# Using Electronic Health Record Data to Develop and Validate a Prediction Model for Adverse Outcomes on the Wards

**Matthew M Churpek, MD, MPH**[1,2], **Trevor C Yuen**[1], **Seo Young Park, PhD**[3], **Robert Gibbons, PhD**[2], and **Dana P Edelson, MD, MS**[1,*]

[1]Department of Medicine, University of Chicago, Chicago, IL

[2]Department of Health Studies, University of Chicago, Chicago, IL

[3]Department of Medicine, University of Pittsburgh, Pittsburgh, PA

## Abstract

**Objective**—Over 200,000 in-hospital cardiac arrests (CAs) occur in the U.S. each year and many of these events may be preventable. Current vital sign-based risk scores for ward patients have demonstrated limited accuracy, which leads to missed opportunities to identify those patients most likely to suffer CA and inefficient resource utilization. We derived and validated a prediction model for CA while treating intensive care unit (ICU) transfer as a competing risk using electronic health record (EHR) data.

**Design**—A retrospective cohort study.

**Setting**—An academic medical center in the United States with approximately 500 inpatient beds.

**Patients**—Adult patients hospitalized from November 2008 until August 2011 who had documented ward vital signs.

**Interventions**—None.

**Measurements and Main Results**—Vital sign, demographic, location, and laboratory data were extracted from the EHR and investigated as potential predictor variables. A person-time multinomial logistic regression model was used to simultaneously predict CA and ICU transfer. The prediction model was compared to the VitalPAC™ Early Warning Score (ViEWS) using the area under the receiver operating characteristic curve (AUC) and was validated using three-fold cross validation. A total of 56,649 controls, 109 CA patients, and 2,543 ICU transfers were included. The derived model more accurately detected CA (AUC 0.88 vs. 0.78; P<0.001) and ICU transfer (AUC 0.77 vs. 0.73; P<0.001) than the ViEWS, and accuracy was similar with cross-

validation. At a specificity of 93%, our model had a higher sensitivity than the ViEWS for CA patients (65% vs. 41%).

**Conclusions—**We developed and validated a prediction tool for ward patients that can simultaneously predict the risk of CA and ICU transfer. Our model was more accurate than the ViEWS and could be implemented in the EHR to alert caregivers with real-time information regarding patient deterioration.

## Keywords

hospital rapid response team; in-hospital cardiac arrest; physiologic monitoring; quality improvement; track and trigger

## INTRODUCTION

Patients who experience physiologic deterioration on the wards and who are unrecognized by the medical staff can progress to suffer cardiac arrest. Previous studies suggest that many of these events may be preventable and that patients often show signs of deterioration, such as abnormal vital signs, hours before the arrest (1–3). Because the majority of the 200,000 cardiac arrests that occur in the United States each year happen outside of intensive care settings (4), a risk stratification tool that accurately detects these patients before the event could have a major impact on patient safety in the hospital setting. In 2007, the National Institute for Health and Clinical Excellence recommended that physiological track and trigger systems be used to monitor all adult patients in the acute hospital setting (5). In addition, there has been a movement in several countries to create a single unified risk score to be used in all hospitals (6, 7). Although many risk scores for ward patients have been published, most were created subjectively in individual hospitals and have shown variable accuracy in previous studies (8, 9). Therefore, a statistically developed tool may overcome the low sensitivity and high false positive rates that result from using risk scores that are not evidence-based. In addition, the rapid growth of electronic health record (EHR) systems allows the possibility of adding other variables to prediction models in addition to vital signs, such as laboratory values, and could provide a method for calculating risk scores for ward patients in real time. Finally, the information relating to predictor variables in the model could be emphasized in teaching ward staff about deteriorating patients.

Importantly, those patients who become critically ill on the wards and are recognized by their caregivers are sometimes transferred to the intensive care unit (ICU). Thus, for patients experiencing physiologic decline on the wards, ICU transfer represents a competing risk for the outcome of cardiac arrest because these patients are recognized and triaged to the ICU before an arrest can occur. Although ICU transfer is sometimes used as an outcome for risk score studies (10, 11), these patients are already recognized by clinical staff and thus may have different characteristics and physiologic aberrations than those who first suffer a ward cardiac arrest. In addition, criteria for admission to the ICU vary across institutions making this outcome less generalizable. Therefore, a prediction model that can detect cardiac arrest while accounting for the competing risk of ICU transfer is needed.

The aim of this study was to develop and validate a prediction model to detect cardiac arrest using vital signs, laboratory values, and other patient characteristics while simultaneously modeling ICU transfer as a competing risk.

## MATERIALS AND METHODS

### Setting and Study Population

We conducted a cohort study using all adult patients hospitalized on the wards from November 2008 until August 2011 at an academic medical center with approximately 500 beds. A rapid response team (RRT) led by a critical care nurse and respiratory therapist has been in place at our hospital since 2008, but no specific vital sign criteria are used for their activation. Consultation from a hospitalist attending physician and/or pharmacist is available upon request, and the RRT is separate from the team that responds to a cardiac arrest.

The study protocol was approved by the University of Chicago Institutional Review Board, and a waiver of consent was granted based on minimal harm and general impracticability (IRB #16995A).

