Title

AI-Informed 30-Day MACE Risk Predictions in ED Arrivals

ISBN

9783032308122

Authors

Gross, Ben J

Patel, MD Radha

Zheng, Kai

et al.

Publication Date

2027

DOI

10.1007/978-3-032-30813-9_57

Peer reviewed

AI-Informed 30-Day MACE Risk Predictions in ED Arrivals

Ben J. Gross Ph.D.[1], Radha Patel M.D.[2], Kai Zheng Ph.D.[3], Shamim Nemati Ph.D.[1], and Mattheus Ramsis M.D.[4]

[1] UC San Diego, Division of Biomedical Informatics, San Diego, CA, USA
[2] UC San Diego, Division of General Internal Medicine, San Diego, CA, USA
[3] UC Irvine, Department of Informatics, Irvine, CA, USA
[4] UC San Diego, Division of Cardiovascular Medicine, San Diego, CA, USA
b1gross@health.ucsd.edu

Abstract. Undifferentiated chest pain is a major driver of emergency department (ED) overcrowding and contributes to morbidity and mortality. Early risk stratification for major adverse cardiac events (MACE) allows for early discharge and ED decongestion. We developed a multimodal machine learning model integrating EHR data, laboratory results, demographic and comorbidity features and chest x-ray impressions to predict 30-day MACE in patients presenting to the ED with chest pain. Using 49,348 ED encounters from two healthcare systems, our model achieved a high area under the curve (AUC) of 82.0% at the development site and 74.6% at the external validation site. Predictions occurred within 6 hours of ED arrival and demonstrated consistently high negative predictive values across time, supporting safe early-rule out. This model demonstrates the potential of integrating AI modeling to enhance early MACE risk stratification and improve ED overcrowding and efficiency.

Keywords: MACE · Triage CDSS · Undifferentiated Chest Pain .

1 Introduction

Undifferentiated chest pain is one of the most common reasons for emergency department (ED) visits, accounting for approximately eleven million annual presentations in the United States [1]. Although fewer than 10% of these patients are ultimately diagnosed with acute coronary syndrome (ACS), the need to exclude life-threatening conditions contributes substantially to ED overcrowding, delays in care, and increased morbidity and mortality [2,3]. This creates a critical challenge for clinicians, who must rapidly distinguish low-risk from high-risk patients in a resource-constrained and time-sensitive environment.

Clinical risk scores such as HEART, TIMI, and GRACE are widely used to stratify risk for major adverse cardiac events (MACE). However, these tools are limited by reduced sensitivity and negative predictive value (NPV) at common decision thresholds, limited generalizability to heterogeneous ED populations, and an inability to incorporate dynamic clinical data such as evolving laboratory results and imaging findings [4,5,6]. As a result, they may overestimate

risk in low-risk patients and contribute to unnecessary admissions and resource utilization.

The increasing availability of electronic health record (EHR) data provides an opportunity to develop more accurate, data-driven approaches to early risk stratification. In this study, we develop and externally validate a multi-modal machine learning model that integrates structured clinical data, laboratory values, and imaging-derived features to predict 30-day MACE in ED patients with chest pain. Our goal is to improve early identification of low-risk patients and support safe discharge decisions to reduce ED overcrowding.

2 Methods

We conducted a multi-center retrospective cohort study of adult ED patients at two academic centers. The development cohort was derived from University of California San Diego Health (UCSD), including encounters from two EDs between January 2016 and December 2023. External validation was performed using data from University of California Irvine (UCI), including ED encounters from January 2023 through August 2024. Institutional review board approval with waiver of informed consent was obtained at both sites.

We included all adult patients ($\geq$18 years) presenting during the study periods. Encounters were eligible if a chest radiograph (x-ray) and at least one troponin measurement were obtained within 6 hours of ED arrival and prior to any MACE. Predictor variables included the first available values from the comprehensive metabolic panel, complete blood count, lactate, and B-type natriuretic peptide (BNP), measured prior to the prediction time. Up to two troponin measurements were included if both occurred before the prediction time, defined as the time of both troponin measurement and chest x-ray availability.

The primary outcome was 30-day MACE, defined using surgical ICD codes indicating events such as acute myocardial infarction, heart failure, or stroke occurring within 720 hours of ED arrival, including subsequent readmissions.

