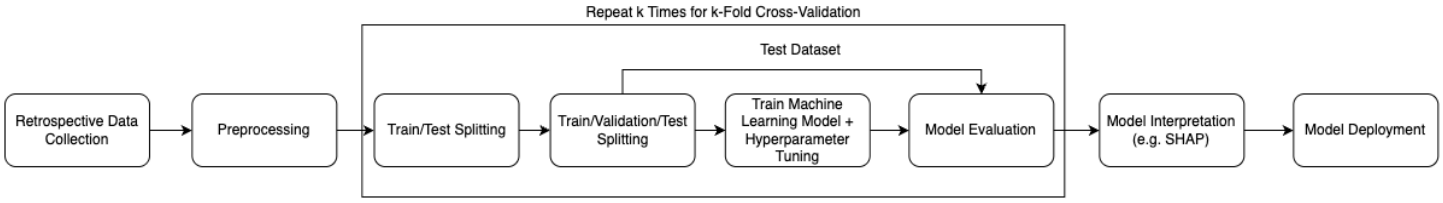
Categories	Event	Score
Last known Fall	No fall or unknown	0
	Within the last year	1
	Within the 6 months	2
	Within the month	3
	During the current hospitalization	4
Mobility	No limitations	0
	Dizziness or generalized weakness	1
	Immobilized or requires assist by 1 person	2
	Use of assistive device or require assistance of 2 persons	3
	Hemiplegic, paraplegic, quadriplegic	4
Medications	No medications	0
	Cardiovascular or central nervous system medication	1
	Cardiovascular and central nervous system medication	2
	Diuretics	3
	Chemotherapy in the last month	4
Mental status/level of consciousness/ awareness	Awake, alert, oriented to date, place, and person	0
	Oriented to person and place	1
	Lethargic and oriented to person only	2
	Memory loss or confusion and requires reorientation	3
	Unresponsive or non-compliance with instruction	4
Toileting needs	No needs	0
	Need for catheter or diversion device	1
	Use of assistive device (bedside commode, pan, or urinal)	2
	Incontinence	3
	Diarrhea, frequency, or urgency	4
Volume/ electrolytes	No problem	0
	Nothing by mouth greater than 24 hours	1
	Use of intravenous fluids or tube feeds	2
	Nausea/vomiting	3
	Low blood glucose or electrolyte abnormality	4
Communication /sensory	No deficit	0
	Visual (glasses) or hearing deficit	1
	Non-English patient, unable to speak, or slurred speech	2
	Neuropathy	3
	Blindness or recent visual changes	4
Behavior	Appropriate behavior	0
	Depression or anxiety	1
	Behavioral non-compliance with instruction	2
	Ethanol or substance use	3
	Impulsiveness	4

Supplement Figure 1. Workflow depicting how data are incorporated into model training and updates.