### Data Sources

All routinely collected vital signs (respiratory rate, blood pressure, heart rate, oxygen saturation, use of supplemental oxygen, temperature, and mental status) and laboratory values (white cell count, hemoglobin, platelets, sodium, potassium, chloride, bicarbonate, anion gap, blood urea nitrogen, and glucose) were collected from the EHR (EPIC; Verona, WI) to be used as potential predictor variables. Patient age, whether they had previously been in an ICU during the current hospitalization, and other demographic characteristics were collected from administrative databases. Mental status was collapsed from four drop-down menu fields in the EHR (orientation, level of consciousness, motor response, and responsiveness) into one score (alert, responsive to voice, responsive to pain, and unresponsive (AVPU)) for each observation as previously described (12). Only vital signs measured on the wards were used in the study.

### Outcomes

The primary outcome of interest was cardiac arrest, defined as a loss of pulse with attempted resuscitation, on the wards. This event was identified using a prospectively validated quality improvement database that has been described previously (12, 13). ICU transfer was identified using the hospital's admission-transfer-discharge database. Ward to ICU transfers that occurred directly from the operating room following surgery were not counted because they were not unexpected events. These patients were coded as having neither cardiac arrest nor ICU transfer at that time and were continued to be followed throughout the rest of their hospital stay, thus allowing them to have an event at a later time. Only the index cardiac arrest or ICU transfer for each patient was used in the study. For patients who suffered both a ward cardiac arrest and ICU transfer during their hospitalization, only data during their stay on the wards prior to cardiac arrest were used. Patients who were transferred to the ICU immediately after suffering a cardiac arrest were only counted as cardiac arrest patients to make the groups mutually exclusive. Patients who died on the wards without any attempt to resuscitate were omitted from the analysis as these likely represented comfort care deaths that would not benefit from identification.

### Statistical Analysis

Patient characteristics were compared using t-tests, Wilcoxon rank sum tests, and chi-squared tests where appropriate. A person-time (i.e. discrete time) multinomial logistic regression model (14–16) was used to predict cardiac arrest and the competing risk of ICU transfer simultaneously. This survival analysis technique involves separating time into discrete periods where each patient contributes a record for each period that they remained on the wards. It has been shown to provide similar results to a Cox proportional hazards model, especially when the time intervals are short (14, 15). The advantages of using this

approach include the ability to easily model competing risks, time-varying covariates, and non-proportional hazards. In addition, it utilizes the same number of observations for each patient per time period, thus removing the bias that might occur if sicker patients have more frequent vital sign observations. We separated the data into eight-hour time periods, but performed a sensitivity analysis using four-hour intervals to ensure that the model coefficients were not sensitive to the time period choice. The vital signs and laboratory values measured closest to the beginning of each time period were used for that period, and a normal (median) value was imputed if the patient did not have any previous values of a particular variable.

Potential predictor variables included in the study were age, whether or not the patient had a previous ICU admission, vital signs, and laboratory values. Time from first ward vital sign entered the model as a linear term, and a sensitivity analysis was performed by entering higher degree polynomial terms to determine if they altered the predictor variable coefficients significantly. All continuous predictor variables were modeled linearly. Backward stepwise variable selection was used, starting with a full model with all predictors, and a p-value of <0.05 for the cardiac arrest portion of the multinomial model was used for variable selection to decrease the chance of overfitting. The linear combination of the regression coefficients was linearly transformed to create a positive score by adding twenty and then multiplying by four for ease of presentation. Change in the score over time was graphed using lowess smoother curves for the 48 hours before cardiac arrest, ICU transfer, and for a random 48 hours in controls.

Model accuracy was determined in two different ways using the full vital sign dataset. First, each patient's highest predicted probability of cardiac arrest or ICU transfer during the entire hospitalization was calculated using the model coefficients from the cardiac arrest and ICU transfer portions (i.e. cardiac arrest and ICU transfer scores) of the multinomial model. These results were used to calculate the sensitivity, specificity, and area under the receiver operating characteristic curve (AUC) for each outcome. Second, these results were recalculated by using whether a cardiac arrest or ICU transfer occurred within 24 hours of each measurement time as the definition of having the event of interest. Both of these definitions have been used in previous studies investigating the accuracy of ward risk scores (3, 8, 9, 17, 18).

Using the variables identified as described above, the model was internally validated using three-fold cross validation, which involved splitting the data into three equal parts, estimating the model coefficients on two-thirds of the data and then evaluating its accuracy in the other third. This results in a less biased estimate of validation accuracy than split-sample validation (19). The cross-validation procedure was repeated 100 times, each time randomly splitting the dataset into three parts. This resulted in 300 AUCs, which were used to determine the model validation accuracy and its uncertainty. The final model was also compared to the VitalPAC™ Early Warning Score (ViEWS), the most accurate risk score from a recent comparison study (17), for detecting both cardiac arrest and ICU transfer using the AUC.

All tests of significance used a 2-sided P<0.05. Statistical analyses were completed using R (R Foundation for Statistical Computing; Vienna, Austria) and Stata version 12.0 (StataCorp; College Station, Texas).