The final cohort included 30,614 encounters at UCSD and 18,734 at UCI. MACE occurred in 240 (0.8%) encounters at UCSD and 479 (2.6%) at UCI. Cohort characteristics stratified by site and outcome are presented in Table 1.

We also track whether patients are given MACE related IV medications during the encounter itself or within 30 days of the encounter start time. These intravenous medications include Heparin, Lasix (Furosemide), and Bumex (Bumetanide). The length of stay in the ED (ED LOS) is measured for all encounters in our cohorts.

In addition to the aforementioned EHR model input features, we also input the gender, race, ethnicity, and age of the patient at the encounter start time. We also note the presence or absence of 15 MACE-related comorbidities from the problem list of the patient at the encounter start time, such as hypertension, diabetes, hyperlipidemia.

Lastly, the model uses the most recent chest x-ray impression available at the time of prediction to capture key MACE risk assessments not present in the

Table 1. Cohort demographics and outcomes are shown below. These are broken down by study site and the presence of 30-day MACE, measured from hospital arrival time.

	30-day MACE		No 30-day MACE	
	UCSD (n=240)	UCI (n=479)	UCSD (n=30,374)	UCI (n=18,255)
Patient Characteristics				
Age, median (years)	65.1	65.3	61.6	62.5
Age, IQR (years)	[56.7–73.5]	[56.0–74.6]	[49.9–72.7]	[47.8–75.6]
Female (%)	36.3	34.4	44.3	49.2
Black Race (%)	9.58	3.1	14.3	3.3
Asian Race (%)	7.5	21.5	7.6	20.6
White Race (%)	44.6	43.0	49.8	45.4
Hispanic Ethnicity (%)	31.3	37.0	23.0	39.8
Patient Outcomes				
ED LOS, median, (hours)	7.3	10.7	7.0	8.7
ED LOS, IQR, (hours)	[4.6–16.8]	[6.5–25.7]	[4.9–10.9]	[5.5–21.8]
Prediction time, median (hours)	3.0	2.5	2.8	2.6
Prediction time, IQR (hours)	[2.0–4.5]	[1.8–3.2]	[2.0–4.0]	[2.0–3.5]
MACE meds during encounter (%)	69.6	78.7	22.7	14.4
MACE meds within 30 days (%)	75.4	82.3	25.0	15.6

EHR. This impression is analyzed using Claude 3.5 Sonnet with instructions to provide a simple MACE risk assessment level (Low Risk, Borderline, High Risk, or Imminent) as an input for our model.

These 40 quantitative and categorical model input features are used to train an XGBoost model [7] to predict MACE within 30 days of the encounter start time. We use an 80/20 training/testing split of our UCSD encounters, thus training the model on 24,491 encounters and testing on the remaining 6,123 encounters. All 18,734 UCI encounters were kept together as an external validation set.

The modeling pipeline was implemented in a HIPAA-compliant cloud environment (AWS EC2), with secure access to Claude via the Amazon Bedrock Converse API.

3 Results

This model obtains an Area Under the Curve (AUC) of 82.0% on the UCSD testing set, and an AUC of 74.6% on the UCI external validation set (Table 2). Model sensitivity is targeted at 70% on the testing set, which determined the threshold risk score for positive model prediction. At this threshold the Positive

Table 2. Model Performance at Development Site (UCSD) and Validation Site (UCI).

Study Site	AUC	AUC-pr	Sensitivity	Specificity	PPV	NPV
UCSD	82.0%	4.8%	70.8%	77.3%	2.4%	99.7%
UCI	74.6%	9.7%	30.1%	93.3%	10.6%	98.1%

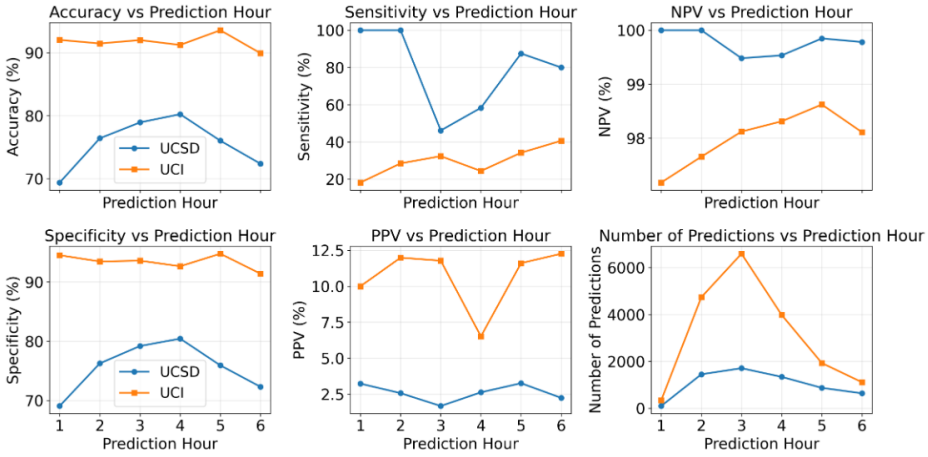

Fig. 1. Model performance by time of prediction (hours from ED arrival).