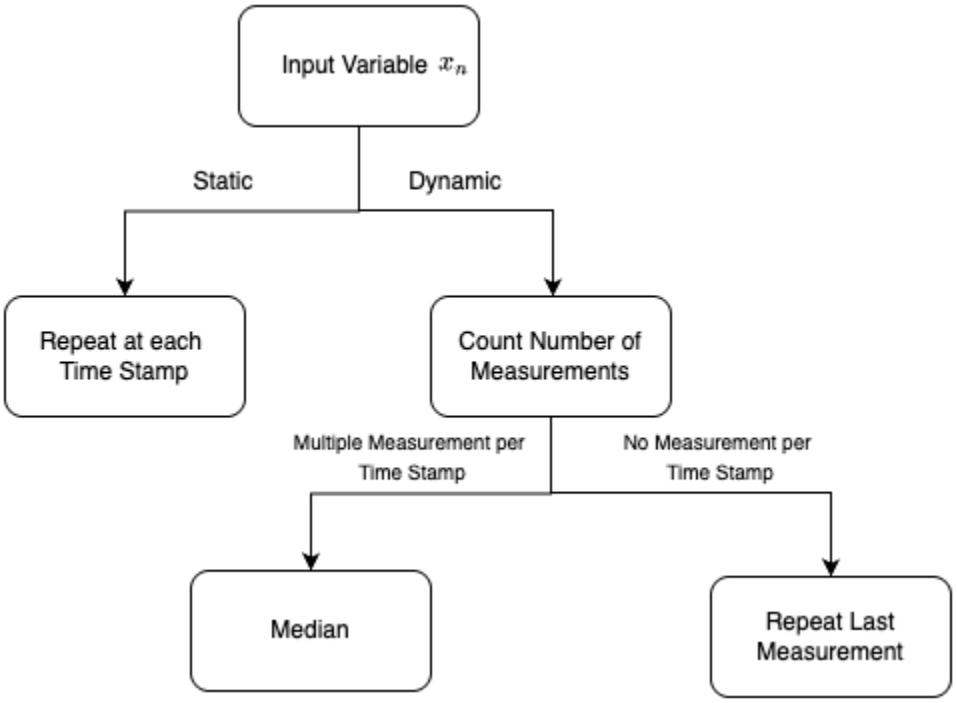
This figure illustrates how multivariate time-series data are used to make one-step-ahead fall risk predictions. Each variable (e.g., vital signs, labs) is aggregated into fixed time intervals from admission to discharge. The model uses all available data up to a prediction timestamp to generate fall risk predictions for the next time interval. The process repeats continuously throughout the hospital stay, enabling dynamic, real-time updates to the risk estimate.

Supplement Figure 2. Workflow for ML model training and evaluation.



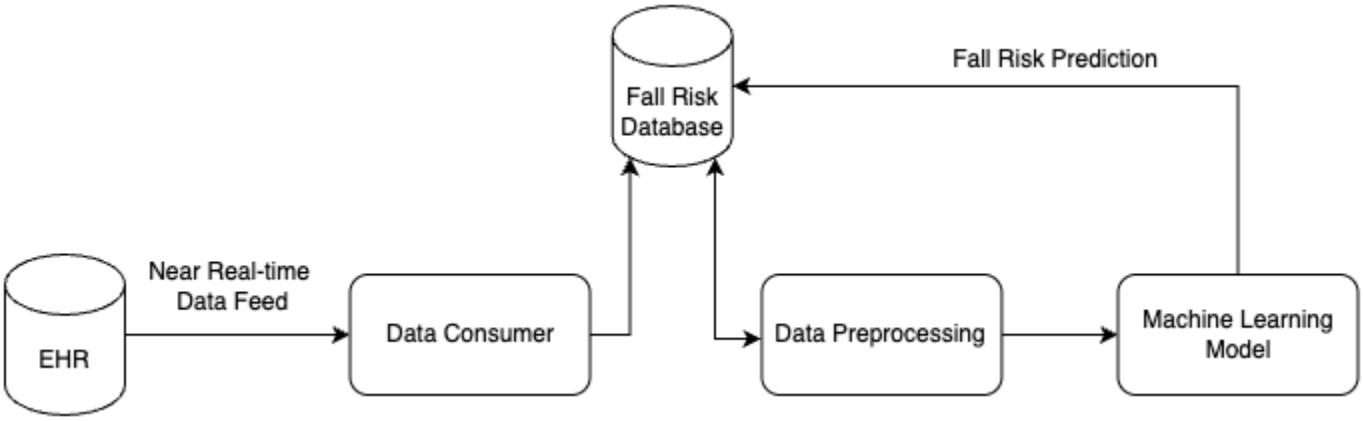
This diagram outlines the end-to-end process for developing and evaluating machine learning models. It begins with retrospective data collection and preprocessing, followed by train/test splitting. Within the training set, data are further split into training, validation, and test subsets. The model is trained with hyperparameter tuning, and performance is assessed through k-fold cross-validation. After model evaluation, interpretation tools such as SHAP are used for transparency. Finally, validated models are prepared for deployment

Supplement Figure 3. Processing strategy for static and dynamic input variables.