## RESULTS

A total of 56,649 controls, 109 cardiac arrest patients, and 2,543 ICU transfer patients were included in the study (Table 1). There were 564 RRT calls over the study period, yielding an

average of 9.51 calls per 1,000 admissions. Compared to patients who did not experience a ward cardiac arrest or ICU transfer, those who suffered a cardiac arrest were older (mean age 64±17 vs. 54±18, p<0.001), were more likely to be black (62.4% vs. 43.4%, p=0.001), had a longer length of stay (median 11 [IQR 4–25] vs. 3 [1–5], p<0.001), and had a lower survival to discharge rate (27.5% vs. 99.9%, p<0.001). Cardiac arrest patients were also more likely to have an RRT call during the study period (8.3 vs. 0.3%, p<0.001) than control patients. Patients who experienced a ward to ICU transfer were older (mean age 60±16 vs. 54±18, p<0.001), more likely to be white (60.3% vs. 35.7%, p<0.001), and more likely to be a surgical patient (43.0 vs. 33.8%, p<0.001) when compared to those who suffered neither event. Compared to controls, ICU patients also had a longer length of stay (median 12 [IQR 7–21] vs. 3 [1–5], p<0.001), were more likely to have an RRT call during the study period (11.3 vs. 0.3, p<0.001) and had a lower survival to discharge rate (81.2% vs. 99.9%, p<0.001).

Our derived model contained age, prior ICU admission, respiratory rate, diastolic blood pressure, heart rate, oxygen saturation, supplemental oxygen use, mental status, temperature, hemoglobin, platelet count, white blood cell count, blood urea nitrogen, potassium, anion gap, and time, and the model regression coefficients are shown in Table 2. The percentage of values imputed ranged from 0.2 to 1.9% for vital signs and from 9.6% to 14.1% for laboratory results. Changing the time intervals to every four hours resulted in minimal (<5%) change to the regression coefficients (data not shown). The derived model more accurately detected cardiac arrest (AUC 0.88 vs. 0.78; P<0.001) and ICU transfer (AUC 0.77 vs. 0.73; P<0.001) than the ViEWS when using each patient's highest score, and results were similar when using whether the event occurred within 24 hours as the definition for model accuracy (Table 3). Model validation results from the 300 cross-validation samples for cardiac arrest and ICU transfer were similar to the AUC calculated using the whole sample (AUC 0.87 (IQR 0.86–0.88) for cardiac arrest and 0.76 (IQR 0.75–0.76) for ICU transfer). Table 4 shows the sensitivity and specificity of the derived model and ViEWS at different cut-points using each patient's highest score. At a specificity of 93%, our model had a higher sensitivity than the ViEWS for cardiac arrest patients (65% vs. 41%). The derived model had a specificity of 95% at the cut-off with 60% sensitivity, while the ViEWS had a specificity of 85% to achieve this level of sensitivity. Empirical density plots of the linearly transformed regression model scores for observations within 24 hours of a cardiac arrest and those not within 24 hours of the event are shown in Figure 1. At 48 hours before each event, the mean cardiac arrest score was higher for those who suffered a cardiac arrest (48.4) compared to those who experienced an ICU transfer (44.6) and those who experienced neither event (40.8). In addition, scores for the cardiac arrest and ICU transfer patients increased leading up to the event, while scores for controls decreased (Figure 2).

## DISCUSSION

We have developed a validated prediction model for ward patients that utilizes vital signs, laboratory results, and patient characteristics to simultaneously predict cardiac arrest and ICU transfer. The model was more accurate than the ViEWS for detecting both outcomes, and model validation accuracy was similar to the accuracy calculated using the entire dataset. This prediction model uses readily available data and could be implemented in the EHR to detect critically ill ward patients in real-time. Patients at highest risk of cardiac arrest could then be targeted for interventions such as an automated call to the RRT or a critical care consult.

Our separation of cardiac arrest and ICU transfer patients in the model development process is novel and an important strength of the study. Most previous studies of early warning scores have either studied one outcome, most often ICU transfer or mortality, or have

combined all outcomes into one group (10, 11, 17, 20). However, patients who suffer a ward cardiac arrest are unique because most of them were either inappropriately triaged or had physiologic decline not identified by their caregivers. This may be because certain types of physiologic abnormalities are more likely to be missed by medical staff. Comparing the regression coefficients between the cardiac arrest and ICU transfer portions of our developed model provides insight into some of the differences between these groups. For example, lower temperature was a significant predictor of cardiac arrest but not ICU transfer, while hypoxia was significantly associated with ICU transfer but not cardiac arrest. It may be that caregivers are more apt to recognize patients who have certain abnormalities (e.g. hypoxia) and so these patients are appropriately transferred to the ICU, while patients with evidence of other types of physiologic deterioration (e.g. hypothermia) are not recognized to the same degree. Importantly, because of these differences, if the goal is to develop a clinical decision support tool to identify those patients at highest risk of being missed by the current safety systems in the hospital, then creating a model to detect cardiac arrest is the optimal target. Modeling the risk of ICU transfer generates a model of recognized critical illness, and modeling them simultaneously provides important information regarding the differences between these patient populations and a means to calculate the likelihood of a ward patient experiencing either event.