Predictive Value (PPV) is low (2.4% for the testing set, 10.6% for the validation set) but our NPV remains high, which aligns with the model's goal to allow low-risk patients to be safely discharged sooner. In the validation set, the sensitivity falls to 30.1%, suggesting a large distribution shift between these datasets. We note that the majority of model predictions occurred within 3 hours of ED arrival, enabling actionable early risk stratification during ED evaluation.

For each study site, we bin the model predictions by hours from hospital arrival to analyze how the model performance changed with the prediction time (Fig 1). Since the NPV is the most critical metric for whether our model can allow earlier patient discharge, we compute a Spearman rank correlation [8] between hourly NPV and prediction hour of the model. The correlation was not statistically significant (p=0.35 at UCSD, p=0.16 at UCI), indicating no evidence of a significant change in NPV by model prediction time.

4 Conclusion & Future Work

ED overcrowding remains a persistent issue that impacts quality of care and increases patient morbidity and mortality. Chest pain, the second most common presentation to the ED, adds substantial pressure to already strained clinical workflows and excessive patient volume. By accurately distinguishing between low- and high-risk patients, real-time risk stratification tools offer the prospect of reducing ED crowding by discharging low-risk patients sooner.

To improve our model, we would like to add more input features that would be available soon after a patient arrives in the ED. These desired features include ECG waveform data, chest x-ray image data, and clinical notes. Performance may be improved through site-specific retraining or recalibration at the validation site.

Acknowledgments. We thank Dr. Brian Kwan and Dr. Brett Meyers for providing valuable expertise in understanding MACE related workflows at UCSD Health. BJG acknowledges grant funding from the National Library of Medicine (award no. T15LM011271). SN acknowledges grant funding from the National Library of Medicine (award no. R01LM013998) and the National Institute of General Medical Sciences (award no. R35GM143121).

Disclosure of Interests. SN is a cofounder of Clairyon Inc., a predictive analytics startup. The terms of this arrangement have been reviewed and approved by the UC San Diego in accordance with its conflict-of-interest policies. All other authors declare no conflicts of interest.

References

1. Yukselen, Z., Majmundar, V., Dasari, M., Arun Kumar, P., Singh, Y.: Chest pain risk stratification in the emergency department: current perspectives. *Open Access Emerg Med* **16**, 29–43 (2024). doi:10.2147/OAEM.S419657.
2. Greenwood-Ericksen, M., Kamdar, N., Swenson, K., et al.: Emergency department boarding, inpatient census, and interhospital transfer acceptances. *JAMA Netw Open* **8**(5), e2512299 (2025). doi:10.1001/jamanetworkopen.2025.12299.
3. Moskop, J.C., Iserson, K.V., Ashton, C., et al.: Another look at the persistent moral problem of emergency department crowding. *Ann Emerg Med* **74**(3), 357–364 (2019).
4. Brown, M.D., Tomaszewski, C.A., Mekwan, J., et al.: Clinical policy: critical issues in the evaluation and management of emergency department patients with suspected non–ST-elevation acute coronary syndromes. *Ann Emerg Med* **72**(5), e65–e106 (2018).
5. Ke, J., Chen, Y., Wang, X., Wu, Z., Chen, F.: Indirect comparison of TIMI, HEART and GRACE for predicting major cardiovascular events. *BMJ Open* **11**(8), e048356 (2021). doi:10.1136/bmjopen-2020-048356.
6. Ola, O., Ball, F., Lopardo, N., et al.: Use of the HEAR score for 30-day risk-stratification in emergency department patients. *Am J Med* **136**(9), 918–926.e5 (2023).
7. Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: KDD, pp. 785–794 (2016).
8. Zar, J.H.: Spearman rank correlation. In: Encyclopedia of Biostatistics (2005).