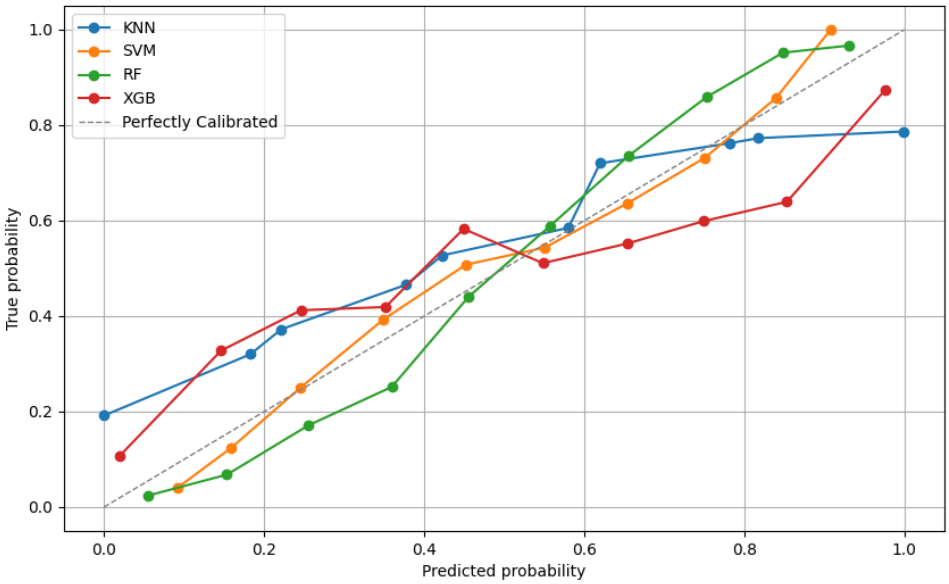
This diagram illustrates how input variables are handled based on their type. Static variables (e.g., age, sex) are repeated across all time stamps. Dynamic variables (e.g., vital signs, labs) are processed by counting the number of available measurements per time stamp. If multiple measurements exist, the median is used; if no measurement is available, the last observed value is carried forward. This standardization ensures consistent time-series input for model training.

Supplement Figure 4. System design



This figure depicts the system design for implementing near real-time fall risk prediction. Data from the electronic health record (EHR) are continuously streamed to a data consumer module, which transfers information to both a fall risk database and a data preprocessing unit. Preprocessed data are then input into a machine learning model, which generates fall risk predictions.

Supplement Figure 5. Calibration plots stratified by K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGB) machine learning models.



Supplement Table 4. Brier scores for fall risk machine learning models

Model	KNN	SVM	RF	XGB
Brier Score	0.22	0.19	0.16	0.15

Supplement Table 5. TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) check list.

Section/Topic			Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	**1**
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	**2, 3**
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	**5**
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	**6**
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	**6, 7**
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	**6, 7**
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	**6, 7**
	5b	D;V	Describe eligibility criteria for participants.	**6, 7**
	5c	D;V	Give details of treatments received, if relevant.	**n/a**
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	**6**
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	**n/a**
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	**7-9**
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	**n/a**
Sample size	8	D;V	Explain how the study size was arrived at.	**n/a**
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	**7-9**
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	**7-9**
	10c	V	For validation, describe how the predictions were calculated.	**7-9**
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	**7-9**
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	**n/a**
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	**n/a**
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	**7-9**
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	**6-9**
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	**10**
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	**10-12**
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	**10-12**
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	**n/a**
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	**10-12**
	15b	D	Explain how to the use the prediction model.	**10-12**
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	**Table 2**
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	**n/a**

Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	**15**
Interpretation	9a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	**10-12**
	9b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	**13**
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	**13-14**
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	**n/a**
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	**1**