Including laboratory values in the model contributes important knowledge to the field, as previous research has focused mostly on vital signs. We found that blood urea nitrogen, anion gap, potassium, hemoglobin, white blood cell count, and platelet count were all significant predictors in the final model. This is similar to the findings of a retrospective case-control study by investigators from Kaiser Permanente, who developed a model for the combined outcome of ICU transfer or death on the wards if the patient was full code (21). Their final model, which included blood urea nitrogen, hematocrit, white blood cell count, and a proxy for measured lactate (anion gap divided by serum bicarbonate), had an AUC of 0.78 for the combined outcome. A limitation of their study is that they included longitudinal chronic illness data in their model, which are not readily available in many hospitals, making their model less generalizable. The laboratory values in our model are also similar to other published criteria, such as the Acute Physiology and Chronic Health Evaluation (APACHE) II score, which includes potassium, creatinine, white blood cell count, and hematocrit (22). In addition, the Pneumonia Severity Index includes hematocrit and blood urea nitrogen (23), and Ranson's criteria includes these measures plus white blood cell count (24). Importantly, we only included laboratory results that are routinely collected because other variables such as lactate, although likely predictive of adverse events, are usually collected because the caregivers are already concerned about the patient. In addition, these variables have large amounts of missing values.

In this study, we compared our model to the ViEWS, a recently described early warning score by Prytherch and colleagues (17). We chose this score as a comparison because it was the best performing early warning score in a recent study comparing it to 33 other systems. In addition, an abbreviated form of the ViEWS, without mental status, has been externally validated and has been shown to be accurate for detecting mortality in recent studies.(25) Our newly developed system had a sensitivity that was 24% higher than the ViEWS at the same specificity (93%). In addition, in order to achieve the same sensitivity (60%) as our model, the ViEWS cut-off had a 10% lower specificity (95% vs. 85%). Therefore, use of the ViEWS would result in over 5500 more "false alarms" for patients who were not going to experience either a cardiac arrest or ICU transfer in our dataset. This large improvement is particularly important in light of the limited resources available for RRTs in many hospitals. Our model was more accurate than the ViEWS at other sensitivity and specificity cut-offs, as outlined in Table 4, which is an important consideration for hospitals with differing levels of resources available to respond to deteriorating patients.

In practice, our score could be implemented in the EHR to provide real-time feedback to clinicians regarding how likely it is that their patients will suffer a cardiac arrest. In addition, the variables most predictive of adverse events in our model could be used to educate physicians and nursing staff regarding early warning signs of impending deterioration and/or the need for ICU transfer. This is especially important with the increasing number of patient hand-offs that occur in many hospitals today, as providers often have less first-hand knowledge about the patients they are caring for. Automatic score generation by the EHR would also decrease the error rate in score calculation that occurs even with simpler vital sign-based systems (26, 27). It could also be used to send automatic notifications to physicians and the hospital's RRT, circumventing the "failure to call" problem that has been described in the literature (28, 29). Thus, the EHR provides a medium to improve both the risk stratification of patients and the notification of caregivers. A hospital could then individualize their response to ward patients based on available resources, including calls to the RRT, critical care consults, and increased frequency of monitoring.

Our study has several strengths compared to previous work. First, our cohort of patients was large enough to develop a model specifically for cardiac arrest. Because cardiac arrest events are measured and verified prospectively in our hospital we can be confident that all the included patients met the standard definition of cardiac arrest. In addition, we performed cross-validation of our model, which provides a more efficient estimate of model accuracy compared to the commonly used split sample validation (19). Importantly, our competing risk model is a novel way to develop risk scores for ward patients and used readily available data. Finally, the inclusion of laboratory results, previous location, and age demonstrates the potential of integrating information found in the EHR with vital sign data. It is likely that additional information in the EHR, such as vital sign and laboratory response to various interventions on the wards (e.g. fluid boluses), will further improve the accuracy of risk models and we plan to perform future studies investigating these variables.

There are several limitations to our study. First, this is a single-center study at an academic center so results may not be generalizable to other settings. For example, our patient population, use of hospitalists as part of the RRT, and lack of specific RRT calling criteria differ from some hospitals, and thus, our results need to be validated in future studies. We are currently in the process of collecting data from other hospitals to externally validate our results. Second, manual chart review of the over 2000 ICU transfers was not feasible, so we cannot know with certainty that all the transfers were unexpected. However, we excluded as ICU transfers those who underwent surgery prior to ICU admission, thus removing a significant number of those who were likely expected admissions. Third, mental status was not explicitly documented as AVPU in our hospital and so errors in translation could occur. However, we coded mental status using language in our EHR that directly mapped to the AVPU terminology and did so prior to any analyses. In addition, if errors did occur, they would have affected our model as well as the ViEWS, as AVPU is included in both algorithms. Fourth, we excluded those patients who died on the wards without a cardiac arrest or ICU transfer from our study because they were most likely comfort care deaths. Due to the size of our study we were unable to confirm that this was the case in all patients. However, we performed a sensitivity analysis by adding them back in as having neither a cardiac arrest nor an ICU transfer and the results did not change significantly. Finally, the relatively large potential variable to outcome ratio meant that we had to model all continuous predictors linearly and could not investigate the predictive value of trends in the variables over time. It is likely that accuracy could be improved with the addition of these features, but this was not feasible in this study due to concerns of overfitting the model.

## CONCLUSIONS

In conclusion, we have developed a validated prediction model that identifies critically ill patients on the wards using readily available vital sign, demographic, and laboratory data. This model has excellent accuracy for detecting cardiac arrest and can simultaneously predict a patient's risk of ICU transfer. Our model, once externally validated, could be implemented in hospitals with EHR systems to identify at risk patients in real time, triggering alerts to physicians and the RRT to intervene and potentially improving patient outcomes.

## Acknowledgments

## REFERENCES

1. Berlot G, Pangher A, Petrucci L, Bussani R, Lucangelo U. Anticipating events of in-hospital cardiac arrest. Eur J Emerg Med. 2004; 11(1):24–28. [PubMed: 15167189]

2. Hodgetts TJ, Kenward G, Vlachonikolis IG, Payne S, Castle N. The identification of risk factors for cardiac arrest and formulation of activation criteria to alert a medical emergency team. Resuscitation. 2002; 54(2):125–131. [PubMed: 12161291]

3. Cretikos M, Chen J, Hillman K, Bellomo R, Finfer S, Flabouris A, et al. The objective medical emergency team activation criteria: a case-control study. Resuscitation. 2007; 73(1):62–72. [PubMed: 17241732]

4. Merchant RM, Yang L, Becker LB, Berg RA, Nadkarni V, Nichol G, et al. Incidence of treated cardiac arrest in hospitalized patients in the United States. Crit Care Med. 2011; 39:2401–2406. [PubMed: 21705896]

5. Excellence: NIfHaC. NICE clinical guideline No. 50. London: 2007. Acutely ill patients in hospital: recognition of and response to acute illness in adults in hospital.

6. Jenkins PF, Thompson CH, Barton LL, Jenkins PF, Thompson CH, Barton LL. Clinical deterioration in the condition of patients with acute medical illness in Australian hospitals: improving detection and response. Medical Journal of Australia. 2011; 194(11):596–598. [PubMed: 21644875]

7. Subbe CP. Better ViEWS ahead? It is high time to improve patient safety by standardizing Early Warning Scores. Resuscitation. 2010; 81(8):923–924. [PubMed: 20637973]

8. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI. Review and performance evaluation of aggregate weighted 'track and trigger' systems. Resuscitation. 2008; 77(2):170–179. [PubMed: 18249483]

9. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI, Higgins B. A review, and performance evaluation, of single-parameter "track and trigger" systems. Resuscitation. 2008; 79(1):11–21. [PubMed: 18620794]

10. Cuthbertson BH, Boroujerdi M, McKie L, Aucott L, Prescott G. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? Crit Care Med. 2007; 35(2):402–409. [PubMed: 17205002]

11. Cuthbertson BH, Boroujerdi M, Prescott G. The use of combined physiological parameters in the early recognition of the deteriorating acute medical patient. J R Coll Physicians Edinb. 2010; 40(1):19–25. [PubMed: 21125034]

12. Churpek MM, Yuen TC, Huber MT, Park SY, Hall JB, Edelson DP. Predicting cardiac arrest on the wards: a nested case-control study. Chest. 2012; 141(5):1170–1176. [PubMed: 22052772]

13. Edelson DP, Litzinger B, Arora V, Walsh D, Kim S, Lauderdale DS, et al. Improving in-hospital cardiac arrest process and outcomes with performance debriefing. Arch Intern Med. 2008; 168(10):1063–1069. [PubMed: 18504334]

14. Efron B. Logistic-Regression, Survival Analysis, and the Kaplan-Meier Curve. Journal of the American Statistical Association. 1988; 83(402):414–425.

15. Ingram DD, Kleinman JC. Empirical Comparisons of Proportional Hazards and Logistic-Regression Models. Statistics in Medicine. 1989; 8(5):525–538. [PubMed: 2727473]

16. Gibbons RD, Duan N, Meltzer D, Pope A, Penhoet ED, Dubler NN, et al. Waiting for organ transplantation: results of an analysis by an Institute of Medicine Committee. Biostatistics. 2003; 4(2):207–222. [PubMed: 12925517]

17. Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS--Towards a national early warning score for detecting adult inpatient deterioration. Resuscitation. 2010; 81(8):932–937. [PubMed: 20637974]

18. Churpek MM, Yuen TC, Park SY, Meltzer DO, Hall JB, Edelson DP. Derivation of a cardiac arrest prediction model using ward vital signs*. Crit Care Med. 2012; 40(7):2102–2108. [PubMed: 22584764]

19. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol. 2001; 54(8):774–781. [PubMed: 11470385]

20. Bleyer AJ, Vidya S, Russell GB, Jones CM, Sujata L, Daeihagh P, et al. Longitudinal analysis of one million vital signs in patients in an academic medical center. Resuscitation. 2011; 82(11): 1387–1392. [PubMed: 21756971]

21. Escobar GJ, LaGuardia JC, Turk BJ, Ragins A, Kipnis P, Draper D. Early detection of impending physiologic deterioration among patients who are not in intensive care: development of predictive models using data from an automated electronic medical record. J Hosp Med. 2012; 7(5):388–395. [PubMed: 22447632]

22. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med. 1985; 13(10):818–829. [PubMed: 3928249]

23. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. N Engl J Med. 1997; 336(4):243–250. [PubMed: 8995086]

24. Ranson JH, Rifkind KM, Roses DF, Fink SD, Eng K, Spencer FC. Prognostic signs and the role of operative management in acute pancreatitis. Surg Gynecol Obstet. 1974; 139(1):69–81. [PubMed: 4834279]

25. Kellett J, Kim A. Validation of an abbreviated Vitalpac Early Warning Score (ViEWS) in 75,419 consecutive admissions to a Canadian regional hospital. Resuscitation. 2012; 83(3):297–302. [PubMed: 21907689]

26. Prytherch DR, Smith GB, Schmidt P, Featherstone PI, Stewart K, Knight D, et al. Calculating early warning scores--a classroom comparison of pen and paper and hand-held computer methods. Resuscitation. 2006; 70(2):173–178. [PubMed: 16806641]

27. Subbe CP, Gao H, Harrison DA. Reproducibility of physiological track-and-trigger warning systems for identifying at-risk patients on the ward. Intensive Care Med. 2007; 33(4):619–624. [PubMed: 17235508]

28. Hillman K, Chen J, Cretikos M, Bellomo R, Brown D, Doig G, et al. Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. Lancet. 2005; 365(9477): 2091–2097. [PubMed: 15964445]

29. Marshall SD, Kitto S, Shearer W, Wilson SJ, Finnigan MA, Sturgess T, et al. Why don't hospital staff activate the rapid response system (RRS)? How frequently is it needed and can the process be improved? Implement Sci. 2011; 6:39. [PubMed: 21496276]
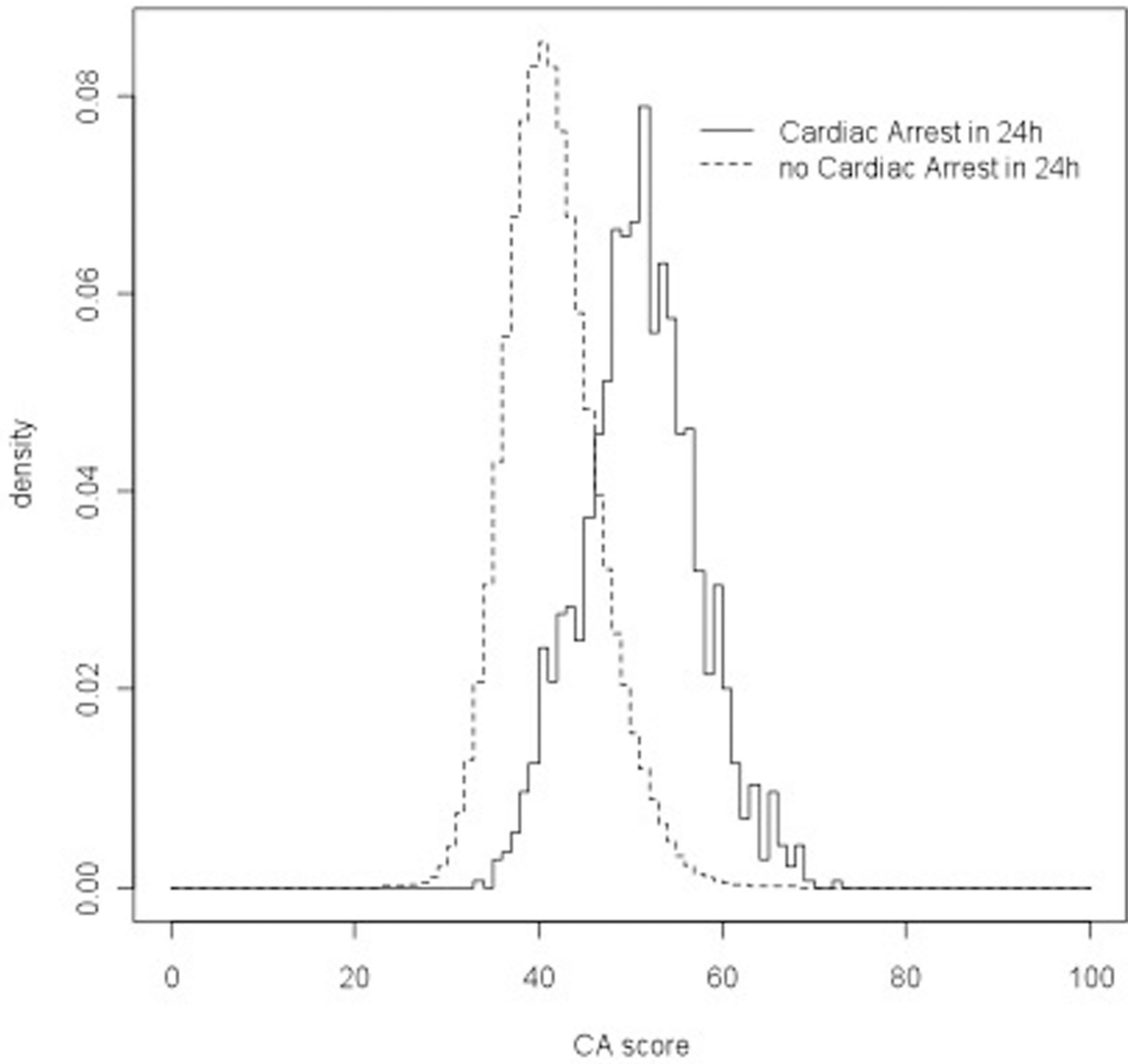
**Figure 1.**
Empirical density plots of the cardiac arrest score for cardiac arrest and for patients who experienced neither a cardiac arrest nor an ICU transfer.
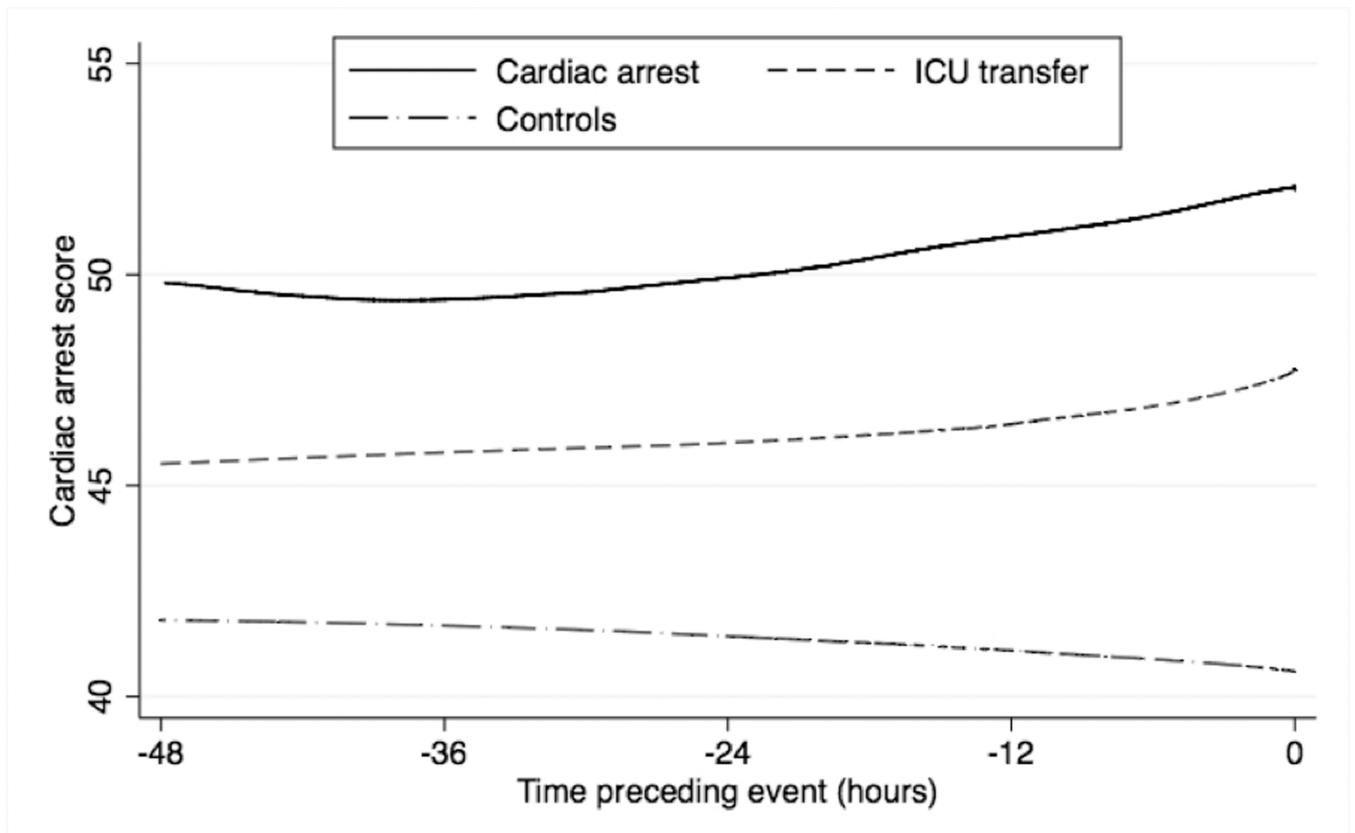
**Figure 2.**
Lowess smoother curves of the average cardiac arrest score within 48 hours of cardiac arrest or ICU transfer. The average score over a random 48 hours is shown for controls.

**Table 1**

Patient Characteristics.

| | Controls (N=56,649) | ICU Transfer (N=2,543) | Cardiac Arrest (N=109) |
|---|---|---|---|
| Age, mean (SD), years | 54 ± 18 [b, c] | 60 ± 16 [a, c] | 64 ± 17 [a, b] |
| Female sex | 32,149 (56.8) [b] | 1,270 (49.9) [a] | 64 (58.7) |
| Race | [b, c] | [a, c] | [a, b] |
| Black | 24,602 (43.4) | 1,094 (43.0) | 68 (62.4) |
| White | 20,206 (35.7) | 1,026 (40.3) | 24 (22.0) |
| Other | 1,456 (2.6) | 92 (3.6) | 3 (2.8) |
| Unknown | 10,385 (18.3) | 331 (13.0) | 14 (12.8) |
| Surgical patient | 19,123 (33.8) [b] | 1,094 (43.0) [a, c] | 36 (33.0) [b] |
| Length of study period, median (IQR), hours | 53 (26 – 114) [b] | 32 (8 – 92) [a, c] | 53 (24 – 132) [b] |
| Hospital length of stay, median (IQR), days | 3 (1 – 5) [b, c] | 12 (7 – 21) [a] | 11 (4 – 25) [a] |
| At least one RRT call during study period | 149 (0.3) [b, c] | 288 (11.3) [a] | 9 (8.3) [a] |
| Survived to discharge | 56,602 (99.9) [b, c] | 2,065 (81.2) [a, c] | 30 (27.5) [a, b] |

[a] $P<0.05$ compared to controls.

[b] $P<0.05$ compared to ICU transfers.

[c] $P<0.05$ compared to cardiac arrests.

**Table 2**

Regression coefficients and p-values for variables in the final multinomial regression model.[a]

| Variable | Cardiac arrest | | ICU transfer | |
|---|---|---|---|---|
| | **Estimate** | **P-value** | **Estimate** | **P-value** |
| Time (hours) | 0.00 | 0.41 | 0.00 | 0.99 |
| Prior ICU stay (1=Yes, 0=No) | 1.37 | <0.001 | 0.12 | 0.02 |
| Heart rate (beats/min) | 0.03 | <0.001 | 0.04 | <0.001 |
| Diastolic blood pressure (mm Hg) | −0.02 | 0.01 | −0.01 | <0.001 |
| Respiratory rate (breaths/min) | 0.14 | <0.001 | 0.14 | <0.001 |
| Oxygen saturation (%) | 0.07 | 0.14 | −0.05 | <0.001 |
| Temperature (°C) | −0.31 | 0.03 | −0.01 | 0.87 |
| Mental status (AVPU) | 0.43 | 0.27 | 1.16 | <0.001 |
| On room air (1=Yes, 0=No) | −0.64 | 0.02 | −0.32 | <0.001 |
| Age (years) | 0.03 | <0.001 | 0.02 | <0.001 |
| BUN (mg/dL) | 0.01 | 0.005 | 0.01 | <0.001 |
| Anion gap (mEq/L) | 0.13 | <0.001 | 0.07 | <0.001 |
| Hemoglobin (g/dL) | −0.17 | 0.007 | −0.01 | 0.37 |
| Platelet count (K/uL) | −0.002 | 0.007 | −0.001 | <0.001 |
| Potassium (mEq/L) | 0.17 | 0.31 | 0.13 | <0.001 |
| White blood cell count (K/uL) | 0.01 | 0.01 | 0.01 | <0.001 |

[a]Predictors modeled as continuous variables unless otherwise noted.

Abbreviations: AVPU, Alert, Responsive to Voice, Responsive to Pain, Unresponsive; BUN, Blood Urea Nitrogen

**Table 3**

Areas under the receiver operating characteristic curves (AUCs) for cardiac arrest or ICU transfer compared to those who experienced neither outcome for the derived model and ViEWS.

| Outcome | Score | | |
|---|---|---|---|
| **Ever experienced event** | **Cardiac arrest model**[a] | **ICU transfer model**[a] | **ViEWS** |
| Cardiac arrest | 0.88 (0.84–0.91) | - | 0.78 (0.73–0.83) |
| ICU transfer | - | 0.77 (0.76–0.78) | 0.73 (0.72–0.74) |
| Experienced event within 24 hours | | | |
| Cardiac arrest | 0.88 (0.88–0.89) | - | 0.74 (0.72–0.75) |
| ICU transfer | - | 0.76 (0.76–0.76) | 0.73 (0.72–0.73) |

[a]Score for cardiac arrest and ICU models were derived from the cardiac arrest and ICU transfer portions of the multinomial logistic regression model.

**Table 4**

Sensitivity and specificity for identifying patients who suffered cardiac arrest at different cut-points for the derived model and ViEWS.

|  | Sensitivity (%) | Specificity (%) |
|---|---|---|
| Derived model |  |  |
| > 49 | 81 (72–88) | 82 (82–83) |
| > 50 | 77 (68–85) | 86 (86–86) |
| > 51 | 71 (61–79) | 89 (89–89) |
| > 52 | 66 (57–75) | 92 (92–92) |
| > 53 | 65 (55–74) | 93 (93–93) |
| > 54 | 60 (50–69) | 95 (95–95) |
| > 55 | 56 (46–66) | 96 (96–96) |
| ViEWS |  |  |
| >7 | 72 (62–80) | 73 (73–73) |
| >8 | 60 (50–69) | 85 (85–85) |
| >9 | 41 (32–51) | 93 (93–93) |
| >10 | 29 (21–39) | 97 (97–97) |